

# My Past Dictates my Present: Relevance, Exposure, and Influence of Longitudinal Data on Facebook

Shujaat Mirza  
New York University  
shujaat.mirza@nyu.edu

Christina Pöpper  
New York University Abu Dhabi  
christina.poepper@nyu.edu

**Abstract**—Online social networks accumulate unprecedented amounts of data that continue to exist on user profiles long after the time of posting. Given that these platforms primarily provide a venue for people to connect and foster online friendships, the influence and the risks associated with longitudinal data may impact users and their reasons for using these platforms. To better understand these issues, we conducted two user studies of Facebook users analyzing the history of past postings w.r.t. to their perceived relevance, longitudinal exposure, and impact on the users’ befriending behavior. The studies give us a cross-cultural undergraduate student sample (n=89, campus study) and a Mechanical Turk sample of two cultural backgrounds from the US and India (n=209, MTurk study). Our findings reveal that a sizable group of participants consider their past postings irrelevant and, at times, embarrassing. However, participants’ awareness and usage of longitudinal privacy control features (e.g., *Limit Past Posts*) are limited, resulting in overexposure of their past postings and personal information. Importantly, we find support that these overexposed, yet irrelevant, past postings (of both participants and friend requesters) have the potential to influence users’ fundamental behavior on the platform: friend network expansion. Participants greatly valued friend requester’s past postings, particularly in the absence of prior personal interactions, but are influenced by their backgrounds (American users rely significantly more than their Indian counterparts on the requesters’ past postings for their befriending behavior). We close by discussing the implications of our findings on the future of longitudinal privacy controls.

## I. INTRODUCTION

In current Online Social Networks (OSNs), remembering, rather than forgetting, is the default option, resulting in new threats to the users’ privacy. With over 2.4 billion monthly active users, Facebook is the largest social networking service [1]. Over 300 million photos are posted to the service daily and as many as 293,000 statuses updated per minute [2]. Information posted by users is indexed and easily searchable using powerful tools, such as Facebook’s *Timeline*, with just a click of a button. Much can be inferred about users through the data that exists on their Facebook profiles [3]. Whereas the platform offers an option for users to make their profiles ‘private’, studies have demonstrated the public availability of a substantial number of user profiles [4], [5]. Even in the case of

‘private’ accounts, a selection of up to nine ‘featured photos’ is still public and visible to everyone [6].

The accumulated data on the users’ profiles is known to serve both archival and exhibition purposes; however, it is unclear if this longitudinal data could influence users’ current behavior on the platform. We narrow this knowledge gap by focusing on one such instance: the impact of past postings on participants’ befriending behavior. Both the participants’ own posts and the posts of people sending friend requests are studied. Since friend network is considered a fundamental component of the platform, studying the impact of (requester’s and recipients’) past postings on its expansion is crucial. However, the role of accumulated data cannot be contextualized fully without a detailed understanding of its relevance and longitudinal exposure; thus, this study is the first to explore these interlinked aspects together, as outlined below.

Prior works on the relevance of past postings have made orthogonal findings by focusing only on the effect of time passed since publication [7], [8]. To gain a more complete picture, we further investigate the role of the actual content of the participants’ past postings as well as the different preferences users may have towards these postings in the current context: re-sharing, availability, exposure, and deletion preferences. The detailed evaluation helps us understand whether the participants’ perceived relevance of their past postings is in line with the influence past postings have on their befriending behavior.

Previous work on longitudinal exposure of past postings has identified the difficulty users face in correctly setting multi-level Facebook privacy settings [9], [10], [11], [12], [13], [14]. However, little is yet known about the awareness and usage of the longitudinal privacy control feature, *Limit Past Posts*, that can restrict the visibility of all past postings. Since aged information may have limited relevance but a significant potential to impact users’ befriending behavior, we aim to understand how much control participants have over its longitudinal exposure.

In brief, we seek to uncover with respect to the longitudinal data on Facebook network: (1) its role in users’ befriending behavior; (2) its relevance for users’ present context; and (3) the control users have over its exposure. To date, these issues have not been adequately investigated together. Gaining further knowledge will help assess the impact of longitudinal data and contribute towards development of appropriate longitudinal privacy controls for social media platforms. Unlike prior studies on the topic of privacy on Facebook [9], [10], [11],

[12], we do not limit our scope to the US population only, but also include non-WEIRD (Western, Educated, Industrialized, Rich and Democratic) [15] populations in our sample.

Our major contributions in this work are threefold:

- 1) To the best of our knowledge, we are the first to explore the influence of past postings on the expansion of the users' friend network (RQ1). Our findings indicate that even though users mainly consider past postings irrelevant (to certain extent) to be shared in the present context, these have the potential to influence the befriending behavior (Section IV-A). We uncover that the participants' curiosity to learn new information through past postings of requesters is considerable and outweighs their fear to share their own history of past postings with the requester upon friend request acceptance. We also learn that users from different cultures associate different levels of value to the past postings in their befriending behavior.
- 2) We capture the extent of participants' perceived relevance of past postings in detail (RQ2) and reveal that the attitudes differ depending upon the actual content of the posting, with sensitive posts more likely to become irrelevant (Section IV-B).
- 3) We uncover participants' lack of awareness and usage of the network's longitudinal privacy management feature, *Limit Past Posts*, (RQ3) and demonstrate that both overexposure and underexposure of aged information occur (Section IV-C).

## II. CURRENT CONTEXT AND PAST POSTINGS

### A. Terminology

For the purpose of our study, *postings* consist of photos, textual status updates, life updates and events/check-ins. *Timeline* is where users share these postings on Facebook. *Longitudinal privacy* relates to the user's ability to have control over the postings' sharing preferences after they have been published on Facebook. *Longitudinal exposure* refers to the visibility of postings made in the past. We use the term *context* to express the temporal circumstances: Postings are considered to be made in the *current context* if they are published on Facebook in the user's *present time* (i.e., during the current day or week). In contrast, postings made in the *old context* refer to those that were posted on the platform anytime *in the past* (i.e., before the users' present time). *Audience* refers to the group of people who are able to see the posting and it can range from 'Public – anyone on or off Facebook' to 'Only me – publisher of the posting'. A posting is termed *underexposed* if its actual audience is a subset of publisher's intended audience. An *overexposed* posting is the one that is visible to a larger set of people than the publisher's intended audience. *Befriending behavior* encompasses all activities and behaviors that occur between receiving a friend request and accepting, rejecting or choosing to not respond to it.

### B. Research Questions and Hypotheses

Our work was guided by a set of research questions (**RQ<sub>x</sub>**) and hypotheses (**H<sub>x</sub>**).

In their study on the development of Facebook befriending models, Rashtian et al. [16] identified having mutual friends

and being active on the platform as factors that impact the befriending decision. Users' reliance on past postings as an investigation action to look for commonalities has not been studied in detail so far. Postings made in the past were likely made in a different context while having a divergent audience in mind. It is therefore crucial to understand if these postings could impact the friend request acceptance process. We explore the following research question in detail:

**RQ1: Impact of Longitudinal Exposure.** *How do users' postings made in the old context play a role in their befriending decisions in the current context?*

We are interested in understanding how frequently users factor in past postings of the requester before making a decision on their request. As prior work has reported that users inherently are more likely to trust people they already have associations and commonalities with [17], [16], it is reasonable to presume that users' attitude towards requests received from strangers might be different than those from acquaintances. However, what has not been studied is whether the reliance on past postings of the requester is negligible for acquaintances. We also aim to understand whether users worry about privacy concerns that arise from sharing their history of past postings with the requester upon acceptance of the request. In more detail, we evaluate the following hypotheses:

**H1A General Impact.** Postings made in the old context are taken into account by the majority of the users in their befriending decisions in the current context.

