Uncovering Personal Histories: A Technology-Mediated Approach to Eliciting Reflection on Identity Transitions

OLIVER L. HAIMSON, University of Michigan
MEGH MARATHE, University of California

When studying identity transitions, interview participants can find it difficult to reflect on their transitions and recall specific details related to past experiences. We present a new approach to enable participant reflection on past identity transitions, and a means to fill in blanks by eliciting data that may not otherwise come up: showing participants sentiment visualizations of their social media data. After detailing our methods of constructing sentiment visualizations, we discuss our experiences using them in a study on gender transition. For most participants, the visualizations elicited substantial reflection, and enabled recalling forgotten data and new interpretations of transition experiences. We guide researchers on how to use this method when studying other identity transitions; this may be especially powerful for marginalized people who undergo substantial identity changes. This article proposes a way to uncover participants' personal histories, which can help HCI researchers to better understand and support marginalized people's experiences.

CCS Concepts: • **Human-centered computing → Human computer interaction (HCI)**; *Empirical studies in collaborative and social computing;*

Additional Key Words and Phrases: Reflection, personal histories, sentiment visualization, interviews, research methods, identity transitions, social media

1 INTRODUCTION

Many life transitions (e.g., gender transition, relationship changes, and changes in one's health) are processes that take place over months or years and involve substantial changes in a person's identity. Identity change is defined as change in the meanings that "define who one is" [15]. When studying such identity transitions, it can be difficult to elicit participants' reflection on their transition because people have difficulty recalling emotions and other specific details related to past experiences. Retrospective interviewing techniques can be useful [64, 92], but carry limitations

ACM Transactions on Computer-Human Interaction, Vol. 30, No. 2, Article 23. Publication date: March 2023.

23

because people have difficulty remembering experiences [50] and do not always remember the emotions they felt during the past [65, 101]. A researcher may know how to ask about an identity transition process as a whole, but it can be hard to dig deeper when identity transition processes differ greatly from person to person. Furthermore, identity transitions do not occur in a vacuum, but overlap and intersect with other life transitions, and it can be difficult for participants to recall how and when these intersecting experiences occurred and how they felt during those time periods.

Information visualization can be a powerful tool for reflection and insight on personal data [17, 46, 96]. Visualizations can trigger memories and enable self-reflection on past experiences, such as by visualizing "familiar patterns" that correspond to people's memories and help them recognize their personal history [94]. Encouraging reflection on past mood or experiences can sometimes lead to increased wellbeing [44, 86], even if those past experiences are negative [48]. Visual stimuli for elicitation and reflection can be created by participants [6, 20], by researchers [21], or collaboratively between the two [95].

We present a new research approach to enable interview participants to reflect on past identity transitions, and a means to fill in blanks by eliciting data that may not otherwise come up in qualitative data collection: showing participants sentiment visualizations of their longitudinal social media data. A unique value of qualitative research is its ability to combine multiple layers of data, through data triangulation using social media and/or visual content in combination with interview data [54]. Though qualitative interviews are already a rich source of data, triangulating interview data with participant reflection on longitudinal social media sentiment visualizations adds an important data layer that we found brings up unique new insights. We detail our methods for constructing sentiment visualization graphs and using them in interviews with participants, so that other researchers can use this approach in their work. We discuss our findings that emerged using this method in $n = 20$ interviews with participants who had recently experienced or were currently in the process of undergoing gender transition. After showing interview participants graphs visualizing approximations of their sentiment over time (via their social media data), participants brought up and reflected on experiences from the past that they had not mentioned prior to viewing the graphs. We found that for most participants the visualizations elicited substantial reflection on their personal history, and enabled recalling forgotten data and communicating new interpretations of identity transition experiences. By *reflection*, we mean that the sentiment graphs prompted participants to think seriously about aspects of their lives that the graphs made visible (that were previously invisible to us as researchers) and their emotions around those experiences.[1]

This method's novelty lies in combining four elements, which together form an elicitation tool to use in qualitative interviews: (1) personal social media data; (2) sentiment analysis; (3) information visualization; and (4) longitudinal representation. This novelty manifests in the elicitation tool's *presentation*, *source*, and *purpose*. In terms of *presentation*, we visualize longitudinal trends and patterns in participants' social media data, whereas other researchers have typically employed individual social media posts [63, 79–81], photographs [6, 7], and diagrams [20] for elicitation. That is, we employ a novel data visualization and presentation technique—a longitudinal and aggregated representation of sentiment detected in the participant's social media posts—instead of asking participants to focus on the content of a social media post or image from the past. Our approach condenses a long time period into graph form to make longitudinal sentiment patterns visible. In terms of *source*, our visualizations draw on social media data authored by the participant, whereas other studies that use aggregated visualizations depended on data generated by sensors and other self-tracking devices [29, 33, 97]. In terms of *purpose*, our elicitation method aims at enabling

---

[1]There is no standard definition of reflection in HCI [9, 10], so we draw upon aspects of Baumer's reflective informatics [9] and Sengers et al.'s reflective design [84] in this definition.

participants to reflect in a *descriptive and expansive* manner about their personal data and life history (e.g., how did my past experiences, and the patterns they convey, shape my retelling of my life's history and my experiences?), whereas much prior work on self-reflection tends to be aimed at having participants think in an *evaluative* manner (e.g., how well am I doing in terms of achieving my fitness goals? What content do I post on social media and why?). We seek to help participants in "the creation and passing on of life stories" by reconstructing their personal transition histories from the perspective of their current identity [5]. Our approach enables reflection, particularly on identity transitions that were longer *processes* rather than events, in unique ways. We guide researchers on how to use this method in their research to fill gaps when studying identity changes. Our method may be especially powerful for marginalized groups and people who have experienced substantial identity changes in their lives, such as in our case study with transgender participants.

This method provides a unique advantage for uncovering personal histories: visualizing participants' sentiment patterns based on their own self-reported data helps participants remember and reflect on particular identity change experiences from the past. In addition, this method helps participants recall and describe experiences that the interviewer would otherwise have no way of knowing about based on the research topic, the participant's self-reported data, and the participant's prior interview responses. Using sentiment measures is important because it provides longitudinal presentation of a proxy for people's emotional state over time, which enables a unique way for participants to reflect on how they may have been feeling over time and in different time periods. This method can be used in studies that aim at understanding people's past experiences and emotions, which is vital knowledge in designing technology. We expect that this method will be particularly useful for understanding difficult identity transitions that involve stigma and marginalization, because people are likely to document such transitions in narrative form [59], and because identity transitions are processes that unfold over time (rather than more time-bounded life events).

We discuss how using this method enabled us to construct what we call a *transition history from below* (drawing from historian E. P. Thompson's "history from below" [91]), that is, narrative and analytical insight on transition based upon an examination of past events that centers the perspectives of those who experienced them. We show that a gender transition history from below provides insights on physical and mental health complexities during and temporal experiences of gender transition that are missing from mainstream narratives written by medical experts.

Our contributions in this article include: (1) a description of a novel method for eliciting participant reflection on past events and the emotions surrounding them: sentiment visualizations of social media data over time; (2) an understanding of how participants responded to this method during a trial of its use with 20 interview participants in the gender transition context; and (3) a discussion of how participant reflection using sentiment visualizations contributes to literature on identity change and qualitative interview elicitation methods. This article contributes methodological insights that combine historical, qualitative, and computational methods to inform HCI research that uncovers and shines analytical light on participants' personal histories. Our method for eliciting reflection on participants' pasts, especially those who have experienced substantial disruptions, can help us further understand how HCI research and design can support marginalized experiences in the future.

2 RELATED WORK

We situate this work within prior research that has used related approaches for reflection, such as visual tools, visualizations, and technological tools. We then discuss previous work related to collaborative sensemaking of visualized data and difficulties people may have interpreting such data. Finally, we describe how our work builds from prior HCI research on identity transitions.

2.1 Existing Reflection Research Methods

The method described in this article is of direct relevance to the broader field of oral history and the sub-field of narrative inquiry that seeks to elicit the kinds of histories from below that we explore in the context of identity transitions. Narrative researchers analyze participants' "stories or narratives or descriptions of a series of events," by using "words as data" [18, 75]. In this way, researchers can focus on the particulars that characterize a participant's narrative instead of aspiring to generalization and universality, be more inclusive of different ways of knowing, and acknowledge the mutually shaping role between participant and researcher [18, 75]. Narrative researchers who study identity are particularly interested in examining how people revise and reinterpret their identities and life stories over the course of their lives [5, 18]. Our method seeks to help elicit the types of personal histories and temporal narratives that are valuable to understanding narrative identity, by reminding participants of their emotional journey through visualizations of sentiment scores based on their personal social media data. In doing so, our method uses the participants' written words as data, aims at highlighting the specifics of participants' stories, and provides an artifact that participants and researchers can attend to together.

Our method also shares elements with interview methods in which researchers ask participants to view and reflect on social media data or features, which often encourages recollection and reflection and increases the "thickness" of the data [27, 34, 54, 63, 79–81, 102]. Scrolling through past data brings digital traces to life and helps participants understand social media use as part of their life change processes over time [79]. Yet our approach makes an important departure from this prior work. By showing participants aggregated visualizations of their social media posts, we invite participants to examine the trends and patterns in their social media data instead of focusing on the content of particular posts and images. That is, our method enables participants to make sense of the bigger picture (i.e., trends and patterns) reflected in their social media posts, in contrast to taking them through individual data points (i.e., particular posts and images). This article's method also shares similarities with usability methods such as cognitive walkthroughs [55], in which people evaluate a technology by "walking through" a series of tasks while talking aloud. In the Elicitation Interview technique, researchers ask people to interpret a visualization of non-personal data; yet participants looked for personal connections to the data, often related to particular time periods [43]. Previous research has also used visual probes in qualitative research [6, 7, 20], such as "longitudinal participant-driven photo elicitation" [35]. Visual tasks and graphical elicitation techniques are helpful for uncovering "data related to emotions and emotional experiences" [20], often prompt reflection, and enable people to move beyond thinking in words, which can bring up new insights and reflections that may not have been revealed otherwise [6, 21]. This work answer's Crilly et al.'s [21] call for future work to focus on new populations, contexts, and types of visual elicitation to understand how to use these techniques in varying ways. We detail a new method that addresses some limitations in existing qualitative research methods by providing a way to fill in blanks and elicit reflection on experiences the researcher would not know to ask about.

2.2 Visual Tools for Reflecting on Participants' Data

People often look at past social media data as a way to reflect on their lives [102]. People share mostly mundane/everyday moments on social media, with occasional "critical moments" [92] – this makes visualized social media data particularly informative for viewing the past. In prior research, two tools used social media data to enable self-reflection: De Choudhury et al.'s Moon Phrases [22], a prototype using social media data to visualize emotion and linguistic expression over time, and Li et al.'s Grafitter [56], a tool that collected and visualized social media data. Another system visualized and enabled discussion around aggregate social media sentiment for

particular topics in real time [16]. Yet rather than being a tool built for self-reflection based on intentional personal data collection, the present study presents social media data collection and visualization as a research method for reflection during participant interviews. De Choudhury and Massimi [23] and Baumer et al. [11] employed somewhat similar methods in survey-based studies, in which they gathered survey participants' reactions to aggregate graphs of empirical findings, and asked for participants' reactions, explanations, estimates, and whether the findings were "representative of their own behavior" [23]. Baumer et al. [11] argued that data interpretation from people who are being studied, who may either corroborate or challenge the researchers' interpretation, can benefit analysis. Our study uses elements similar to these methods, but extends them by showing participants individual rather than aggregate graphs, and with a goal of reflection elicitation.

2.3 Visualizing People's Activities Over Time

Several studies have experimented with different types of visualizations that display certain aspects of people's activities over time. Begole et al. [12] algorithmically detected and visualized people's daily activity over time to display work behaviors to colleagues. Vrotsou et al. [100] used a similar approach to display people's activity patterns to researchers in the form of a visualization. Researchers have also explored innovative ways to visualize personal informatics, such as 3D printed material artifacts [51]. Fischer et al. [33] studied household "Internet of things" data visualized over time, and discussed this data with the households. In this study, they documented the work required for participants to make sense of visualized data and relate it to household activities, and concluded that what data means is determined by the interaction of people and data [33]. Participants in Tolmie et al.'s [97] interviews and ethnographic study reflected on domestic sensor data over time, a process that required "situated reasoning" and articulation work. People in that study seemed to need to account for peaks in graphs, and did so by describing circumstances that might have caused them [97]. In this study, we take these sorts of approaches a step further and enable participants to reflect on their visualized longitudinal personal data in interviews to provide further insights that would otherwise be hard to gain.

2.4 Types and Goals of Technologically-mediated Reflection

Previous research has examined different forms and myriad goals of technology-mediated reflection [48, 52], meaning technology that enables people to reflect on past experiences generally recorded online or using an application of some kind. Prior work has primarily examined reflection in the context of personal informatics systems, defined by Li et al. as systems that "help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge" [57]. Li et al. [57] developed the stage-based model of personal informatics systems to describe the five stages that people go through when using personal informatics systems: preparation, collection, integration, reflection, and action. Researchers examined personal informatics data with goals such as modeling relationships between activities and mood [44, 86], tracking mood over time to influence future wellbeing [44, 86], self-tracking to influence behavior change [70], and measuring stress levels at past moments [77]. Researchers have often been interested in self-tracking for goals related to reflection, such as understanding people's self-reflection behaviors on self-tracked data [58], and enabling self-trackers to find meaning [30] and insights in and reflect on their own data through visualizations [17]. Users are interested in self-tracking for goals like self-monitoring, self-reflection, and behavioral/psychological/medical change or improvement [61]. Other lines of research have sought to understand how reflection on past data and experience can impact current mood [44, 48, 52]. Finally, researchers have developed tools to enable people to spontaneously reminisce on past social media content [72] and to easily express wellbeing status to one's social

network, a process, which also enabled self-reflection [4]. Elsden et al. [29] studied how people interact with and make meaning from past quantified data (e.g., from sensor-based apps and personal informatics tools), and found that personal data's meaning and experience changes over time: data can become meaningful digital possessions and personal accounts of the past. Reflecting on past personal data requires what Elsden et al. [29] call "data-work:" the language, effort, and reflection required to make sense of one's quantified data from the past. We make connections with personal informatics research because our method shares similarities with personal informatics methods—that is, both involve tracking and reflecting on personal data over time. However, because it requires no self-tracking from the participant, our method can be used in contexts beyond traditional personal informatics data collection contexts. Additionally, our contribution is an expansion of qualitative methods, rather than a personal informatics contribution.

2.5 Collaborative Sensemaking of Visualized Data

In research, personal data is best understood collaboratively between researchers and the person who created the data [97]. In Fischer et al.'s study [33], making sense of visualized data was a collaborative process between household members and researchers. Attempting to interpret such data without the people who created it eliminates substantial context [33]. Interacting with visual content, whether the visual is made by the participant or the researcher, not only increases data understanding, but can also improve communication between the two people [6]. Pina et al. [74] argued for expanding self-tracking design to collaborative family contexts, rather than only individual contexts. Similarly, we present a way that interpreting participants' data can be a collaborative process between researchers and participants.

2.6 Difficulties Interpreting Visualized Data

Visualized data are not interpretable for all users [6, 20, 78]. Novices often have difficulty understanding and reflecting on visualized self-tracked data, and may consider the reflection process a burden without rewards if the visualization is too abstract and removed from their experiences [78]. Thus, interest in and use of personal informatics can fade quickly [78]. Additionally, people may not have data reflection goals for themselves [17]. Graphic elicitation techniques can be particularly difficult for those who do not read graphs well [17], or have difficulty thinking spatially [20] or viewing time linearly [6]. The analytic sensemaking required to interpret visualized data can cause substantial burden for some [86]. Some people require help interpreting their data, which in our process is done collaboratively with the researcher when reflecting on visualizations of past emotional data. Eliciting reflection on sentiment data during interviews, and collaborative sensemaking with a researcher, could enable these people to still receive some positive benefits of reflection (e.g., improved emotional wellbeing [48], finding meaning and insight in past events [14, 17, 30]) without the burdens (e.g., having to track data themselves, interpreting graphs on their own without explanation [31, 57]).

2.7 A New Method to Inform Identity Transitions Research

Finally, many studies within social computing [3, 26, 37, 62, 67, 77, 83] and from the broader social science literature [45, 68, 90] have examined identity transitions, life events, and the emotions surrounding them, and researchers in these areas may find this new method useful. Major life events inventories like the Social Readjustment Rating Scale [45] and the Major Life Events Taxonomy [41] are not feasible in an interview context, but researchers may still wish to learn about meaningful experiences from participants' pasts without running through an exhaustive list. In Elsden et al.'s study [29], which showed participants quantified data from sensor-based apps and personal informatics tools, many participants discussed and reflected on life transitions that happened in

the past, and translated graphs into personal stories, often about meaningful life changes: "the data, in various ways—through maps, graphs, peaks, and troughs, absences—offered a legible reflection of these life changes." Using sentiment visualization graphs of participants' social media data as an elicitation tool can be useful to expand, deepen, and fill in gaps with interview data in research settings like these.

Studies about identity transitions are a growing research area in HCI and social computing [42]. For example, studies have focused on experiences of pregnancy and parenthood [1, 28, 36], transition to college [69, 85], health status changes [60, 62], transition to and away from the military [26, 83], and gender transition [39], to name a few. Particularly when identity transitions are difficult, uncommon, or isolating, people often document their transition processes, whether publicly or privately, as a way to benefit both themselves and the community of others who are also experiencing that identity transition [37, 59]. Interviews, which usually last only an hour or two (and often less), cannot possibly elicit as much personal information as people share when they regularly document an identity change in writing as it occurs. Our approach can harness longitudinal personal data in a way that can be easily presented to a participant to enable reflection and fill in data gaps. Our method uses data collection and visualization techniques inspired by and similar to those often employed in personal informatics, yet rather than making a personal informatics or design contribution, we contribute a new qualitative data elicitation method to help uncover participants' personal histories.

3 METHODS

This study was approved by our institution's Institutional Review Board. We interviewed 20 transgender and/or non-binary Tumblr users who kept a genre of blog called "transition blogs" in which they had been documenting their gender transitions for a period of months or years prior to the interview. We chose gender transition as our case study because it is an identity transition that is a process rather than an event, but that also includes many events during that process (e.g., disclosures to different audiences), and that involves substantial change over time. We chose Tumblr as this study's data source because it is a site that this population used heavily to share lengthy, personal text data over long periods of time, which enabled meaningful longitudinal sentiment analysis and visualization.

We first found a sample of transition blogs by searching on Tumblr for transgender and gender-transition related tags, such as #transition, #mtf, and #ftm, and then by searching for other tags that emerged in that initial sample of transition blogs. We then used theoretical sampling [88] to choose a sample of interview participants from the larger sample of blogs who were diverse demographically and had different sentiment patterns over time. That is, when choosing interviewees, we selected some bloggers with overall positive trajectories, some with less positive trajectories, some whose sentiment patterns were volatile, and some with less volatility. We contacted each potential interviewee through Tumblr's messaging system to ask if they would be interested in being interviewed for the study. We informed each participant about the blog data collection and analysis, and gave them the opportunity to opt out. Those who agreed were interviewed via their preferred method of video chat (n = 19) or phone (n = 1). Interviews lasted on average 60 minutes (SD = 13.8 minutes, range: 40–88 minutes) and covered many topics, such as gender transition and disclosure experiences on social media sites, that are not covered in this paper but discussed in our other work [37]. We compensated each interview participant with a $25 Visa gift card.

Interview participants were 50% trans women, 35% trans men, and 15% non-binary trans people. Participants were 65% white, 15% Black, 15% Asian, and 10% Hispanic/Latino/a/x (percentages add up to greater than 100% because some participants were of multiple races/ethnicities). The average age was 26.65 (SD = 7.02 years, range: 19–43). A total 18 participants were American (one residing
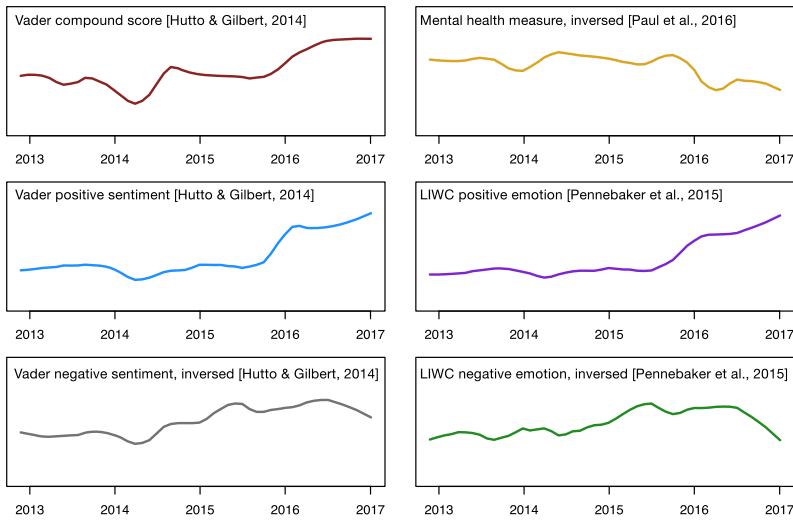
Fig. 1. Example sentiment graph (for P4), with sentiment measures labeled. ©Oliver Haimson.

in Europe at the time of the interview) and two were Canadian. Participants had been blogging on average for 709 days (*sd* = 589, range: 189–2,305 days), began blogging between 2012 and 2016, and were active bloggers up until 2016 or 2017 (when interviews were conducted). Participants had posted on average 361 posts (*sd* = 635, range: 18–2,617 posts).

Prior to each interview, using computational sentiment analysis methods, we analyzed and visualized the interviewee's textual Tumblr transition blog content over the time period they had maintained their blog. We used Tumblr's API [99] and the PyTumblr API client [76] to collect this data, an approach that appeared to be allowed according to Tumblr's API License Agreement as of January 2017 [98]. Tumblr blogs can include many types of content, but for this article we only analyzed participants' text posts and photo caption content (if it was more than 10 words long) because we found that these were most likely to include meaningful personal text. We used six different sentiment analysis measures to visualize participants' textual blog data over time (see Figure 1). Some were created particularly for analyzing social media text [47], while some were created for text analysis more broadly [73]. See Appendix A for each sentiment measure's descriptive statistics. Future work could examine whether displaying only one or several graphs, rather than six, may be a better approach.

— Vader compound score—a composite, unidimensional score of sentiment (red) [47];
— Vader positive sentiment (blue) [47];
— Vader negative sentiment, inversed (grey) [47];
— A negative mental health measure using Urban Dictionary words but verified to work across platforms [71], inversed (yellow). This turned out not to be a very good indicator of emotional wellbeing and gave somewhat random graphs for most participants;
— LIWC positive emotion (purple) [73];
— LIWC negative emotion, inversed (green) [73].