**H1B Impact of Offline Interactions.** Offline interactions minimize the users' reliance on past postings in their befriending decisions.

**H1C Impact of Requester's Postings.** Past postings made by the users themselves are not as critical for the expansion of friend network as the ones made by the requester.

As the accumulated data on Facebook grows significantly, it is important to understand the relationship between information aging and sharing preferences in order to better contextualize its impact. Ayalon and Toch [18], [7] found out that willingness to share drops with the time passed since publishing of the post and called for an expiration date for the content. In another study, Bauer et al. [8] reported that participants' predictions about how their preferences would change correlated poorly with their actual changes in preferences over time and participants found value in these posts for reminiscence. Based on these seemingly orthogonal findings, the two studies disagreed on the idea of setting expiration times for the postings. Both studies account for the effect of time alone on the relevance of past postings and do not factor in the effect of content. This motivates our second research question.

**RQ2: Comfort with Longitudinal Exposure.** *How do users' preferences for exposure of past postings on their Timelines change based on the actual content?*

We explore in detail the different aspects of relevance of longitudinal data through understanding users' visibility preferences. Re-sharing a past posting to the *Timeline* implies that the user deems it relevant enough to be highlighted in the current context. If a user's preference for a past posting is continued availability, then it is likely that the posting is deemed relevant for exhibiting or reminiscent purposes. In contrast, if the user decides to restrict exposure or delete

the posting, then it is likely to have had limited relevance or complete irrelevance. We are also interested to understand if the actual content of postings impacts users' exposure preferences. For the purpose of our study, we focus on postings concerning *personal* (relating to one's self, family, etc.) and *sensitive* (pertaining to political/religious views, etc.) content. Research shows that postings expressing controversial views can offend people and damage relationships as opposed to the ones revolving around personal issues [19]. People's religious or political beliefs also undergo revisions over time, which could also impact their visibility preferences for the past postings. We hypothesize the following:

**H2A** *General Discomfort.* Facebook users are not comfortable sharing their postings from the old context into the current context.

**H2B** *Discomfort w.r.t. Content.* Users' discomfort to share postings concerning sensitive topics is much more apparent than in the case of those containing personal content.

To make sense of the findings of the above questions, it is important to uncover users' command over their longitudinal data exposure. Thus, we study users' awareness of Facebook's longitudinal privacy features that allow control over visibility of past postings on users' *Timelines*:

**RQ3: Awareness of Longitudinal Controls.** *What is users' understanding of Facebook's privacy features for postings published in the current context and ones that were posted in the old context? Is there disparity between the users' perceived and the actual privacy settings?*

In the past, users have reportedly struggled with correctly configuring access control settings [11], [12]. With the introduction of new longitudinal privacy features (*Limit Past Posts*) and multiple revisions to the existing exposure control options ('*friends except*', '*custom settings*', '*specific friends*', etc.) [20], the task of configuring exposure settings correctly is becoming a challenge for the users. We hypothesize the following:

**H3A** *Lack of Awareness.* Facebook users' awareness of the network's privacy features is not as comprehensive for the postings made in the old context as it is for the postings made in the current context.

**H3B** *Mismatch in Exposure Settings.* Facebook users' perception of their profile's privacy settings does not match the actual settings.

### III. METHODS

#### A. Questionnaire Design & Approach

The survey questions revolved around three major themes: influence of past postings on the befriending behavior (**RQ1**), relevance of past postings based on their content (**RQ2**), and users' understanding of longitudinal privacy controls and features (**RQ3**). We list an abridged version of the questions from the survey in Appendix A.

In the first part of the study, we prompted respondents to scroll back by 3 years on their Facebook *Timelines* to identify postings concerning sensitive and personal nature (**H2A-B**). For each kind of posting, we asked the participants using a Likert scale from 1 (definitely) to 5 (definitely not) for

their preferences to keep the post available, to change its exposure settings, to re-share it in the present context, and to delete it from their profile. If no such post was found, the participants were prompted to answer an alternate set of questions inquiring the non-existence of such postings on their profile. We asked follow up questions to understand their preferences more thoroughly. A period of 3 years was chosen as it provided us with postings that were neither too recent nor very old, and was partly inspired by prior work [7].

The second stage revolved around exploring participants' behavior upon receiving friend requests from strangers and acquaintances (**H1A-C**). There are two types of postings that could impact a befriending decision: past postings of the requester and those of the recipient. Since Facebook interface does not provide a way to track the history of previously accepted or denied requests and we did not want to use automated, privacy-invasive ways to collect data of users' profiles, we made respondents roughly recall the recent instances when they received a friend request and estimate their actions, such as whether they visited the *Timelines* of the requesters to look through the past postings before making a decision on the request. We grouped their actions and provided broader categories as answer options to reduce burden of recalling a specific instance. Participants were also asked to list the types of postings that generally influence their decision positively or negatively. For aided recall, a set of possible choices, derived from findings of a study on unfriending behavior [21], were offered to the participants in addition to the open ended text box.

In the final stage, to test hypothesis **H3A**, we displayed a list of privacy features afforded to the users by Facebook. While some of those features deal with postings that are to be published in the current context, others focus on configuring privacy settings for postings made in the old context, e.g., *Limit Past Posts*. We asked users to report their awareness and usage of each of these features to analyze if their understanding is consistent across both contexts, current and old. To test hypothesis **H3B**, respondents were asked to list their perception of existing privacy settings for different information types that exist on their *Timelines*. Participants were also provided with choices such as "*I do not know*" and "*I have not posted this information*" in case they did not remember their choices. Afterwards, respondents were asked to visit their profiles and report the actual privacy settings for each of the above information types.

In order to be able to assess the quality of responses, we chose to introduce two controlled questions into the survey (see part 5 of the Appendix A). One of these asked the respondents to choose a specific option as a choice for a question. The other asked about users' usage of a non-existent feature. We discarded the entire set of responses from those participants that became victim to both of the attention questions.