We used R to visualize these sentiment plots. The sentiment measures had different ranges, so we normalized each and plotted them so that they would appear on the same scale. This way, only patterns and deviations in the graphs would stand out. We inversed the negative measures (Vader negative sentiment [47], LIWC negative emotion [73], and the negative mental health measure

from [71]) so that on each graph, positive sentiment was represented by higher vertical values, and negative sentiment was represented by lower vertical values. We plotted the graphs using the R function Loess.smooth [25], with a one degree local polynomial, a Gaussian fitting, and the smoothness parameter varying depending on the amount of time the person had been blogging. Adjusting the smoothness parameter required tradeoffs between granularity and graph interpretability. We prioritized graph readability, which meant a relatively wide smoothing window, particularly for those who had been blogging for years. This emphasized and elicited reflection on more long-term sentiment trends and larger sentiment changes rather than short-term variability and day-to-day or week-to-week fluctuations. Future research should examine how smoothing parameters impact participant reflection. The X-axis varied for each participant, and displayed the full amount of time that person had been blogging on Tumblr (thus, the labels designate years for some participants, and months for others. We designed the sentiment graphs carefully to be legible to audiences unfamiliar with scientific visualizations. We chose to use simple line graphs because they are interpretable to most people as a familiar way of representing change over time [32]. We initially had designed sentiment graphs that were more complex, but simplified them substantially after pilot testing them with a small group, and settled on the design shown here before displaying graphs to participants. We intentionally did not include labels differentiating the plots, because we wanted participants to reflect on each without knowing which particular measures they represented. We did not include Y-axis values on the graphs because the sentiment measures would not have been meaningful to participants. We described to each participant that the plots showed six different ways of measuring an approximation of emotional wellbeing over time, in which lower values and dips represented lower emotional wellbeing, while higher values and peaks represented higher emotional wellbeing.

Toward the end of each interview (after having interviewed the participant for roughly 40–60 minutes) we shared each participant's individual sentiment graphs with them, told them that the graphs were generated using emotional wellbeing scores calculated from their Tumblr blog's text, and asked them to reflect on the graphs. This timing is in line with recommendations that visual elicitation usually works best toward the end of interviews [21]. For video chat interviews, we shared our screen with participants and showed them the graphs. When there were technical difficulties (n = 1) or with the participant who preferred phone (n = 1), we instead e-mailed the sentiment graphs image to the participant. Questions began with a variation of "Do any of the lines seem to resonate with how you were feeling over time?" After this, questions were highly tailored for each person (for an example dialogue, see Section 4.3) and focused on patterns in the particular sentiment graphs. This study was in the context of gender transition, and thus sentiment graphs tended to follow certain patterns, such as a generally positive trajectory over time as people proceeded through their transitions. For other types of life transitions, researchers must be aware of and sensitive to the types of sentiment patterns that may occur, so that they can understand common patterns vs. outliers, and be prepared for negative emotions that may arise when participants view less positive sentiment trajectories (which we discuss more in Section 5.1). This method should be further tested in other life transition contexts.

As detailed in Section 4, though initially we had a goal of understanding, which sentiment graphs were most accurate, assessing accuracy was a difficult task for participants, and interviewees instead tended to reflect on past experiences. Thus, after several interviews we shifted to a goal of reflection: using the sentiment graphs primarily as tools for eliciting reflection on past identity transitions and people's emotions around those experiences. It is important to note that sentiment analysis methods have accuracy limitations [8], and sentiment visualizations involve simplifying complex transition journeys into depictions of positive and negative emotions. Despite these

limitations and complexities, which we discussed with participants, we found the method valuable for eliciting participant reflection.

Interviews were audio-recorded and transcribed. We analyzed interview data using open coding and line-by-line analysis, in which we allowed codes to emerge from the data [88]. We then examined connections between codes, and organized the codes into larger themes [88]. In the analysis presented in this article, the three emerging categories we focus on related to applying this method are (1) sentiment graphs as a reflective tool, (2) sentiment graphs help participants recall new data, and (3) participants' difficulties relating to sentiment graphs. We then grouped these under the larger category of how participants responded to the sentiment graphs as an elicitation tool.

4 RESULTS

We present a description of our experiences applying this method to support our argument that sentiment graphs of social media data can be a useful tool that researchers can use to enable interview participants to speak about and reflect on important events from their pasts. We first describe how sentiment graphs enabled participants to reflect on their personal histories and the emotions surrounding their past experiences. We next detail how participants recalled and described new data corresponding to patterns in sentiment graphs. Such novel data were related to topics or people that the participant had not previously mentioned during the interview, and that the interviewer would not have known to ask about. Finally, we describe some participants' difficulties in connecting the graphs with their own experiences. Interestingly, even participants who had trouble relating to the graphs were able to use them as a tool for reflection.

4.1 Reflection on Personal Histories Using Sentiment Graphs

Most participants (16 out of 20) were able to make connections between the sentiment changes they saw in the graphs, and particular identity changes they had experienced during those time periods. Thus, longitudinal sentiment graphs of social media data can serve as a reflective tool for uncovering personal histories.

Given that this study focused on gender transition, many participants drew connections between the sentiment graphs and their gender transition journeys, particularly with respect to hormone replacement therapy. Several participants noted a connection between increased sentiment in their graph and starting hormone replacement therapy, such as P14, who said, "and then in October, I was really excited. I finally started hormone therapy." P18 described a period of time when she temporarily lost access to feminizing hormones, which corresponded to the decrease in positive sentiment shown around November in most of the graphs in Figure 2.

The sentiment graphs related to not only hormone replacement therapy, but many different types of life experiences. In addition to the impacts of hormones, P18 made several other clear connections between experiences in her life and the sentiment graphs we showed her: "I know that right before Thanksgiving, that was when I officially dropped out of school, at least for the rest of that semester. So, I was really down in the dumps right around that time." The holiday season also impacted her mood: "I know that right around December, it went up a little bit, as it always does for most people, just because of the holidays and everything else." For these reasons, she stated that she considered the grey and green sentiment graphs to be "good indicators…when it came down to my emotional wellbeing."

While each of the examples in the section above involve reflection, some participants used the graphs to go beyond simply corresponding patterns in the graphs to past experiences, and further reflected on those experiences in meaningful ways. In this way, we used sentiment graphs over time as a tool that enabled people to reflect on their personal histories and the emotions that came up for them around life experiences. While we do not present quotes from all participants, for
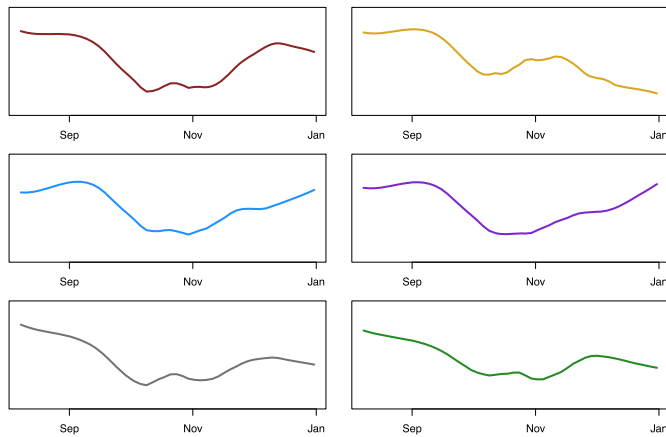
Fig. 2. P18's sentiment graph. Top left (red): Vader compound. Middle left (blue): Vader positive sentiment. Bottom left (grey): Vader negative sentiment, inversed. Top right (yellow): mental health measure, inversed. Middle right (purple): LIWC positive emotion. Bottom right (green): LIWC negative emotion, inversed. ©Oliver Haimson.

13 out of 20, visualizations prompted substantial reflection. P20 used the graphs as an opportunity to put herself back "in that place" in time and reflect on how she was feeling during a particular time period:

> *Okay, so that was second semester into my sophomore year. Let me just put myself in that place... At the time, I was at a super high in my life. I felt like I was kind of taking over the world. That's when my blog was starting to take off... [referring to a dip in the graph] I'm not sure what a dip would be. I did have to go to my brother's graduation and kind of present in this weird, ambiguous, gender neutral presentation at my school that I go to. That was a big low. After that, everything was great.*

This is an example of the sentiment graphs' ability to enable reflection for participants in a way that traditional interview techniques and existing elicitation techniques often do not. "Putting people back" into particular periods of their lives to reflect is a difficult endeavor, and sentiment graphs can help.

Several of P3's sentiment graphs showed an increase in positive sentiment throughout the second half of 2016 leading into 2017. P3 attributed this increase in positive sentiment to her process of coming out as trans and starting on hormone replacement therapy:

> *At least a majority of those graphs you can see...it looks like I'm at a low and then I come out and essentially... that line starts skyrocketing up faster than any other growth than any of those lines. [referring to the yellow line] Forget that one, it doesn't matter. The other ones they just kind of jump up and that's kind of how I felt after coming out. like it took a lot to get there, I was nervous and I was scared and didn't know what to do and then I did it. And that smile [increasing pitch sounds] kind of just grew on me... I think that line [dip] in the middle is probably where I got hormones, I got that like September, I think. Yes, seven months yesterday. It feels amazing.*

Interestingly, P3 was able to easily disregard the yellow graph that did not support this positive trajectory ("Forget that one, it doesn't matter."), which turned out to be a correct assumption on her part given the yellow graph's lack of relationship with positive sentiment for most participants.

For other participants, the sentiment graphs brought back up past feelings of anger, distress, and jealousy. P2 reflected on the complex feelings, both negative and positive, that he had experienced in the past year related to hormone replacement therapy:

> *This should be this year then... A lot of my earlier [blog] posts were a lot more 'I need to start transitioning because I'm really upset.' You know? A lot of guys, and probably girls, I don't know, experience a sense like you see other people transitioning and you get extremely jealous. And you feel bad because I should be happy for you but I'm actually fucking pissed. So, I think, at the start of my blog that was definitely were I was. Then, I had that, 'Oh my god! It's happening!' And then it probably evened out a little bit.*

Thus, most participants were able to connect their past experiences to sentiment graphs of their social media data. Furthermore, the visualizations helped place participants back into particular periods of their lives. The majority of participants used sentiment graphs as a tool for substantial reflection.

4.2 Sentiment Graphs Help Participants Recall New Data

Sentiment graphs often brought up experiences that participants had not yet mentioned in the interview. Importantly, some participants had not written about these experiences on their blogs, and hence the researcher would not have known to ask them about these events. That is, sentiment graphs helped participants recall and reflect on new data that they had not thought to bring up before viewing the graphs. These new revelations were likely related to the graphs' visual depiction of time, as evidenced by participants' descriptions of events in relation to the graphs' visual patterns and the related time periods.

We had been interviewing P14 for almost an hour when we showed them the sentiment graphs. For P14, the graphs became a means to reflect on their diagnosis with polycystic ovarian syndrome (PCOS).

> *I feel like the gray and the light-blue graph [are representative of how I was feeling] because there were times in this period where I was getting a little bit more worried because of the tests I was going through and finding out that I have PCOS. And I was worried about like what might happen, if I wouldn't be allowed to go on testosterone, or those kinds of things. And there was also concerns of if I'm more likely to have cancer due to having PCOS, or the fact that my mom has the CHEK2 gene. So, I had to get tested for that. But I ended up being negative for that, but I was like afraid that I might not be able to actually start hormone therapy. And so, I'm wondering if these dips, especially these dips around October, were from me voicing those concerns [on my blog]...*

Surprisingly, this major health diagnosis that had caused P14 substantial distress within the past six months had not come up during the hour-long interview we conducted prior to showing them the sentiment graphs. We would not have known to ask P14 about PCOS, and this data would not have been elicited without the graphs.

Participant P5 noticed that their sentiment graphs all showed a decrease the previous April, and stated that *"April was a sad month for me last year."* When asked to elaborate, they described,

> *I was coming out and my great grandmother passed, and I was dating somebody and it was all bad, so that's so funny you can see that in the graphs. Yeah, actually, looking at the months I can see oh yup there's the dip, there's the recuperation and there's the trail off where I just disappeared for a couple months.*

Interestingly, while P5 did write about some experiences (e.g., coming out) on their blog; other experiences, such as their great grandmother's death, were not mentioned at all. It might be that

the computational sentiment measures picked up P5's negative affect during that time period when they were writing about other topics. It might also be that P5 used the sentiment visualizations to find connections to experiences during that time period. Either way, the sentiment graphs enabled us to elicit important information about difficult experiences that P5 had not mentioned until this point in the interview, after we had been talking for roughly an hour.

P4 also found connections between patterns in the sentiment graphs and recalled experiences that she had not previously mentioned. Responding to a period of positive sentiment shown in the graphs, P4 described positive, yet complicated, relationship experiences:

> *My wife who is now my ex-wife [and I] started a poly-[amorous] relationship and that went really well for a good bit of the year. It kind of all exploded in September that's why that dip there might be telling, and after that I met a new person and right toward the end of the year things started looking up again.*

While P4's sentiment graphs showed a positive sentiment trend during this time period, each of the six graphs differed slightly, and none visualized the level of complexity P4 described in her narrative. Accuracy notwithstanding, the sentiment graphs helped shed light on new data.

Sentiment graphs enabled reflection even for participants who found it difficult to interpret the graphs (we describe such difficulties in Section 4.3). For participant P6, the graphs brought up reflection on a period of depression that they had not mentioned previously in the interview: "I mean I guess the green one stands out to me in terms of I definitely was really depressed during that, during the year before I started going on testosterone or maybe a little bit before that." Similarly, viewing sentiment graphs reminded P2 of relocating to a new country and working full-time.

> *I think some of it might be impacted by the fact too, over the summer I was really busy – I worked like 40 hours a week at [coffee shop]. On the other hand, I worked as a guy, which was very affirming for me. But I was working 40 hours a week in customer service. Then, around September/October/November-ish I was here, so I just came to [European country]. And I think that accounts for the first months. You know when you show up in a new place, you're kind of not familiar with everything, you're away from your parents, from your family, and your friends. You're kind of in a new place and in this case, it's like my second language, so you're kind of alienated from everyone because I can go out and talk to people or I cannot because I'm very self-conscious of how I sound in [European language] or something.*

Thus, sentiment graphs helped participants recall new data and reflect on experiences that the researchers would not have otherwise asked them about. This was true even when participants had difficulty interpreting the graphs.

4.3 Participants' Difficulties Relating to Sentiment Graphs

This section describes the difficulties we encountered when showing sentiment graphs to interview participants. In total, seven participants (out of 20) initially had difficulty relating to the sentiment graphs, but upon further discussion and explanation, most were able to use the graphs to reflect on past experiences to some extent.

When developing this method, our initial goal was to determine which of the six sentiment measures was most accurate and most resonated with people's experiences. We soon learned that assessing accuracy via line graphs over time was a very difficult task for most participants. These initial difficulties enabled us to dig deeper, and ultimately use sentiment graphs as a tool for *reflection* rather than accuracy.
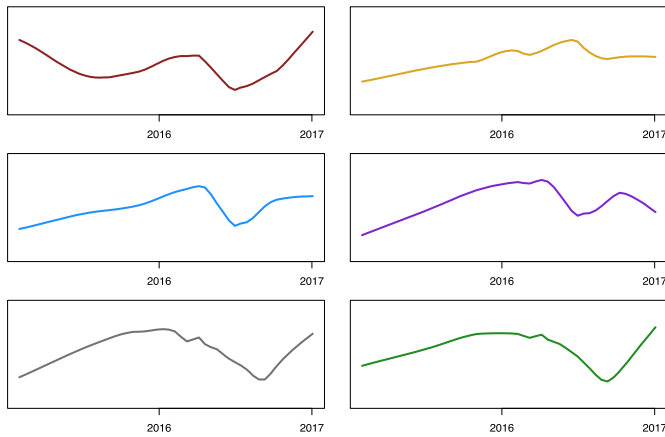
Fig. 3. P3's sentiment graph. Top left (red): Vader compound. Middle left (blue): Vader positive sentiment. Bottom left (grey): Vader negative sentiment, inversed. Top right (yellow): mental health measure, inversed. Middle right (purple): LIWC positive emotion. Bottom right (green): LIWC negative emotion, inversed. ©Oliver Haimson.

P3 initially had difficulty matching the experiences she remembered with the sentiment graphs we showed her. We drew her attention to the "dip" in positive sentiment that seemed to have occurred in mid-2016 (see Figure 3), and asked her to describe what might have been happening in her life during that time. However, she remembered having experienced negative emotions not in mid-2016, but instead in December, calling the graphs' accuracy into question.

> *Interviewer: [describing graphs in non-scientific terms] The higher the line is, that means that represents feeling better. When it goes down – in some of the lines there's a dip in 2016 – that would represent not feeling as well...*
>
> *P3: So would that dip be roughly December, that big dip?*
>
> *Interviewer: The dip actually looks more like July or August. Was there something in December that happened?*
>
> *P3: Yeah, December was when I got kind of disowned from my family, so I know I kind of went off the deep end there a little bit. Not all, just my Mom's side pretty much. They're all die-hard Jehovah's witnesses and they did not tolerate that very well. You say that dip is around July, well that's crazy.*
>
> *Interviewer: Yeah, it sounds like if things were kind of rough in December, these graphs aren't doing a great job of representing that.*
>
> *P3: To a degree, I had just come out too, so in that sense, I was really happy with myself. I mean I had my own crap I had to deal with but I might not have led on. I don't know I kind of bottle things up inside 'cause I don't want people to know too much. But, yeah look at these graphs, they're so weird, they all seem to follow the same pattern.*
>
> *Interviewer: Yeah, they are pretty highly correlated. So, they are different, slightly different... I know that this is a difficult question to even answer, so I'm not sure if it is really possible to be able to tell. But, do any of these resonate with you more than others?*
>
> *P3:* **It is a difficult question. You're asking me if that squiggly line represents how I feel.**

This excerpt brings up a fundamental complexity that arises when using sentiment graphs to determine the accuracy of sentiment measures: It is difficult, if not impossible, for many people to
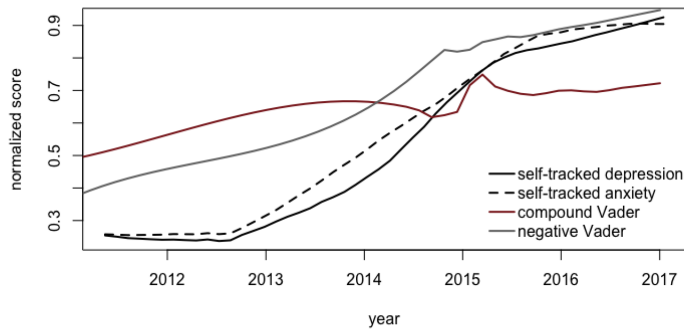
Fig. 4. P6's self-tracked sentiment and sentiment as measured from Tumblr blog posts. ©Oliver Haimson.

connect their feelings over a long period of time with a "squiggly line" on a graph. P3's interview represents a turning point in our study where we shifted from trying to understand which graphs were most accurate for measuring emotional wellbeing over time, and instead began to use these graphs as a reflective tool that helped people talk about their experiences from the past and how those made them feel. For example, despite our conversation excerpted above, P3 was able to reflect on her experiences well using the graphs as a starting point (see Section 4.1 for example quotes).

Other participants described being "bad with numbers" or unskilled at making sense of information visualizations. On surface level, this method would seem to be a bad fit for participants like these. For example, P6 stated that they were "kind of bad with numbers," "bad with dates," had to "think really hard about what was happening during those years," and described that "when it comes to sequence and time, I can tell sequence but I can't tell time." Timeline-based visuals are not always successful, as some people do not relate well to time being visualized in linear ways [6].

Despite or perhaps because of these difficulties with dates and numbers, P6 had a strong interest in self-tracking their emotional wellbeing over time, and actually kept their own spreadsheet listing their emotional state over the past six years:

> I broke it down into semesters... So, Fall, Spring, and Summer... And it tells me what my job was during that time, it tells me what work I was doing during that time, it tells me stuff about my mental health, what medication I might have been taking for mental health. If I was seeing a therapist, which one... If there was some major issues that me and my partner were facing at that time.

P6 volunteered to share this spreadsheet with us and allow us to analyze their self-tracked data, and we found that P6's self-tracked measures of depression and anxiety correlated somewhat with the Vader negative sentiment [47] ($r = .75$, $p = .02$ with self-tracked depression, $r = .72$, $p = .03$ with self-tracked anxiety) and Vader compound score [47] ($r = .59$, $p = .10$ with self-tracked depression, $r = .64$, $p = .06$ with self-tracked anxiety) measures from our sentiment analysis of their Tumblr blog, even with a very small sample size ($n = 9$ time periods where there was enough data from both sources). Figure 4 displays these measures over time for P6. Interestingly, despite the Vader measures being most highly correlated with their self-tracked measures, P6 thought the LIWC positive and negative emotion measures resonated more with them when they looked at the graphs during our interview (see Figure 5).

P6's example shows that even when sentiment visualizations are somewhat accurate according to a ground truth (self-tracked data), some participants may have difficulty making sense of sentiment graphs. However, soon after saying that they were bad with dates, P6 was able to
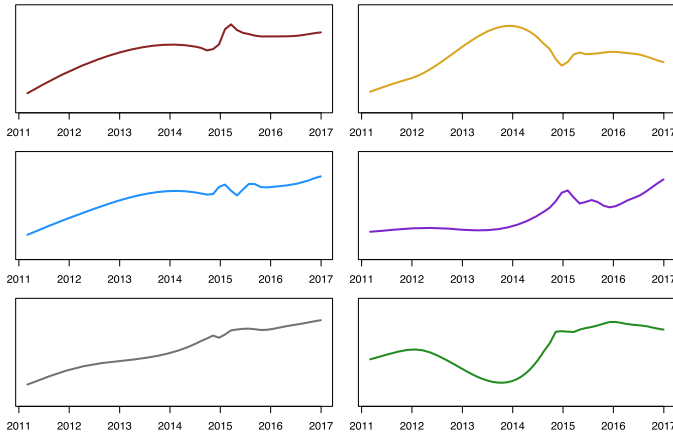
Fig. 5. P6's sentiment graph. Top left (red): Vader compound. Middle left (blue): Vader positive sentiment. Bottom left (grey): Vader negative sentiment, inversed. Top right (yellow): mental health measure, inversed. Middle right (purple): LIWC positive emotion. Bottom right (green): LIWC negative emotion, inversed. ©Oliver Haimson.

recall and reflect on a period of depression they experienced before starting hormone therapy (see Section 4.2). Thus, sentiment graphs may be useful for eliciting reflection even when participants express difficulty interpreting the graphs.

Another participant, P13, stated, *"I feel like I'm looking at random lines. I'm sorry."* Yet he still used the sentiment graphs to reflect on experiences like spending his first summer away from family, and starting college.

> P13: *I was looking at the blue [graph]. Yeah, the blue one, and I was like that kind of looks like me, because just remembering what I was going through in July and how things did get better once the school year started and such.*
> Interviewer: *What was going on over the summer?*
> P13: *Just like relationship issues and I'm – like I said, I go to college. I keep saying this, but I go to college in the middle of nowhere, but I'm from southern California where there's everything, basically, and so, last summer was my first summer away from my family. The entire summer, so it was just kind of hard for me not to be able to be around them.*

As with all elicitation methods, this method works better with some participants than with others. P2, after considerable reflection based on sentiment graphs (see Sections 4.1 and 4.2), ended with a caveat about his limited ability to interpret the graphs: *"But I don't really know that I can derive much from these from these sorts of things or these sort of symbols."* This apparent contradiction shows that sentiment graphs can be a useful tool for reflection and elicitation, *whether or not* they are particularly accurate and *whether or not* the participant fully understands the graphs.