#### B. Pilot Studies

To evaluate the effectiveness and clarity of the questions, we conducted two pilot studies with 10 colleagues from our academic community. In the first study, 5 respondents were handed out the paper versions of the survey and asked to provide answers using the think-aloud technique [22] while one

TABLE I. DEMOGRAPHICS OF THE PARTICIPANTS OF THE TWO STUDIES: CAMPUS STUDY (N=89) & MTURK STUDY (N=209). IN (INDIA) AND US (UNITED STATES) UNDER BACKGROUND FOR THE MTURK STUDY ARE MEANT TO FURTHER SPECIFY THE BACKGROUND.

|                    | Category      | Campus Study (89) | MTurk Study (209) |
|--------------------|---------------|-------------------|-------------------|
| <b>Age</b>         | 18 - 22       | 98% (87)          | 2% (5)            |
|                    | 23 - 27       | 2% (2)            | 34% (69)          |
|                    | 28 - 37       | -                 | 42% (87)          |
|                    | 38 - 47       | -                 | 16% (34)          |
|                    | 48 - 57       | -                 | 3% (7)            |
|                    | 58+           | -                 | 3% (7)            |
| <b>Gender</b>      | Male          | 56% (50)          | 52% (108)         |
|                    | Female        | 44% (39)          | 47% (99)          |
|                    | Not disclosed | -                 | 1% (2)            |
| <b>Back-ground</b> | Asia          | 43% (38)          | 44% (91, IN)      |
|                    | North America | 18% (16)          | 51% (107, US)     |
|                    | Europe        | 15% (13)          | 3% (6)            |
|                    | Middle East   | 12% (11)          | -                 |
|                    | Others        | 12% (11)          | 2% (5)            |

of the researchers sat next to them. The researcher used semi-structured interviews to probe the participants to gauge if their understanding of tasks was consistent with the researchers' intentions. Taking the feedback in, we designed the online survey for the next 5 respondents to be filled in the absence of a researcher to resemble the environment of the actual study. Following their feedback, the duration of study was shortened to 25 minutes to allow for focused responses.

### C. Recruitment and Demographics

The detailed demographic breakdown of the two studies is shown in Table I. For the campus study, we recruited 91 participants from our academic community. After discarding two cases void of attention, we were left with 89 sets of responses. Our university is characterized by a diverse set of nationalities and cultural backgrounds, which allowed us to recruit people who grew up in regions scattered all over the world, thus, allowing the data set to contain a broad range of views and perspectives that are likely representative of the complex user base of the service. For our MTurk study, we were left with 209 participants after discarding 32 responses that were either incomplete or lacked attention. The dominance of US and Indian workers on the MTurk platform is well documented in studies before [23].

### D. Procedure

The survey was powered by *Qualtrics* [24] service. In addition to the minimum age limit of 18 years, the other attribute required of the participants to undertake the study was to be a regular user of the Facebook service for at least 3 years.

*Campus Study:* In 2018, we hosted 6 sessions in the lab at our New York University Abu Dhabi campus, where up to 20 machines were set up for participants to fill in the survey. We paid each participant 50 AED as subsistence allowance for completing the survey, which is consistent with hourly rates of other similar academic research activities at the university. There was a show-up allowance of 10 AED in case participants withdrew consent or discontinued participation for any reason.

*MTurk Study:* In 2019, we presented our survey as a human intelligence task (HIT) on Amazon's Mechanical Turk (MTurk), a crowd-sourcing service [25]. The only changes made to the survey from the Campus study were that two additional questions on availability and exposure preferences were introduced. Participation was limited to workers who had an approval rating of at least 99% and had more than 1,000 tasks approved. Once Turkers accepted the HIT, they were redirected to the Qualtrics [24] survey. We estimated the survey to take about 25 minutes to complete and paid US \$5 to each participant. On average, participants took 18.9 minutes to finish the survey.

### E. Ethical Considerations

Each participant electronically authorized the IRB-approved consent form at the start of the study. They were informed that no data will be recorded from their profile, but only the answers they provide to the survey questions. The consent form also informed the participants about their right to withdraw consent or discontinue participation at any time, listed the duration of the study and the incentives for participation. For the campus study, we purposefully used broader categories (continents instead of country, etc.) to ensure anonymity.

### F. Data Analysis

We first performed the Shapiro-Wilk test on all dependent variables and found that the distribution was not normal in most cases. Therefore, we chose to perform non-parametric tests to compare which groups are significantly different from each other. Depending upon the type of data, these tests ranged from Mann-Whitney U test to Kruskal Wallis rank sum test to Spearman's rank correlation coefficient. The details of this analysis are reported in the next section.

## IV. RESULTS

### A. Impact of Past Postings: Case of Users' Befriending Behavior

We start by reporting results on the participants' reliance on longitudinal data in their befriending decisions, relating to **RQ1** (*Impact of Longitudinal Exposure*).

1) *Classification: Privacy-awareness vs. Indifference:* To understand the effect of past postings on befriending behavior, we considered two dimensions: the impact of past postings of the requester and those of the request recipient. Based on respondents' answers, we could categorize them into different groups. Users that are interested in looking into the requester's postings before making their decision for majority of the requests are labeled as *Curious*. Those users who are concerned at the prospect of sharing their past postings with the requester upon request acceptance for majority of the requests are labeled *Concerned*. Participants were marked for these categories if their behavior applied to the majority ( $\geq 5$ ) of the received friend requests. As shown in Table II, users' behavior can be classified into four categories. The categories that housed most participants were 1 (*Curious & Concerned* — 39%) and 4 (*Incurious & Unconcerned* — 28%). Classifying participants into groups based on the intersections of their interest in friend-requesters' past posts and their concern about

TABLE II. CLASSIFICATION OF PARTICIPANTS' ATTITUDES TO PAST POSTS DURING BEFRIENDING BEHAVIOR, BASED ON THE (MTURK, CAMPUS) STUDIES (N=298). CURIOSITY CAPTURES RECIPIENT'S INTEREST IN REQUESTER'S PAST POSTINGS AND CONCERN REFLECTS RECIPIENT'S HESITANCE IN SHARING THEIR HISTORY OF POSTINGS WITH THE REQUESTER UPON REQUEST ACCEPTANCE.)

|           | Concerned   | Unconcerned |           |
|-----------|-------------|-------------|-----------|
| Curious   | 39%(83, 34) | 19%(25, 32) | 58%(174)  |
| Incurious | 14%(37, 4)  | 28%(64, 19) | 42%(124)  |
|           | 53%(158)    | 47%(140)    | 100%(298) |

sharing their own reveals diversity in general attitudes towards past postings.

The most popular category, *Curious & Concerned*, contains participants who visit past postings of requesters for the majority of the received friend requests to derive insights for their decision-making process. At the same time, they are also concerned about sharing their own complete history of past postings with the requesters upon the acceptance of the request. Overall, 53% (158) of participants are concerned about sharing their own longitudinal data and as many as 58% (174) of participants are curious to learn from requester's past postings in their befriending behavior, providing support for **H1A** (*General Impact*) as majority of users factor in past postings in their decision-making process. That being said, we identified a decent number of users who were at the other extreme of the spectrum: they neither express curiosity for requester's past postings nor show concern for sharing their own past postings with the requester (category 4).

2) *Influence of Posting Types: Positive vs. Negative: Inappropriate posts* (32%) and *polarizing posts* (23%) turned out to be the major red flags that participants look out for in the requester's past postings. In addition, for the cases of strangers, participants were more cautious and termed *lack of past posts* (14%) to negatively impact their decision. Postings that positively influence participants' decisions tended to depict *common interests* (36%), *positive personality traits* (33%), and *background affinity* (23%). In the open-ended text box, some participants reported looking for posts that establish their link to the requester in real life, such as mutual friends.

3) *Whose Past Postings Matter More Often: Requester's or Recipient's?:* Comparing frequencies of participants' interest in friend-requesters' past posts and their concern about sharing their own past posts can provide insights into the relative usefulness of two types of past postings in users' befriending behavior. Since offline interactions can influence users' behavior, we controlled this parameter by analyzing the cases for requests received from strangers. For every 10 friend requests received from strangers, 60% of the MTurk study respondents (124) reported visiting the profiles of majority of requesters to review their past postings before making a decision, if any, on their request. In contrast, 42% of the respondents (87) were concerned that the requester will be able to fully access the history of past postings upon their decision to approve the request. A similar trend was apparent even more in the campus study with 75% of the respondents (66) opting to visit past postings of the majority of the requesters as opposed to 30%

TABLE III. ROLE OF OFFLINE INTERACTIONS

|        | Motivation to review |              | Hesitance to share |              |
|--------|----------------------|--------------|--------------------|--------------|
|        | Stranger             | Acquaintance | Stranger           | Acquaintance |
| MTurk  | 60% (124)            | 40% (83)     | 42% (87)           | 22% (46)     |
| Campus | 75% (66)             | 31% (28)     | 30% (27)           | 4% (3)       |

(27) who were concerned to share their own history of past postings. The Mann-Whitney U test confirmed the statistical significance of the difference between attitudes ( $U = 1745$ ,  $p\text{-value} < .001$ ), lending support to **H1C** (*Impact of Request Sender's Postings*). Thus, the new information learned through the requester's past postings is more likely to be critical for a user's befriending decision than the privacy concerns arising from sharing their detailed history of past postings with the requester upon acceptance of the request.

4) *Whose Past Postings Matter More: Stranger's or Acquaintance's?:* Table III details the percentage of participants who are motivated to review requesters' past postings in majority of the cases and are hesitant to share their own history of postings with them. As the numbers demonstrate, the value of postings diminishes significantly if users have an offline connection with the requester ( $U = 1854$ ,  $p\text{-value} < .01$ ). This effect applied to both attitudes: in-person interactions reduced the users' motivation to review the requester's longitudinal data and increased their willingness to share their own longitudinal data with them (**H1B** - *Impact of Offline Interactions*). For the question about hesitance to share one's own postings, we observed a tendency for answers about strangers to be polar (*Every time* (32%) and *None* (31%)), suggesting a blanket judgment one way or the other, rather than the participant thinking about each specific case on the occasion.

5) *Differences in Attitude of MTurk participants: US versus India:* In contrast to participants with Indian background, American participants tended to be proactively looking into past postings of strangers before making a decision on the request. The non-parametric Mann-Whitney U test confirmed that the differences between the two groups are statistically significant ( $U = 3750$ ,  $p\text{-value} < .01$ ).

## B. Relevance of Past Postings

Relevance captures whether a participant thinks an old post should still be in their timeline or be reshared (for whatever reason), or the degree to which it should be in the timeline/be reshared. In Table IV, we report the participants' attitude towards past postings, relating to **RQ2** (*Comfort with Longitudinal Exposure*). All options that could be checked as assessment wrt. posts' visibility were selected by the participants. Discomfort or unease was inferred by participants' selection of Definitely Not, Probably Not or Possibly on the Likert scale. Next, we detail results on all four cases.

1) *Re-sharing Preference:* Of the 197 MTurk participants who were able to find a personal post, 45% (89) expressed discomfort at the idea of re-sharing these past posts to the current context. Of the 86 campus study participants who were able to find personal post, 55% (48) expressed discomfort with the idea of re-sharing these posts to the current context. For sensitive postings, 52% (78 of 150) of the MTurk respondents

TABLE IV. USERS' ATTITUDE TOWARDS VISIBILITY OF PAST POSTS CONTAINING PERSONAL AND SENSITIVE CONTENT, BASED ON THE MTURK STUDY (N=209). MULTIPLE ANSWERS COULD BE CHECKED.

| Attitude to posts           | Postings' content |                 |
|-----------------------------|-------------------|-----------------|
|                             | Personal (197)    | Sensitive (150) |
| Unease at re-sharing        | 45% (89)          | 52% (78)        |
| Unease at keeping available | 25% (49)          | 35% (52)        |
| Desire to change exposure   | 18% (38)          | 22% (33)        |
| Desire to delete            | 22% (43)          | 33% (50)        |

TABLE V. REASONS FOR UNEASE AT RE-SHARING PAST POSTS; PARTICIPANTS OF THE MTURK STUDY (PERSONAL = 89; SENSITIVE = 78) AND THE CAMPUS STUDY (PERSONAL = 48; SENSITIVE = 31). MULTIPLE ANSWERS COULD BE CHECKED.

| Reasons for unease     | MTurk study |           | Campus study |           |
|------------------------|-------------|-----------|--------------|-----------|
|                        | Personal    | Sensitive | Personal     | Sensitive |
| Irrelevance            | 65%         | 60%       | 81%          | 65%       |
| Embarrassing to me     | 32%         | 32%       | 23%          | 3%        |
| Embarrassing to others | 18%         | 18%       | 65%          | 3%        |
| I am not sure why      | -%          | -%        | -%           | 28%       |

and 61% (31 of 51) of the campus study respondents expressed lack of comfort for re-sharing the posts on their timelines.

Table V lists the major reasons behind respondents' unease to re-share past postings to the current context. Interestingly, for the MTurk study, the sensitivity level of the post did not impact this behavior, whereas in the case of the campus study, resharing sensitive posts gave participants less specific feelings of unease than for private posts.

2) *Availability Preference*: While users' willingness to share past postings in the current context was low, we also aimed to understand if users considered those posts relevant enough to be kept online. 25% (49 of 197) of the MTurk respondents expressed lack of comfort for keeping the personal posts available on their timelines. As for the reason behind this, both "The post is irrelevant (e.g., I do not see a reason to keep it online)" and "The post depicts outdated views" were selected by roughly 39% of the respondents. 25% did not want their friends to find the post. As for sensitive posts, 35% (52 of 150) of the MTurk respondents expressed lack of comfort for keeping the posts available on their timelines. 46% of these cited "I do not make posts concerning such a topic anymore" whereas 41% chose "The post is irrelevant (e.g., I do not see a reason to keep it online)" as one of the reasons. 39% reported "The post depicts outdated views" as the reason.

3) *Exposure Preference*: When asked about their preference to change exposure settings, 18% and 22% of the MTurk participants selected to change exposure settings of the personal and sensitive posts, respectively. Differences between newly chosen and existing settings was statistically significant ( $U = 566$ ,  $p\text{-value} < .05$ ). Roughly 60% of these respondents opted to restrict access to their postings and changed exposure settings from *Public/Friends of Friends* to more private options.

*Influence of age*: Upon investigating the relationship between the respondents' age and the urge to change exposure settings of postings, we noticed a moderate negative correlation (Spearman coefficient:  $-0.3$ ,  $p\text{-value} < .001$ ) between the two, suggesting that desire to change exposure settings is higher for younger participants. Subsequently, we asked participants

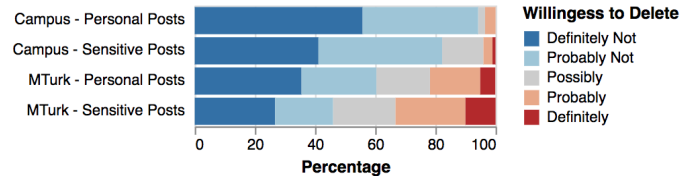


Fig. 1. Deletion Preference: personal vs. sensitive postings for both studies.

about their new preferences for the audience of these posts. Interestingly, we noticed positive correlation (Spearman coefficient:  $0.3$ ,  $p\text{-value} < .03$ ) between age and the preferred size of audience, suggesting that younger participants preferred to make their past posts private, whereas elder participants were comfortable keeping their posts open for wider audiences.