5 DISCUSSION

"History from below," also known as "people's history," is a way of constructing historical narratives by drawing from the experiences of everyday people, especially those who are marginalized and oppressed, rather than from those in power or leadership positions [13, 91]. When considering personal experiences with identity transition, our approach can be considered a *transition history from below,* which emphasizes centering marginalized people's experiences and posits them as the experts of their narratives. A contrasting approach would be to draw from "experts" such as

clinicians and medical or mental health professionals who specialize in the types of identity transitions one wants to study. For example, a long history of research drew false conclusions about trans experiences because its data sources were medical professionals who worked with trans populations instead of trans people themselves [24, 87, 89]. In the research approach we describe in this article, we consider people to be the experts of their own experiences, and we show how to combine multiple data sources—participants' historical social media data combined with interviews in which they reflect on this data—to understand identity change experiences.

We construct a gender transition history from below based on interviews with a diverse set of trans participants. This approach to constructing history enables the inclusion of insights that are important to the lived experience of transition, such as the physical and mental health complexities and temporal intricacies of gender transition. Participants described how complex intersecting medical needs could have significant impacts on their mental and physical wellbeing during transition: The discovery that a newly diagnosed health condition (polycystic ovary syndrome) would restrict access to gender-affirming hormone replacement therapy caused P14 substantial distress (Section 4.2). Participants also reflected on the mental health impacts of their lack of agency over gender presentation at various points in their transition: having to attend her brother's graduation using an ambiguous gender presentation was a significant low for P20 (Section 4.1). Furthermore, multiple participants described how the lack of access to gender-affirming medical treatment impacted their wellbeing: P2 and P6 experienced depression, anger, and even jealousy in relation to their own and other people's experiences with hormone replacement therapy (Sections 4.1 and 4.2). Participants also described experiencing multiple intersecting life transitions during the same period of time: a family member's passing away, a bad relationship, and publicly disclosing their trans identity were all an inextricable part of P5's gender transition, with overlapping mental health consequences. A transition history from below thus complicates mainstream narratives that portray gender transition as a linear and straightforward journey of physical transformation.

"Putting people back" into particular periods of their lives to reflect is a difficult endeavor, and this difficulty can limit the retrospective study of long-lasting identity transitions (e.g., gender transition, health status changes, and relationship changes) [50, 65, 101]. Showing interview participants sentiment graphs of their social media data over time is a promising new approach for eliciting reflection on participants' identity transitions. While existing methods enable participant reflection in many important ways, our method combines sentiment analysis and visualization of longitudinal personal data to uncover unique insights from participants related to their past experiences and identity changes. We showed how we used this method in the context of a study focused on one particular type of identity transition: gender transition. Yet each of the participants in our study also experienced other life experiences and transitions during their gender transition processes, and sentiment graphs enabled us to elicit reflection about these myriad, intersecting identity changes. Researchers will find this method valuable for other populations, particularly those who have faced difficult experiences in their pasts.

Previous identity transitions work has typically used interviews and surveys to understand people's past experiences with changing identity (e.g., [26, 37, 60, 69, 83, 85]). These studies are often focused on one type of identity transition in isolation, enabling the interviewer to specifically ask about that particular experience, as we did about gender transition. Our sentiment visualization elicitation method helps resolve two challenges facing such studies. First, recalling patterns of emotion over time is often difficult in interviews [21]. Second, identity transitions do not occur in isolation, and people may not remember, or think to bring up when asked, other things that were happening in their life at that time. By visualizing longitudinal emotion patterns without naming specific experiences, our method helps participants reflect on their experiences in the context of

the emotions surrounding them—and further, recall other life experiences that intersected with their identity transition (e.g., P5's great grandmother passing away).

Our method provides participants with a visual tool to focus on the local and specific aspects of their identity transitions. Narrative identity researchers have found that people's lives are influenced by cultural norms and scripts that demarcate the milestones and themes that are considered appropriate and positive during particular phases of life (e.g., falling in love during adolescence) [5, 18]. Events and processes that are non-normative (such as gender transition) or traumatic are typically left out of mainstream cultural scripts. People are socialized to believe that they must adopt one of "a handful of dominant, broadly acceptable scripts of 'master narratives' " for coping with non-normative and traumatic occurrences, if they want to obtain acceptance and approval for their life stories [5, 93]. Oral history researchers found that listeners were more accepting toward life-threatening events if narrators spoke of their bravery or concern for others, as opposed to their feelings of fear and sadness [93]. Transgender people and people who have undergone stigmatizing or traumatic life transitions are expected to perform a similar set of master narratives that are dictated by "experts" such as medical professionals, to obtain acceptance and access to life-sustaining resources [24, 82]. In the context of interview research, this may mean that marginalized participants also feel the pressure to conform to these master narratives. By presenting participants with visualizations based on their own social media data, our method provides participants with a potential alternative to mainstream cultural scripts. This alternative script may encourage participants to move away from master narratives and delve instead into their own idiosyncratic and unique life journeys. By drawing from oral history and narrative identity research, we can make this important methodological contribution to HCI.

Prior work on qualitative interview elicitation methods has found that showing participants their social media *posts* from the past elicits reflection on those particular events and posts [72, 79]. We find that showing participants sentiment *graphs* of longitudinal social media data provided participants a new way of viewing their pasts and helped participants reflect on their experiences at a higher level and with a broader focus than looking at the posts themselves. This is likely because of the aggregate nature of the graphs (versus seeing a single post at a time) and because the graphs displayed emotional patterns without mentioning the specific posts corresponding to them. This helped elicit new data about experiences that participants had not mentioned on social media (e.g., P14's health diagnosis) and that the interviewer would not have known to ask about, thus uncovering personal histories and building transition histories from below.

Some of our findings overlap with prior research on using personal data for self-reflection. For example, participants in Choe et al.'s study interacted with a system called Visualized Self, which pulled data from multiple self-tracked sources to visualize patterns over time [17]. Choe et al. found that participants used the system for meaningful reflection, and many of their insights had to do with past identity changes, such as a job change or becoming a parent [17]. Like our participants, participants in Choe et al.'s study too found "peaks or extreme values" in the visualizations particularly useful for reflection [17]. Though Choe et al.'s study was not focused on interview elicitation, graphs of self-tracked data helped participants bring up life experiences that were external to the data being tracked. In our method, sentiment graphs of social media data serve a similar reflection purpose: Graphs were not a way to make life experiences more organized or accurate; instead, by making our participants think with emotion patterns and on occasion doubt their accuracy, sentiment graphs enabled interviewees to think seriously about their pasts and reflect on their experiences within the context of the emotions surrounding them. Indeed, and of particular importance for an interview elicitation method, this helped participants recall new data that they had not mentioned on social media and that we would not have known to inquire about.

Our method promotes collaborative reflection in the interview process, in the sense that researchers presented participants' data over time as prompts, along with interview questions, to collaboratively enable reflection. The collaborative reconstruction of people's stories and narratives with researchers is a part of any interview, and we show that this process may be aided by graph-based elicitation methods. Had they not seen the sentiment graphs, people may not have thought to remember or reflect on particular past experiences in detail, or at all. Prior work has shown that collaborative reflection and sensemaking can benefit people (e.g., improved emotional wellbeing [48], finding meaning and insight in past events [14, 17, 30]); and using our sentiment visualization elicitation method during interviews could provide participants some of these benefits without the burdens (e.g., having to track data themselves, interpreting graphs on their own without explanation [31, 57]). Yet the experience of reflecting with a researcher in an interview context is quite different than self-initiated reflection, and may even generate different narratives or recollections. This brings up open questions around how understanding the past actually happens in research interviews, and how it may be different than other self-reflection processes, an exciting area for future work. Importantly, though our method was collaborative, it was not as collaborative as it could be. For example, future work could involve participants more in the data analysis and visualization process (in line with Thudt et al.'s [95] argument that people should be given a more active role in visualization design processes), perhaps enabling them to choose which data points to include and exclude in the graphs, and even entering additional data points that may not have been present in their social media dataset. Though some participants may feel overwhelmed or unequipped to analyze and visualize their data, for some participants, a more collaborative process would likely invoke even more reflection.

5.1 Ethical Considerations

We took great care to conduct our study ethically and in a way that did not harm participants. All of the participants we interviewed for this study explicitly opted in to data collection and the data we collected were technically public (though much of it was private by obscurity or "publicly private" [53]). Another ethical concern involved whether it was advisable to show sentiment graphs to participants whose sentiment showed a substantial downward trend over the long term or the recent short term. Interviewees in our sample did not have long-term downward sentiment trajectories, likely because gender transition trajectories show an upward slope on average [38] and a bias in our interview sample was that those people who are most distressed may have been less likely to respond to our interview requests. That is, those experiencing depression or mental distress (which is prevalent among trans populations [49]) may not have had the emotional bandwidth to respond to a Tumblr message from a researcher, or to spend emotional energy on an interview. Showing participants negative trajectories may bring up negative emotions and it may not be advisable to use this method. At minimum, researchers should be trained in witnessing and responding to difficult participant emotions in a sensitive and reflexive manner (e.g., [66]).

Making connections to experiences and identities in the past can be difficult for some people, particularly those who have experienced difficult identity changes [40]. It is important for researchers to consider that some interview participants may not want to reflect on experiences or emotions from the past. This research method also requires caution because it involves showing visualized sentiment data to people who may not have otherwise chosen to view such data. There are important differences between tracking oneself and having one's data collected by a researcher. Most research in which people view their visualized emotions over time [17, 57] focus on people who *want* to self-track and view their own data. Yet with the current method, while participants have consented to data collection and interviews, they generally are *not* people who have intended to track their emotions over time,neither examine nor reflect on sentiment patterns or their pasts. In

our study, we specifically informed participants in advance that we were collecting and analyzing their blog data (but not that we were visualizing it), and at the beginning of the interview we told each participant explicitly that they could exit the interview at any time or skip any question if they were not comfortable answering. None of our participants mentioned feeling uncomfortable with viewing their sentiment visualization graphs. Yet researchers using this method should be aware that participants may feel discomfort, and researchers must approach the method with sensitivity, allowing participants to opt out at all stages.