4) *Deletion Preference*: We found statistically significant difference ( $U = 12330$ ,  $p\text{-value} < .01$ ) between MTurk respondents' desire to delete personal and sensitive posts (**H2B - Discomfort w.r.t. Content**). Whereas 22% (43 of 197) of the MTurk respondents preferred to take the chosen personal post down, this number increased to 33% (50 of 150) for sensitive posts. In addition, roughly 20% of the participants were not entirely sure about their preference for this question in both cases. Figure 1 represents how willingness to delete is much higher for sensitive posts than for the personal ones. One potential reason behind this trend could be that sensitive posts containing political content are much more likely to become outdated with the passage of time. Whereas a similar trend was observed in the campus study, it was strikingly different in another aspect: Campus students' desire to delete their past posts was considerably less than the MTurk respondents, suggesting they had more confidence in their past postings.

Taken together, these dimensions of relevance lend support to the hypothesis that users' willingness to share past postings in the current context is considerably low (**H2A - General Discomfort**). Results on the hypothesis **H2B** were mixed, given that we obtained different results for the different potential behaviors. Whereas no significant differences were observed for re-sharing preference in the case of MTurk study, differences among deletion preference found statistically significant support for both user studies.

### C. Understanding of Privacy Features and Settings

Finally, we report results on the participants' awareness and understanding of Facebook privacy features and settings, relating to **RQ3 (Awareness of Longitudinal Controls)**.

1) *Privacy Tools and Exposure to Postings in Old Contexts*: While only 9% of the participants (campus study: 8, MTurk study: 18) lacked awareness about privacy controls for the current context (*selecting audience for new postings, reviewing postings you are tagged in, etc.*), as many as 35% of the participants (campus study: 28, MTurk study: 77) had never heard about the longitudinal privacy control *Limit Past Posts*. In addition, 28% of the participants (campus study: 23, MTurk study: 61), most for any feature, had never used this longitudinal privacy feature even though they were aware of its existence. Possible reasons for this could be the obscurity about the effectiveness of the feature, lack of initiative from the service about informing the users, or even the lack of need felt

TABLE VI. SUMMARY OF HYPOTHESES TESTING. STATISTICAL SIGNIFICANCE INDICATED AS: (\*) FOR  $p < .05$ , (\*\*) FOR  $p < .01$ , (\*\*\*) FOR  $p < .001$ , (\*) FOR MIXED RESULTS WITH STATISTICAL SIGNIFICANCE OBSERVED ONLY IN SOME BEHAVIORS. STATISTICAL TESTS ABBREVIATIONS: **MWU**: MANN-WHITNEY U TEST; **KWH**: KRUSKAL-WALLIS H TEST; **RTT**: RIGHT TAILED TEST. STUDY: M: MTURK; C: CAMPUS.

| Summary of Hypotheses Testing |                                |            |            |       |
|-------------------------------|--------------------------------|------------|------------|-------|
| Hypotheses                    | Description                    | Stat. sig. | Stat. test | Study |
| H1-A                          | General Impact                 | *          | RTT        | M + C |
| H1-B                          | Impact of Offline Interactions | **         | MWU        | M + C |
| H1-C                          | Impact of Requester's Postings | ***        | MWU        | M + C |
| H2-A                          | General Discomfort             | *          | RTT        | M + C |
| H2-B                          | Discomfort w.r.t. Content      | *          | MWU        | M + C |
| H3-A                          | Lack of Awareness              | ***        | KWH        | M + C |
| H3-B                          | Mismatch in Exposure Settings  | *          | MWU        | M + C |

by the users for such a feature. These differences ( $p$ -value  $< .001$ ) support our hypothesis **H3A** (*Lack of Awareness*) that the Facebook users' awareness of the platform's privacy features is not as comprehensive for past postings as it is for the postings made in the current context.

2) *Overexposure and Underexposure of Users' Data*: We observed inconsistency between users' perceived and actual exposure settings resulting in different information types to be classified into two categories: *overexposed* and *underexposed* w.r.t. users' perception. The box-plots in Figure 2 detail these findings for our campus study. Basic information (birthday, gender, etc.) and political/religious views fall into the *underexposed* category. For 'basic information', participants expected exposure to 'Friends of Friends' whereas in reality, it turned out to be a more private option (Friends Only). Similarly, for the 'political/religious views', actual settings (Only me) turned out to be more private than the participants' perceived settings (All friends except a few).

The more concerning category, *overexposed*, includes 'posts & photos' and 'personal information'. The majority of the participants, from both studies, believed that only friends could access personal information data type. However, to their surprise, the participants discovered that the information type was accessible by not only 'friends of friends' but also by the 'public' in majority of the cases. As it can be observed for Personal Information in Figure 2, two boxes' notches do not overlap at all, indicating at the 95% confidence level that the medians differ. Wilcoxon-Mann-Whitney sum test further confirmed these differences ( $p$ -value  $< .01$ ). Similarly, we found that posts and photos of the participants of both studies were overexposed to larger audiences. While the box-plot for 'Posts & photos' does not tell much more than that the median for both plots happens to be roughly the same, the Wilcoxon-Mann-Whitney sum test indicates the existence of a difference ( $p$ -value = .03) between two samples. Regardless of the category (overexposed/underexposed), hypothesis **H3B** (*Mismatch in Exposure Settings*) can be accepted given the significant mismatch.

An overview of hypotheses and their statistical significance levels is provided in Table VI.

## V. DISCUSSION

### A. Effect of Irrelevant Longitudinal Data on Befriending Decisions.

Our results in **RQ2** suggest that almost half of the participants did not see past postings relevant enough to be re-shared in the current context. Roughly one-quarter of the participants showed unease at keeping the identified past postings available on their *Timelines*. Users found neither current nor reminiscent relevance in these postings and instead indicated preference to restrict access or perform deletion of these sizable number of postings. Posts containing sensitive content were more likely to be deemed irrelevant (to certain extent).

Presence of irrelevant postings implies that users struggle to adequately manage exposure to their longitudinal data. Indeed, participants' lack of understanding of privacy features finds support in the findings of **RQ3** as multiple categories of information were found to be *overexposed* w.r.t. the intended audience. Similarly, users lacked both the awareness and the usage of *Limit Past Posts* feature, rendering their outdated, and often embarrassing longitudinal data accessible to *Friends of Friends* and *Public*.

To understand the influence of past postings on the expansion of users' friendship network, we contextualize these above findings with those in **RQ1**, which revealed that the majority of participants valued the insights learned from the history of past postings of the requester in their befriending decisions. Given the presence of irrelevant (in some sense), overexposed postings on users' *Timelines*, an unfair representation of a user is highly likely, resulting in negatively influencing their befriending experience.

*Presence of indifferent group*: The second largest number of participants belonged to the indifferent group that were neither curious about requesters' past postings nor concerned about sharing their own longitudinal data upon request acceptance. This number could be influenced by a subset of users who do not accept requests from strangers whatsoever and therefore, saw no value in past postings. For the rest, insights acquired from past postings are not the major factor influencing their decisions. Their befriending model could either rely on other investigative actions such as sending private messages and looking for mutual friends, as identified by Rashtian et al. [16] or accepting any and all requests without investigation.

### B. Future Work in Longitudinal Privacy Management

*Catering to diversity of user base*: While studying cultural differences was not the focus of the study, our findings suggest that request recipients' reliance on past postings is not consistent for American and Indian participants. Users from the US valued getting insights from the requester's past postings significantly more than their Indian counterparts. This trend could possibly hint at American users' openness and curiosity towards requesters' past postings. Conversely, it is also possible that Indian users do not accept requests from strangers and, thus, did not see value in looking into their past postings. While cultural differences might be at play, it is not a wild hypothesis to think that other factors may have a notable influence: education, religiosity, individual freedoms under

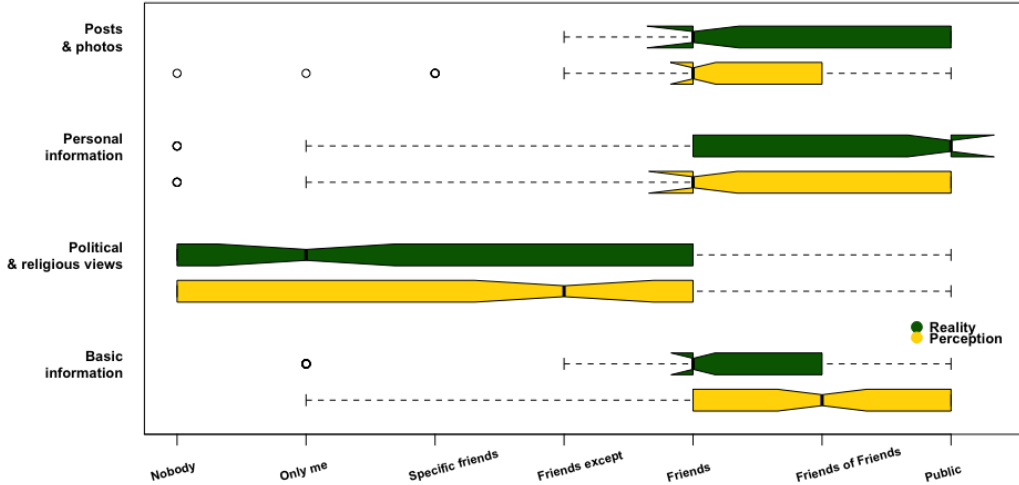


Fig. 2. Contrast between actual and perceived audience of different types of information found on profiles of campus study participants. Based on the inconsistencies visible in the box-plots, the different information types can be classified into two categories: *Overexposed* (Posts and photos & Personal information) and *Underexposed* (Political/religious views & Basic information).

current legislation, etc. To draw conclusions with universal validity, we encourage future work to design studies with the focus on uncovering the interplay between diversity and attitudes of the massive user base of the platform.

*Need for customizable solutions:* Our findings reveal that users associate reminiscence or archival value to some of their past postings and would prefer to keep those available on their *Timelines*. Current longitudinal privacy control, *Limit Past Postings*, is simply ineffective at ensuring that since it restricts access to all past postings. Such one-size-fits-all approaches are counterproductive as reflected by the large fraction of participants who never used the feature despite being aware of it. Proposals that archive the past posts [26], and thus limit their exposure to information owners alone, might not work for all either. Therefore, there is need for controls that can better cater to complex needs of the platform’s users.

Since users also expressed different levels of concern for the exposure of the postings depending upon their content, straightforward solutions, such as setting a default expiration time [18], [27], would not be enough to satisfy most users’ needs. We see value in exploring digital forgetting directions that can realize flexible expiration times [28] by taking different heuristics, such as the posting’s content, audience, and user’s privacy attitude into account.

*Tackling lack of initiative on users’ side:* Our results show evidence that user’s awareness of privacy features is not as comprehensive for postings made in the old context as it is for posting made in the current context. We report that the vast majority of users visited the Facebook privacy settings rarely, i.e., hardly once a year. The lack of initiative on the user side to learn more about the existing privacy options could be attributed to the difficulty of setting up the existing privacy management schemes correctly and efficiently. In line with Bauer et al.’s recommendation [8], we agree that efforts should be dedicated to the design of effective interfaces that help users avoid regrettable online disclosures while minimizing effort required on their part. In this regard, proposals, such as

Wang et al.’s [29], to “nudge” users to consider the content and context of their online disclosures are worth further research.

### C. Limitations

*Recruitment:* Since we recruited Facebook users for the purpose of our study, our results may not be applicable to other OSNs, especially those lacking the *Timeline* feature to access past postings. That being said, as Facebook is the largest social media platform with more than 2 billion active users [1], our findings are highly relevant, timely and impactful. Recruitment from campus for research introduces its own limitations such as homogeneity in age, behaviors, life experiences, etc. This is why we followed up our first study with a second one on the Amazon MTurk platform to recruit participants with more representative age groups, professions, and experiences.

*Validity:* We made a decision to design the study keeping in mind the privacy of participants, which meant avoiding direct access to users’ accounts and instead relying on self-reported information. Regarding both participants’ review of requesters’ past posts and their concerns about requesters viewing their own (the participants’) posts (whatever the concerns may actually have been), we acknowledge that we do not have evidence that the participants actually made different decisions based on them. We believe the validity concerns here are not as high as for privacy-invasive procedures. If the study had required participants to agree to an automated way of gathering information from their profiles, we would have introduced a bigger bias in our results: privacy-aware users would have been less likely to participate in the study, resulting in less generalizable results. In addition, since Facebook does not afford users any feature to keep track of accepted or denied friend requests from the past, we had to resort to respondents’ recalling capabilities. To minimize the validity concerns, we purposefully offered users broad enough categories when recalling their behavior. Alternate approaches that require creation of mock profiles to send dummy requests for observation of participants’ behavior would have introduced other concerns, such as lack of accounting for participants’



biases to demographic background of the dummy requesters, given the cultural diversity of our study participants.

#### D. Related Work

*Influence of Digital Footprint:* Both user studies [30], [19] and media [31], [32] have reported on how digital footprint can have serious consequences, such as termination of careers. In their work, Mohamed et al. [30] identified the challenges posed to the reputation of employees by their growing online activity. Our study demonstrated the role digital footprint could play on the expansion of users' online friendship network. Whereas the effect of content of postings on unfriending decisions [21] has been discussed before, existing models for befriending behavior [33], [16] have not explored the impact of users' growing digital footprint on this process. Our work bridges the gap in knowledge about the role of digital footprint on this central process of OSNs.

*Digital Footprint, Relevance and Exposure:* Mondal et al. [34] studied the longitudinal exposure of shared data to find that a significant fraction of users withdraws a large percentage of old publicly shared data on Twitter. Madejski et al. [35] identified mismatch between users sharing intentions and privacy settings and called for a contextual privacy settings. Liu et al. [11] found that privacy settings match users' expectations only 37% of the time. Similar concerns were raised in other studies ([36], [26]), indicating lack of control experienced by the Facebook users. Our research builds on prior work aimed at understanding users' awareness of privacy tools. We specifically focused on users' awareness and usage of longitudinal privacy controls that allow users to manage content published in the old context.