5.2 Limitations and Future Work

This research and method involve several limitations. Performing video interviews allowed us to reach a larger set of people (given the culturally stigmatized nature of gender transition). However, this meant that we could not be physically present with participants while they viewed sentiment visualizations. Future work should examine whether co-presence of participants and researchers affects the elicitation potential of sentiment visualizations. Next, it may be necessary to perform graph literacy before using this method in interviews, because if participants cannot interpret the graphs well, then their reflections may suffer (though, as we found, in some cases this method can lead to reflection regardless). As another limitation, though the sentiment graphs enabled people to reflect on past experiences and think about the emotions surrounding those experiences, this method did not actually *recapture* emotions that participants felt in the past. Additionally, participants may have limited retrospective recall. Some may post-rationalize their narratives in light of the visualizations that they viewed, and describe their past experiences differently than they otherwise would, potentially in a skewed manner. Participants also may focus more on peaks and dips, rather than steady curves, since the former stands out more. Relatedly, sentiment of social media data likely does not directly indicate a person's actual emotional wellbeing over time, since people may post overly-positive or overly-negative content online. Additionally, inviting participants to create a personal narrative that coheres with a quantified representation of their data may have been leading for some participants; however, many were quick to disregard visual patterns they did not consider meaningful. Next, sentiment analysis is known to have accuracy limitations [8], which may impact results. We did not rigorously test the visualizations' accuracy, and how "correct" the graphs were was not transparent to us or to participants. Thus, as a limitation and ethical issue, it is possible that when using this method participants will answer in ways that they think the researchers want to hear (e.g., stating that a "dip" in the graph corresponds to a life event). This may be especially likely to happen when researchers are in relative positions of power compared to participants from marginalized groups, and researchers should carefully consider and mitigate this possibility. Accuracy issues would be especially problematic if this method were used in medical contexts (e.g., to confirm mental health diagnoses), which we do not recommend. Finally, especially skilled researchers are able to elicit substantial reflection without the use of sentiment visualizations, and this method may not add extra value in all interview settings. Future work should also investigate how randomly-generated graphs perform in comparison to sentiment graphs. Furthermore, this method could be attempted with other types of graphs that may work well to capture emotion, such as emotion glyphs, or interactive visualizations that overlay social media data on graphs and allow users to zoom into and out of time periods. Future research could also examine this method's efficacy with populations experiencing additional types of identity transitions.

6   APPLYING THE RESEARCH METHOD

This research method will likely be useful in studies that require participants to reflect on their pasts, especially for people who have experienced substantial identity changes. It would not work

well for studies focusing on participants' current experiences or evaluations of particular systems. Researchers who wish to use this approach in the future can apply the following steps:

6.1 Data Collection

The first step is to gather participants' data, with their permission or using an opt-out approach. This method was designed for and evaluated on textual social media data, but it could be extended to other types of personal data repositories, such as personal diaries (journals), quantification of wellbeing over time that has not yet been visualized (e.g., P6's self-tracked sentiment data), and potentially even visual social media data (e.g., photos, videos). Choosing a social media site where researchers can access data is vital; some social media sites (e.g., Tumblr, Twitter, and Reddit) allow data collection via APIs, while others (e.g., Facebook, Instagram) do not. The research site must also be a space where participants share substantial personal data over time. Thus, sites like Facebook and Twitter may not be ideal for this method because text posts tend to be shorter and less personal and in-depth. You may also ask participants to provide their data to you, such as by downloading their Facebook data archive. The method will work best with sufficient data to create meaningful sentiment patterns over time, which will depend on the time duration (e.g., several months vs. several years), but in our experience included at minimum about 10 data points total and two data points per month for each person.

6.2 Data Analysis and Visualization

The researcher must choose a sentiment analysis tool (e.g., [47, 73]) that works best for their research context, and generate sentiment measures for each temporal data point. Then, the researcher visualizes the data over time using a simple line graph, or a more complex visualization method of their choice. Though simple line graphs worked well because they were lightweight and easily shareable and explainable to participants, future research could explore more involved visualization methods, such as interactive visualizations that link to participants' particular social media posts and display their sentiment scores. The visualization will generally need to use a smoothing function (e.g., Loess.smooth in R [25]) to improve visual appearance and interpretability. Researchers do not need to use six graphs as in the study presented here, but should use more than one to show that sentiment analysis and visualization are subjective and there is no one "right way" to visualize a person's sentiment over time. The researcher must pilot test their visualizations with several people to make sure that they are clear and easily understandable.

6.3 Use in Interviews

We recommend showing participants the sentiment graphs toward the end of the interview (as also recommended by [21]), to avoid skewing the interview toward certain time periods, experiences, or emotions. The sentiment graphs are then used to enable further reflection and fill in gaps. The researcher should describe sentiment analysis in non-scientific language and describe how to interpret the graphs (e.g., "The peaks may represent a time during which you were feeling good, and the dips may represent a time when your emotional wellbeing was not as good."). After these explanations, the researcher can start the reflection process by asking, "Do any of the graphs resonate with how you were feeling over time?" The researcher should then tailor their questions to the graphs and participant, such as by pointing out particular patterns and asking the participant to explain what might have been happening during that time period, and relating topics that came up earlier in the interview or in the participant's textual data. The activity can take anywhere from a few minutes to 30 minutes or more depending on how much reflection the graphs elicit and how easily the participant interprets the graphs. Researchers may also find value in pairing this method with other elicitation methods, like showing participants specific social media posts

from the past, which may add additional valuable information to prompt reflection. Researchers should carefully consider the ethical considerations described in Section 5.1 before employing this method. It is important to note that some participants will not be good at reading or interpreting graphs, and that this method may not elicit reflection from every participant.

6.4   Analysis

Interview transcripts should be analyzed with the visualizations to understand participants' experiences in depth. General qualitative data analysis techniques can be used (e.g., [88]).

6.5   Potential Difficulties

As happened in the study that we describe here, some participants may have difficulty interpreting the visualizations, or may have trouble thinking about time visually or linearly [6]. These participants may still reflect on their past experiences, but those reflections may not correspond well to the graphs they are shown. However, a mismatch like this will often still generate useful data. Next, given that emotion detection has accuracy limitations [8] and is somewhat of a contentious topic [2, 19], researchers should be prepared for participants who may find their visualizations inaccurate and/or may be critical of sentiment analysis approaches in general. In these cases, it may be necessary to skip this part of the interview. Finally, researchers may have difficulty accessing longitudinal textual social media data, and should consider the feasibility of their data collection approaches.

7   CONCLUSION

We presented a novel approach for enabling interview participants to reflect on their personal histories and identity transitions: showing participants sentiment visualization graphs of their social media data over time. We detailed our methods of constructing sentiment visualization graphs and using them in interviews, so that HCI researchers can use these methods in future work to uncover participants' personal histories that may otherwise remain hidden. We describe how sentiment graphs helped participants reflect on their past experiences and the emotions surrounding them. Indeed, the graphs helped participants recall new data about intersecting life experiences that they had not previously mentioned on social media or brought up in our interviews, and that we would not have known to ask them about. Surprisingly, even participants who initially had difficulty relating their experiences to sentiment graphs used the graphs as tools for reflection. These findings show how sentiment graphs of social media data can be a powerful reflection tool to use with interview participants. Adding this research method to HCI researchers' toolboxes can make previously invisible aspects of participants' narratives visible and leading to the construction of what we call a *transition history from below*. In this way, uncovering personal histories can help us draw from the past to inform future technology design and research that supports people with marginalized and changing identities.

APPENDIX

A   SENTIMENT MEASURES DESCRIPTIVE STATISTICS

Table 1. Sentiment Measures Descriptive Statistics

Sentiment Score	Mean	Standard Deviation
Vader compound [47]	0.41	0.59
Vader positive sentiment [47]	0.16	0.12
Vader negative sentiment [47]	0.07	0.08
negative mental health measure from [71]	0.62	1.63
LIWC positive emotion [73]	4.19	3.99
LIWC negative emotion [73]	1.83	2.61

ACKNOWLEDGMENTS

REFERENCES

[1] Tawfiq Ammari, Sarita Schoenebeck, and Daniel M. Romero. 2018. Pseudonymous Parents: Comparing Parenting Roles and Identities on the Mommit and Daddit Subreddits. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems*.

[2] Nazanin Andalibi and Justin Buss. 2020. The Human in Emotion Recognition on Social Media: Attitudes, Outcomes, Risks. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, Honolulu, HI, 1–16. DOI:https://doi.org/10.1145/3313831.3376680

[3] Nazanin Andalibi and Andrea Forte. 2018. Announcing Pregnancy Loss on Facebook: A Decision-Making Framework for Stigmatized Disclosures on Identified Social Network Sites. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 158:1–158:14. DOI: https://doi.org/10.1145/3173574.3173732

[4] Paul André, M. C. Schraefel, Alan Dix, and Ryen W. White. 2011. Expressing Well-being Online: Towards Self-reflection and Social Awareness. In *Proceedings of the 2011 iConference (iConference'11)*. ACM, New York, NY, 114–121. DOI: https://doi.org/10.1145/1940761.1940777

[5] Jenna Baddeley and Jefferson A. Singer. 2007. Charting the life story's path: Narrative identity across the life span. In *Proceedings of the Handbook of Narrative Inquiry: Mapping a Methodology*, D Jean Clandinin (Ed.). SAGE Publications, Inc., Thousand Oaks, California, 177–202. DOI:https://doi.org/10.4135/9781452226552

[6] Anna Bagnoli. 2009. Beyond the standard interview: the use of graphic elicitation and arts-based methods. *Qualitative Research* 9, 5 (Nov. 2009), 547–570. DOI:https://doi.org/10.1177/1468794109343625

[7] Marcus Banks. 2018. *Using Visual Data in Qualitative Research*. SAGE Publications. Google-Books-ID: hJEStAEA-CAAJ.

[8] Erin O'Carroll Bantum and Jason E. Owen. 2009. Evaluating the validity of computerized content analysis programs for identification of emotional expression in cancer narratives. *Psychological Assessment* 21, 1 (March 2009), 79–88. DOI: https://doi.org/10.1037/a0014643

[9] Eric P. S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, New York, NY, 585–594. DOI:https://doi.org/10.1145/2702123.2702234

[10] Eric P. S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems*. ACM, New York, NY, 93–102. DOI: https://doi.org/10.1145/2598510.2598598

[11] Eric P. S. Baumer, Xiaotong Xu, Christine Chu, Shion Guha, and Geri K. Gay. 2017. When Subjects Interpret the Data: Social Media Non-use as a Case for Adapting the Delphi Method to CSCW. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*. ACM Press, Portland, Oregon, 1527–1543. DOI: https://doi.org/10.1145/2998181.2998182

[12] James "Bo" Begole, John C. Tang, and Rosco Hill. 2003. Rhythm Modeling, Visualizations and Applications. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 11–20. DOI: https://doi.org/10.1145/964696.964698

[13] Sabyasachi Bhattacharya. 1983. 'History from Below'. *Social Scientist* 11, 4 (1983), 3–20. DOI: https://doi.org/10.2307/3517020 Publisher: Social Scientist.

[14] Güler Boyraz, Sharon G. Horne, and Thomas V. Sayger†. 2010. Finding Positive Meaning After Loss: The Mediating Role of Reflection for Bereaved Individuals. *Journal of Loss and Trauma* 15, 3 (April 2010), 242–258. DOI: https://doi.org/10.1080/15325020903381683

[15] Peter J. Burke. 2006. Identity Change. *Social Psychology Quarterly* 69, 1 (2006), 81–96. Retrieved from http://www.jstor.org/stable/20141729.

[16] Ofelia Cervantes, Francisco Gutiérrez, Ernesto Gutiérrez, Esteban Castillo, J. Alfredo Sánchez, and Wanggen Wan. 2015. Expression: Visualizing Affective Content from Social Streams. In *Proceedings of the Latin American Conference on Human Computer Interaction.* ACM, New York, NY, 10:1–10:8. DOI: https://doi.org/10.1145/2824893.2824903

[17] Eun Kyoung Choe, Bongshin Lee, Haining Zhu, Nathalie Henry Riche, and Dominikus Baur. 2017. Understanding Self-reflection: How People Reflect on Personal Data Through Visual Data Exploration. In *Proceedings of the 11th EAI International Conference on Pervasive Computing Technologies for Healthcare.* ACM, New York, NY, 173–182. DOI: https://doi.org/10.1145/3154862.3154881 event-place: Barcelona, Spain.

[18] D. Jean Clandinin. 2007. *Handbook of Narrative Inquiry: Mapping a Methodology.* Sage Publications, Inc., Thousand Oaks, California. DOI: https://doi.org/10.4135/9781452226552

[19] Mike Conway and Daniel O'Connor. 2016. Social media, big data, and mental health: Current advances and ethical implications. *Current Opinion in Psychology* 2016, 9 (June 2016), 77–82. DOI: https://doi.org/10.1016/j.copsyc.2016.01.004

[20] Andrea J. Copeland and Denise E. Agosto. 2012. Diagrams and Relational Maps: The Use of Graphic Elicitation Techniques with Interviewing for Data Collection, Analysis, and Display. *International Journal of Qualitative Methods* 11, 5 (Dec. 2012), 513–533. DOI: https://doi.org/10.1177/160940691201100501

[21] Nathan Crilly, Alan F. Blackwell, and P. John Clarkson. 2006. Graphic elicitation: Using research diagrams as interview stimuli. *Qualitative Research* 6, 3 (Aug. 2006), 341–366. DOI: https://doi.org/10.1177/1468794106065007

[22] Munmun De Choudhury, M. Gamon, A. Hoff, and A. Roseway. 2013. "Moon Phrases": A social media faciliated tool for emotional reflection and wellness. In *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops.* 41–44.

[23] Munmun De Choudhury and Michael Massimi. 2015. "She said yes!" Liminality and Engagement Announcements on Twitter. In *Proceedings of iConference 2015.* Retrieved from http://www.news.gatech.edu/sites/default/files/GT-computing-research-twitter-engagement-2015.pdf.

[24] Dallas Denny. 2004. Changing Models of Transsexualism. *Journal of Gay & Lesbian Psychotherapy* 8, 1–2 (Aug. 2004), 25–40. DOI: https://doi.org/10.1300/J236v08n01_04

[25] R. Documentation. [n.d.]. scatter.smooth function. Retrieved December 15, 2016 from https://www.rdocumentation.org/packages/stats/versions/3.5.2/topics/scatter.smooth.

[26] Bryan Dosono, Yasmeen Rashidi, Taslima Akter, Bryan Semaan, and Apu Kapadia. 2017. Challenges in Transitioning from Civil to Military Culture: Hyper-Selective Disclosure through ICTs. *Proceedings of the ACM on Human-Computer Interaction* 1, 2 (Nov. 2017), 1–23. Retrieved from https://www.cs.indiana.edu/~kapadia/papers/rotc-cscw-2018.pdf.

[27] Stefanie Duguay. 2016. "He has a way gayer Facebook than I do": Investigating sexual identity disclosure and context collapse on a social networking site. *New Media & Society* 18, 6 (June 2016), 891–907. https://doi.org/10.1177/1461444814549930

[28] Abigail C. Durrant, David S. Kirk, Diego Trujillo-Pisanty, and Sarah Martindale. 2018. Admixed Portrait: Design to Understand Facebook Portrayals in New Parenthood. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 12:1–12:14. DOI: https://doi.org/10.1145/3173574.3173586

[29] Chris Elsden, David S. Kirk, and Abigail C. Durrant. 2016. A Quantified Past: Toward Design for Remembering With Personal Informatics. *Human-Computer Interaction* 31, 6 (Nov. 2016), 518–557. DOI: https://doi.org/10.1080/07370024.2015.1093422

[30] Daniel Epstein, Felicia Cordeiro, Elizabeth Bales, James Fogarty, and Sean Munson. 2014. Taming data complexity in lifelogs: Exploring visual cuts of personal informatics data. In *Proceedings of the 2014 conference on Designing interactive systems.* ACM, 667–676. DOI: https://doi.org/10.1145/2598510.2598558

[31] Daniel A. Epstein, An Ping, James Fogarty, and Sean A. Munson. 2015. A Lived Informatics Model of Personal Informatics. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing.* ACM, New York, NY, 731–742. DOI: https://doi.org/10.1145/2750858.2804250

[32] Stephanie D. H. Evergreen. 2016. *Effective Data Visualization: The Right Chart for the Right Data.* SAGE Publications. Google-Books-ID: yJgECwAAQBAJ.

[33] Joel E. Fischer, Andy Crabtree, Tom Rodden, James A. Colley, Enrico Costanza, Michael O. Jewell, and Sarvapali D. Ramchurn. 2016. "Just whack it on until it gets hot": Working with IoT Data in the Home. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM Press, Santa Clara, California, 5933–5944. DOI: https://doi.org/10.1145/2858036.2858518

[34] Justine Gangneux. 2019. Rethinking social media for qualitative research: The use of Facebook Activity Logs and Search History in interview settings. *The Sociological Review* 67, 6 (Nov. 2019), 1249–1264. https://doi.org/10.1177/0038026119859742

[35] Nanna Gorm and Irina Shklovski. 2017. Participant Driven Photo Elicitation for Understanding Activity Tracking: Benefits and Limitations. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing.* ACM Press, Portland, Oregon, 1350–1361. DOI:https://doi.org/10.1145/2998181.2998214

[36] Xinning Gui, Yu Chen, Yubo Kou, Katie Pine, and Yunan Chen. 2017. Investigating Support Seeking from Peers for Pregnancy in Online Health Communities. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 50:1–50:19. DOI:https://doi.org/10.1145/3134685

[37] Oliver L. Haimson. 2018. Social Media as Social Transition Machinery. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (Nov. 2018), 63:1–63:21. DOI: https://doi.org/10.1145/3274332

[38] Oliver L. Haimson. 2019. Mapping gender transition sentiment patterns via social media data: Toward decreasing transgender mental health disparities. *Journal of the American Medical Informatics Association* 26, 8–9 (2019), 749–758. DOI:https://doi.org/10.1093/jamia/ocz056

[39] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2015. Disclosure, Stress, and Support During Gender Transition on Facebook. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing.* ACM, New York, NY, 1176–1190. DOI: https://doi.org/10.1145/2675133.2675152

[40] Oliver L. Haimson, Jed R. Brubaker, Lynn Dombrowski, and Gillian R. Hayes. 2016. Digital Footprints and Changing Networks During Online Identity Transitions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 2895–2907. DOI:https://doi.org/10.1145/2858036.2858136

[41] Oliver L. Haimson, Albert J. Carter, Shanley Corvite, Brooke lyn Wheeler, Lingbo Wang, Tianxiao Liu, and Alexxus Lige. 2021. The Major Life Events Taxonomy: Social Readjustment, Social Media Information Sharing, and Online Network Separation During Times of Life Transition. *Journal of the Association for Information Science and Technology* 72, 7(2021), 933–947.

[42] Oliver L. Haimson, Bryan Semaan, Brianna Dym, Joey Chiao-Yin Hsiao, Daniel Herron, and Wendy Moncur. 2019. Life Transitions and Social Technologies: Research and Design for Times of Life Change. In *Proceedings of the Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing.* Association for Computing Machinery, Austin, TX, 480–486. DOI:https://doi.org/10.1145/3311957.3359431

[43] T. Hogan, U. Hinrichs, and E. Hornecker. 2016. The Elicitation Interview Technique: Capturing People's Experiences of Data Representations. *IEEE Transactions on Visualization and Computer Graphics* 22, 12 (Dec. 2016), 2579–2593. DOI: https://doi.org/10.1109/TVCG.2015.2511718

[44] Victoria Hollis, Artie Konrad, Aaron Springer, Matthew Antoun, Christopher Antoun, Rob Martin, and Steve Whittaker. 2017. What Does All This Data Mean for My Future Mood? Actionable Analytics and Targeted Reflection for Emotional Well-Being. *Human-Computer Interaction* 32, 5–6 (Nov. 2017), 208–267. DOI: https://doi.org/10.1080/07370024.2016.1277724

[45] Thomas H. Holmes and Richard H. Rahe. 1967. The social readjustment rating scale. *Journal of Psychosomatic Research* 11, 2 (Aug. 1967), 213–218. DOI:https://doi.org/10.1016/0022-3999(67)90010-4

[46] D. Huang, M. Tory, B. Adriel Aseniero, L. Bartram, S. Bateman, S. Carpendale, A. Tang, and R. Woodbury. 2015. Personal Visualization and Personal Visual Analytics. *IEEE Transactions on Visualization and Computer Graphics* 21, 3 (March 2015), 420–433. DOI: https://doi.org/10.1109/TVCG.2014.2359887

[47] Clayton J. Hutto and Eric Gilbert. 2014. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In *Proceedings of the 8th International AAAI Conference on Weblogs and Social Media.* Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM14/paper/view/8109.

[48] Ellen Isaacs, Artie Konrad, Alan Walendowski, Thomas Lennig, Victoria Hollis, and Steve Whittaker. 2013. Echoes from the Past: How Technology Mediated Reflection Improves Well-being. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 1071–1080. DOI: https://doi.org/10.1145/2470654.2466137

[49] Sandy E. James, Jody L. Herman, Susan Rankin, Mara Keisling, Lisa Mottet, and Ma'ayan Anafi. 2016. *The Report of the 2015 U.S. Transgender Survey.* Technical Report. National Center for Transgender Equality, Washington, DC. Retrieved from http://www.transequality.org/sites/default/files/docs/USTS-Full-Report-FINAL.PDF.

[50] C. David Jenkins, Michael W. Hurst, and Robert M. Rose. 1979. Life Changes: Do People Really Remember? *Archives of General Psychiatry* 36, 4 (April 1979), 379–384. DOI:https://doi.org/10.1001/archpsyc.1979.01780040021001

[51] Rohit Ashok Khot, Larissa Hjorth, and Florian 'Floyd' Mueller. 2014. Understanding physical activity through 3D printed material artifacts. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems.* ACM Press, Toronto, Ontario, Canada, 3835–3844. DOI:https://doi.org/10.1145/2556288.2557144

[52] Artie Konrad, Simon Tucker, John Crane, and Steve Whittaker. 2016. Technology and Reflection: Mood and Memory Mechanisms for Well-Being. *Psychology of Well-Being* 6, 1 (June 2016), 5. DOI: https://doi.org/10.1186/s13612-016-0045-3

[53] Patricia G. Lange. 2007. Publicly Private and Privately Public: Social Networking on YouTube. *Journal of Computer-Mediated Communication* 13, 1 (Oct. 2007), 361–380. DOI:https://doi.org/10.1111/j.1083-6101.2007.00400.x

[54] Guillame Latzko-Toth, Claudine Bonneau, and Mélanie Millette. 2017. Small Data, Thick Data: Thickening Strategies for Trace-Based Social Media Research. In *Proceedings of the SAGE Handbook of Social Media Research Methods*, Luke Sloan and Anabel Quan-Haase (Eds.). SAGE. Google-Books-ID: 9oewDQAAQBAJ.

[55] Clayton Lewis, Peter G. Polson, Cathleen Wharton, and John Rieman. 1990. Testing a walkthrough methodology for theory-based design of walk-up-and-use interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems Empowering people*. ACM Press, Seattle, Washington, 235–242. DOI: https://doi.org/10.1145/97243.97279

[56] Ian Li, Anind Dey, and Jodi Forlizzi. 2009. Grafitter: Leveraging social media for self reflection. *XRDS: Crossroads, The ACM Magazine for Students* 16, 2 (2009), 2.