## VI. CONCLUSION

We analyze the participants' longitudinal data on Facebook for its perceived relevance, exposure control and influence in their befriending behaviors. Our results indicate that although a significant number of past postings are perceived as irrelevant in some sense, they have the potential to impact the befriending behavior of users in the present context. Inappropriate and polarizing posts turned out to be major red flags that participants scrutinized the requesters' *Timelines* for. Posts depicting common interests and positive personality traits were significant contributors to the acceptance of the request. Additionally, we revisited users' understanding of longitudinal privacy controls and make recommendations for the design of user interfaces and features to minimize regrettable disclosures.

## ACKNOWLEDGMENT

The authors thank the anonymous reviewers for their helpful comments. This research was supported by the Center for Cyber Security at New York University Abu Dhabi (NYUAD).

## REFERENCES

- [1] Statista, *Facebook users worldwide 2018*. Statista, 2018.
- [2] Domo, "Data never sleeps 7.0," 2019.
- [3] J. Titcomb, "Facebook 'likes' reveal personality," 2015.
- [4] R. Farahbakhsh, X. Han, A. Cuevas, and N. Crespi, "Analysis of publicly disclosed information in Facebook profiles," in *2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2013)*, ASONAM, pp. 699–705, 2013.

- [5] L. Thompson, K. Dawson, R. Ferdig, E. Black, J. Boyer, J. Coutts, and N. Black, "The intersection of social networking with medical professionalism," *J. Gen. Intern. Med.*, vol. 23, no. 7, 2008.
- [6] K. N. Nason, H. Byrne, G. J. Nason, and B. O'Connell, "An assessment of professionalism on students' facebook profiles," *Eur J Dent Educ.*, vol. 22, no. 1, pp. 30–33, 2016.
- [7] O. Ayalon and E. Toch, "Not even past: Information aging and temporal privacy in online social networks," *Human-Computer Interaction*, vol. 32, no. 2, pp. 73–102, 2017.
- [8] L. Bauer, L. F. Cranor, S. Komanduri, M. L. Mazurek, M. K. Reiter, M. Sleeper, and B. Ur, "The post anachronism: The temporal dimension of Facebook privacy," in *Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society, WPES*, (New York, NY), pp. 1–12, ACM, 2013.
- [9] C. Fiesler, M. Dye, U. Feuston, A. S. Bruckman, M. De Choudhury, et al., "What (or who) is public? Privacy settings and social media content sharing," in *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, CSCW*, pp. 567–580, 2017.
- [10] R. Gross and A. Acquisti, "Information revelation and privacy in social networks," in *Proceedings of the ACM Workshop on Privacy in the Electronic Society, WPES*, (New York, NY), pp. 71–80, ACM, 2005.
- [11] Y. Liu, K. P. Gummadi, B. Krishnamurthy, and A. Mislove, "Analyzing Facebook privacy settings: User expectations vs. reality," in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference, IMC*, (New York, NY), pp. 61–70, ACM, 2011.
- [12] M. Madejski, M. Johnson, and S. M. Bellovin, "A study of privacy settings errors in an online social network," in *2012 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom*, pp. 340–345, 2012.
- [13] M. Sleeper, R. Balebako, S. Das, A. L. McConahy, J. Wiese, and L. F. Cranor, "The post that wasn't: Exploring self-censorship on Facebook," in *Proceedings of the Conference on Computer Supported Cooperative Work, CSCW*, (New York, NY), pp. 793–802, ACM, 2013.
- [14] K. Vaccaro, D. Huang, M. Eslami, C. Sandvig, K. Hamilton, and K. Karahalios, "The illusion of control: Placebo effects of control settings," in *Proceedings of the CHI Conference on Human Factors in Computing Systems, CHI*, (New York, NY), pp. 16:1–13, ACM, 2018.
- [15] J. Henrich, S. J. Heine, and A. Norenzayan, "Most people are not weird," *Nature*, vol. 466, pp. 29–29, 2010.
- [16] H. Rashtian, Y. Boshmaf, P. Jaferian, and K. Beznosov, "To befriend or not? A model of friend request acceptance on Facebook," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, SOUPS, (Menlo Park, CA), pp. 285–300, USENIX Association, 2014.
- [17] F. Nagle and L. Singh, "Exploring privacy in online social networks," in *2009 International Conference on Advances in Social Network Analysis and Mining, ASONAM*, pp. 312 – 315, 2009.
- [18] O. Ayalon and E. Toch, "Retrospective privacy: Managing longitudinal privacy in online social networks," in *Proceedings of the Ninth Symposium on Usable Privacy and Security, SOUPS*, (New York, NY), pp. 4:1–4:13, ACM, 2013.
- [19] Y. Wang, G. Norcie, S. Komanduri, A. Acquisti, P. G. Leon, and L. F. Cranor, "'I Regretted the Minute I Pressed Share': A qualitative study of regrets on facebook," in *Proceedings of the Seventh Symposium on Usable Privacy and Security, SOUPS*, (New York, NY, USA), pp. 10:1–10:16, ACM, 2011.
- [20] N. Stokes, "The guide to Facebook privacy settings," 2018.
- [21] C. Sibona and S. Walczak, "Unfriending on Facebook: Friend request and online/offline behavior analysis," in *2011 44th Hawaii International Conference on System Sciences, HICSS*, pp. 1–10, 2011.
- [22] J. Nielsen, T. Clemmensen, and C. Yssing, "Getting access to what goes on in people's heads?: Reflections on the think-aloud technique," in *Proceedings of the Second Nordic Conference on Human-computer Interaction, NordiCHI*, (New York, NY), ACM, 2002.
- [23] J. Ross, L. Irani, M. S. Silberman, A. Zaldivar, and B. Tomlinson, "Who are the crowdworkers?: Shifting demographics in mechanical turk," in *CHI '10 Extended Abstracts on Human Factors in Computing Systems, CHI*, (New York, NY, USA), pp. 2863–2872, ACM, 2010.
- [24] Qualtrics, "The research software," 2018.
- [25] K. Crowston, "Amazon Mechanical Turk," 2012.

- [26] X. Zhao, N. Salehi, S. Naranjit, S. Alwaalan, S. Volda, and D. Cosley, "The many faces of Facebook: Experiencing social media as performance, exhibition, and archive," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI*, (New York, NY), pp. 1–10, ACM, 2013.
- [27] R. Geambasu, T. Kohno, A. A. Levy, and H. M. Levy, "Vanish: Increasing data privacy with self-destructing data," in *Proceedings of the 18th Conference on USENIX Security Symposium, SSYM*, (Berkeley, CA, USA), pp. 299–316, USENIX Association, 2009.
- [28] A. Zarras, K. Kohls, M. Dürmuth, and C. Pöpper, "Neuralyzer: Flexible expiration times for the revocation of online data," in *Proceedings of the Sixth ACM Conference on Data and Application Security and Privacy, CODASPY*, (New York, NY, USA), pp. 14–25, ACM, 2016.
- [29] Y. Wang, P. G. Leon, K. Scott, X. Chen, A. Acquisti, and L. F. Cranor, "Privacy nudges for social media: An exploratory Facebook study," in *Proceedings of the 22nd International Conference on World Wide Web, WWW*, (New York, NY), pp. 763–770, ACM, 2013.
- [30] R. E. Mohamed, T. B. Idalino, and S. Chiasson, "When private and professional lives meet: The impact of digital footprints on employees and political candidates," in *Proceedings of the 8th International Conference on Social Media & Society, SMSociety*, (New York, NY, USA), pp. 48:1–48:5, ACM, 2017.
- [31] E. Shugerman, "Acceptances withdrawn for facebook memes," *The Independent*, 2017.
- [32] Staff, "Anthony weiner scandal: A timeline," *CNN*, 2016.
- [33] M. Rahman, V. Mezhuyev, M. Bhuiyan, S. Sadat, S. Zakaria, and N. Refat, "Reliable decision making of accepting friend request on social networks," *IEEE Access*, vol. 6, pp. 9484–9491, 2018.
- [34] M. Mondal, J. Messias, S. Ghosh, K. P. Gummadi, and A. Kate, "Forgetting in social media: Understanding and controlling longitudinal exposure of socially shared data," in *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, SOUPS, (Denver, CO), pp. 287–299, USENIX Association, 2016.
- [35] M. Madejski, M. J. maritzaj, and S. M. Bellovin, "The failure of online social network privacy settings," in *CUCS, Columbia University Computer Science Technical Reports*, 2011.
- [36] H. R. Lipford, A. Besmer, and J. Watson, "Understanding privacy settings in Facebook with an audience view," in *Proceedings of the 1st Conference on Usability, Psychology, and Security, UPSEC*, (Berkeley, CA, USA), pp. 2:1–2:8, USENIX Association, 2008.
- h) Since you are comfortable keeping this post on your timeline, which of the following would describe the underlying reasons? Select all that apply.
- The post holds value to me
  - The post is still relevant
  - I want my Facebook friends to continue accessing it
  - I want to keep the post for archival reasons
  - Other Reason: —
- i) You are not comfortable keeping this post on your timeline, which of the following would describe the underlying reasons? Select all that apply.
- The post depicts outdated views (i.e. my views have changed)
  - The post is irrelevant (e.g. I do not see a reason to keep it online)
  - I do not want my Facebook friends to find this
  - I do not make posts concerning such a topic anymore
  - Other Reason: —
- j) Do you prefer to take this post down?  
 • Definitely to Definitely Not (5-point scale)
- 2) Think of the last ten times you received a Facebook friend request from people (i) you have met in person (acquaintance) and (ii) you have not met in person (strangers). [Once participants responded to questions a) - d) for Part (i), they were asked to answer those for Part (ii).]
- a) In roughly how many of those instances did you visit their Facebook Wall to take a look at their past postings before deciding on whether to accept or decline the request? (i) None (ii) 1-4 (iii) 5-9 (iv) Every time
- b) Once at their profile, what types of past postings influence your friend request decision negatively (i.e. rejecting the friend request)? Select all that apply.
- Polarizing posts (politics, religion)
  - Frequent/unimportant posts (unimportant, too much)
  - Inappropriate posts (sexist, racist, swear, sex, etc.)
  - Everyday life posts (exercise, spouse, child, celebrities, sports, etc.)
  - Lack of past posts
  - Other kinds of postings: —
- c) Once at their profile, what types of past postings influence your friend request decision positively (i.e. accepting the friend request)? Select all that apply.
- Posts depicting common interests (hobbies, exercise, sports, etc.)
  - Posts depicting positive personality traits
  - Posts depicting their background (hometown, college, etc.)
  - Other kinds of postings: —
- d) In roughly how many of the above 10 friend request instances were you concerned that the sender will be able to have full access to history of your past postings upon your decision to approve the request?  
 (i) None (ii) 1-4 (iii) 5-9 (iv) Every time
- 3) Privacy features
- a) How often do you visit the privacy settings of your Facebook profile?  
 (i) Daily (ii) Weekly (iii) Monthly (iv) Yearly (v) Never
- b) Out of all privacy features that Facebook allows, do you know the following options?
- Selecting an audience for stuff you share
  - Reviewing stuff others tag you in
  - Limiting access to the private information in the About section
  - 'Limit Past Posts' to minimize the audience of old posts from Timeline
  - Selecting audience for a post you have already deleted
  - 'Friend Request Setting' to determine who can send you friend requests
  - Limiting access to your posts by certain individuals through the use of 'Restricted List'
- I am aware of this feature and have used it • I am aware of this feature and have not used it • I am not aware of this feature.
- c) Part A: Without consulting your Facebook profile, answer who do you think can see the following types of information on your profile?  
 Part B: Now visit your Facebook profile and confirm the actual audience of the posting.
- Basic information (birthday, birth year and gender)
  - Contact information (emails, address and phone numbers)
  - Relationship status
  - Political and Religious views
  - Personal information (activities, interests, about me, favorite movies, TV shows, books, and quotes)
  - Your posts and photos
  - Choose 'Friends of Friends' option for this question
- Only me • Custom Settings • Specific friends • Friends except • Friends • Friends of Friends • Public (Anyone on or off Facebook)
- 4) Demographics
- a) What is your age? (i) — (ii) I prefer not to disclose
- b) Which gender do you identify with the most?  
 (i) Male (ii) Female (iii) Diverse (iv) I prefer not to disclose.
- c) Which country did you spend most time growing up?  
 (i) — (ii) I prefer not to disclose
- 5) Attention Checks
- a) Please Choose 'Friends of Friends' as an option:  
 (i) Only me (ii) Custom Settings (iii) Specific friends (iv) Friends except (v) Friends (vi) Friends of Friends (vii) Public (Anyone on or off Facebook)
- b) Are you aware of the following feature and have you used it?: "Selecting audience for a post you have already deleted"  
 (i) I am aware of this feature and have used it (ii) I am aware of this feature and have not used it (iii) I am not aware of this feature.

## APPENDIX

### A. Survey Questionnaire

Note: We present an abridged version of the survey here. Questions are renumbered for presentation, and visual details are removed for concision. Instructions and visual aids were given at each stage of the process to assist participants of the survey.

- Guidance was provided on how to scroll back on the timeline by 3 years. Identify the first post of personal nature (relating to one's self, family, etc.) and the first post of sensitive nature (pertaining to religious/political views, etc.) that appears on your timeline. For each of the posts, answer the following questions:
  - Were you able to find a post? (i) yes (ii) no
  - What is the current privacy settings of the post?  
 (i) Only me (ii) Custom Settings (iii) Specific friends (iv) Friends except (v) Friends (vi) Friends of Friends (vii) Public (Anyone on or off Facebook)
  - Since you are here now, do you want to change privacy settings for the post?  
 (i) yes (ii) no
  - Who would you prefer to be the current audience for the post now?  
 (i) Only me (ii) Custom Settings (iii) Specific friends (iv) Friends except (v) Friends (vi) Friends of Friends (vii) Public (Anyone on or off Facebook)
  - Do you feel comfortable resharing this post on your profile?  
 • Definitely to Definitely Not (5-point scale)
  - If you are not comfortable sharing the post, which of the following would describe best the underlying reason? [Participants were only asked this questions if their answer to e) indicated discomfort]
    - I believe the post is too old to be relevant now
    - I feel resharing the post will be embarrassing to me
    - I feel resharing the post will be embarrassing to others tagged
    - I should not have posted it in the first place
    - I am not sure why/Other:
  - Do you feel comfortable keeping this post on your timeline, accessible by everyone? [Based on their answer to this question, participants were asked either question h) or question i)]
    - Definitely to Definitely Not (5-point scale)