[57] Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A Stage-based Model of Personal Informatics Systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 557–566. DOI: https://doi.org/10.1145/1753326.1753409

[58] Ian Li, Anind K. Dey, and Jodi Forlizzi. 2011. Understanding My Data, Myself: Supporting Self-reflection with Ubicomp Technologies. In *Proceedings of the 13th International Conference on Ubiquitous Computing*. ACM, New York, NY, 405–414. DOI: https://doi.org/10.1145/2030112.2030166

[59] Haiwei Ma, C. Estelle Smith, Lu He, Saumik Narayanan, Robert A. Giaquinto, Roni Evans, Linda Hanson, and Svetlana Yarosh. 2017. Write for Life: Persisting in Online Health Communities Through Expressive Writing and Social Support. *Proceedings of the ACM on Human Computer Interaction* 1, CSCW (Dec. 2017), 73:1–73:24. DOI:https://doi.org/10.1145/3134708

[60] Haley MacLeod, Kim Oakes, Danika Geisler, Kay Connelly, and Katie Siek. 2015. Rare World: Towards Technology for Rare Diseases. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM Press, Seoul, Republic of Korea, 1145–1154. DOI: https://doi.org/10.1145/2702123.2702494

[61] Alessandro Marcengo and Amon Rapp. 2014. Visualization of Human Behavior Data: The Quantified Self. In *Innovative Approaches of Data Visualization and Visual Analytics*. 236–265. Retrieved June 22, 2019 from https://www.igi-global.com/chapter/visualization-of-human-behavior-data/78722

[62] Michael Massimi, Jackie L. Bender, Holly O. Witteman, and Osman H. Ahmed. 2014. Life transitions and online health communities: Reflecting on adoption, use, and disengagement. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM Press, 1491–1501. DOI: https://doi.org/10.1145/2531602.2531622

[63] Sarah McRoberts, Haiwei Ma, Andrew Hall, and Svetlana Yarosh. 2017. Share first, save later: Performance of self through snapchat stories. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado). Association for Computing Machinery, New York, NY, 6902–6911. DOI: https://doi.org/10.1145/3025453.3025771

[64] Sandra Metts, Susan Sprecher, and William R. Cupach. 1993. Retrospective Self-Reports. In *Proceedings of the Studying Interpersonal Interaction*. Guilford Press. Google-Books-ID: RxdkR7fC5gEC.

[65] Terence R. Mitchell, Leigh Thompson, Erika Peterson, and Randy Cronk. 1997. Temporal Adjustments in the Evaluation of Events: The "Rosy View". *Journal of Experimental Social Psychology* 33, 4 (July 1997), 421–448. DOI: https://doi.org/10.1006/jesp.1997.1333

[66] Wendy Mitchell and Annie Irvine. 2008. I'm okay, you're okay?: Reflections on the well-being and ethical requirements of researchers and research participants in conducting qualitative fieldwork interviews. *International Journal of Qualitative Methods* 7, 4 (2008), 31–44. DOI:https://doi.org/10.1177/160940690800700403

[67] Wendy Moncur, Lorna Gibson, and Daniel Herron. 2016. The role of digital technologies during relationship breakdowns. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*.

[68] Scott M. Monroe. 1982. Life events assessment: Current practices, emerging trends. *Clinical Psychology Review* 2, 4 (Jan. 1982), 435–453. DOI:https://doi.org/10.1016/0272-7358(82)90023-X

[69] Tsubasa Morioka, Nicole B. Ellison, and Michael Brown. 2016. Identity Work on Social Media Sites: Disadvantaged College Students' First Year College Transition. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM Press, 846–857. DOI: https://doi.org/10.1145/2818048.2819959

[70] Sean A. Munson and Sunny Consolvo. 2012. Exploring goal-setting, rewards, selfmonitoring, and sharing to motivate physical activity. In *Proceedings of the 2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth)*. IEEE, 25–32.

[71] Arindam Paul, Ankit Agrawal, Wei-keng Liao, and Alok Choudhary. 2016. AnonyMine: Mining anonymous social media posts using psycho-lingual and crowd-sourced dictionaries. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.

[72] S. Tejaswi Peesapati, Victoria Schwanda, Johnathon Schultz, Matt Lepage, So-yae Jeong, and Dan Cosley. 2010. Pensieve: Supporting Everyday Reminiscence. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 2027–2036. DOI: https://doi.org/10.1145/1753326.1753635

[73] James W. Pennebaker, Ryan L. Boyd, Kayla Jordan, and Kate Blackburn. 2015. *The development and psychometric properties of LIWC2015.* Technical Report. University of Texas at Austin, Austin, TX. Retrieved from https://utexas-ir.tdl.org/handle/2152/31333.

[74] Laura R. Pina, Sang-Wha Sien, Teresa Ward, Jason C. Yip, Sean A. Munson, James Fogarty, and Julie A. Kientz. 2017. From Personal Informatics to Family Informatics: Understanding Family Practices around Health Monitoring. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing.* ACM Press, Portland, Oregon, 2300–2315. DOI: https://doi.org/10.1145/2998181.2998362

[75] Stefinee Pinnegar and J. Gary Daynes. 2007. Locating narrative inquiry historically: Thematics in the turn to narrative. In *Proceedings of the Handbook of Narrative Inquiry: Mapping a Methodology*, D. Jean Clandinin (Ed.). SAGE Publications, Inc., Thousand Oaks, California. DOI:https://doi.org/10.4135/9781452226552

[76] PyTumblr. n.d.. PyTumblr. Retrieved December 15, 2016 from https://github.com/tumblr/pytumblr.

[77] Tauhidur Rahman, Mi Zhang, Stephen Voida, and Tanzeem Choudhury. 2014. Towards Accurate Non-Intrusive Recollection of Stress Levels Using Mobile Sensing and Contextual Recall. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare.* ICST, Oldenburg, Germany. DOI: https://doi.org/10.4108/icst.pervasivehealth.2014.254957

[78] Amon Rapp and Federica Cena. 2016. Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human-Computer Studies* 2016, 94 (Oct. 2016), 1–17. DOI : https://doi.org/10.1016/j.ijhcs.2016.05.006

[79] Brady Robards and Siân Lincoln. 2017. Uncovering longitudinal life narratives: Scrolling back on Facebook. *Qualitative Research* 17, 6 (Dec. 2017), 715–730. DOI: https://doi.org/10.1177/1468794117700707

[80] Sarita Schoenebeck, Nicole B. Ellison, Lindsay Blackwell, Joseph B. Bayer, and Emily B. Falk. 2016. Playful Backstalking and Serious Impression Management: How Young Adults Reflect on their Past Identities on Facebook. In *Proceedings of the ACM Conference on Computer Supported Cooperative Work & Social Computing.*

[81] Victoria Schwanda Sosik, Xuan Zhao, and Dan Cosley. 2012. See Friendship, Sort of: How Conversation and Digital Traces Might Support Reflection on Friendships. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work.* ACM, New York, NY, 1145–1154. DOI:https://doi.org/10.1145/2145204.2145374

[82] Bryan Semaan, Lauren M. Britton, and Bryan Dosono. 2017. Military Masculinity and the Travails of Transitioning: Disclosure in Social Media. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing.* ACM, New York, NY, 387–403. DOI:https://doi.org/10.1145/2998181.2998221

[83] Bryan C. Semaan, Lauren M. Britton, and Bryan Dosono. 2016. Transition Resilience with ICTs: 'Identity Awareness' in Veteran Re-Integration. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems.* ACM, New York, NY, 2882–2894. DOI: https://doi.org/10.1145/2858036.2858109

[84] Phoebe Sengers, Kirsten Boehner, Shay David, and Joseph 'Jofish' Kaye. 2005. Reflective Design. In *Proceedings of the 4th Decennial Conference on Critical Computing: Between Sense and Sensibility.* ACM, New York, NY, 49–58. DOI: https://doi.org/10.1145/1094562.1094569

[85] Madeline E. Smith, Duyen T. Nguyen, Charles Lai, Gilly Leshed, and Eric P. S. Baumer. 2012. Going to College and Staying Connected: Communication Between College Freshmen and Their Parents. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work.* ACM, New York, NY, 789–798. DOI: https://doi.org/10.1145/2145204.2145322

[86] Aaron Springer, Victoria Hollis, and Steve Whittaker. 2018. Mood Modeling: Accuracy Depends on Active Logging and Reflection. *Personal and Ubiquitous Computing* 22, 4 (Aug. 2018), 723–737. DOI: https://doi.org/10.1007/s00779-018-1123-8

[87] Allucquère Rosanne (Sandy) Stone. 1987. The Empire Strikes Back: A Posttranssexual Manifesto. Retrieved from http://sandystone.com/empire-strikes-back.pdf.

[88] Anselm Strauss and Juliet M. Corbin. 1998. *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory.* SAGE Publications. Google-Books-ID: tBcEjwEACAAJ.

[89] Susan Stryker. 2017. *Transgender History.* Seal Press.

[90] Mark Tausig. 1982. Measuring Life Events. *Journal of Health and Social Behavior* 23, 1 (1982), 52–64. DOI: https://doi.org/10.2307/2136389

[91] Edward Palmer Thompson. 2001. *The Essential E.P. Thompson.* New Press. Google-Books-ID: 6aiw2SDj2Q8C.

[92] Rachel Thomson, Robert Bell, Janet Holland, Sheila Henderson, Sheena McGrellis, and Sue Sharpe. 2002. Critical Moments: Choice, Chance and Opportunity in Young People's Narratives of Transition. *Sociology* 36, 2 (May 2002), 335–354. DOI: https://doi.org/10.1177/0038038502036002006

[93] Avril Thorne and Kate C McLean. 2003. Telling traumatic events in adolescence: A study of master narrative posi-
      tioning. In *Proceedings of the Autobiographical Memory and the Construction of a Narrative Self: Developmental and
      Cultural Perspectives*. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, 169–185.
[94] A. Thudt, D. Baur, S. Huron, and S. Carpendale. 2016. Visual Mementos: Reflecting Memories with Personal Data.
      IEEE Transactions on Visualization and Computer Graphics 22, 1 (Jan. 2016), 369–378. DOI: https://doi.org/10.1109/
      TVCG.2015.2467831
[95] Alice Thudt, Uta Hinrichs, Samuel Huron, and Sheelagh Carpendale. 2018. Self-Reflection and Personal Physicaliza-
      tion Construction. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. ACM, New
      York, NY, 154:1–154:13. DOI:https://doi.org/10.1145/3173574.3173728
[96] Alice Thudt, Charles Perin, Wesley Willett, and Sheelagh Carpendale. 2017. Subjectivity in personal storytelling with
      visualization. *Information Design Journal* 23, 1 (Jan. 2017), 48–64. DOI:https://doi.org/10.1075/idj.23.1.07thu
[97] Peter Tolmie, Andy Crabtree, Tom Rodden, James Colley, and Ewa Luger. 2016. "This Has to Be the Cats": Personal
      Data Legibility in Networked Sensing Systems. In *Proceedings of the 19th ACM Conference on Computer-Supported
      Cooperative Work & Social Computing*. ACM, New York, NY, 491–502. DOI: https://doi.org/10.1145/2818048.2819992
[98] Tumblr. 2014. Application Developer and API License Agreement | Tumblr. Retrieved July 30, 2016 from https://www.
      tumblr.com/docs/en/api_agreement.
[99] Tumblr. n.d.. Tumblr API. Retrieved December 15, 2016 from https://www.tumblr.com/docs/en/api/v2.
[100] K. Vrotsou, K. Ellegard, and M. Cooper. 2007. Everyday Life Discoveries: Mining and Visualizing Activity Patterns in
      Social Science Diary Data. In *Proceedings of the 2007 11th International Conference Information Visualization*. 130–138.
      DOI:https://doi.org/10.1109/IV.2007.48
[101] W. Richard Walker, John J. Skowronski, and Charles P. Thompson. 2003. Life is Pleasant – and Memory Helps to Keep
      it that Way! *Review of General Psychology* 7, 2 (June 2003), 203–210. DOI:https://doi.org/10.1037/1089-2680.7.2.203
[102] Xuan Zhao and Siân E. Lindley. 2014. Curation through use: understanding the personal value of social media. In
      Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM Press, 2431–2440. DOI: https:
      //doi.org/10.1145/2556288.2557291