

Narrative Paths and Negotiation of Power in Birth Stories

MARIA ANTONIAK, DAVID MIMNO, and KAREN LEVY, Cornell University, USA

Birth stories have become increasingly common on the internet, but they have received little attention as a computational dataset. These unsolicited, publicly posted stories provide rich descriptions of decisions, emotions, and relationships during a common but sometimes traumatic medical experience. These personal details can be illuminating for medical practitioners, and due to their shared structures, birth stories are also an ideal testing ground for narrative analysis techniques. We present an analysis of 2,847 birth stories from an online forum and demonstrate the utility of these stories for computational work. We discover clear sentiment, topic and persona-based patterns that both model the expected narrative event sequences of birth stories and highlight diverging pathways and exceptions to narrative norms. The authors' motivation to publicly post these personal stories can be a way to regain power after a surveilled and disempowering experience, and we explore power relationships between the personas in the stories, showing that these dynamics can vary with the type of birth (e.g., medicated vs unmedicated). Finally, birth stories exist in a space that is both public and deeply personal. This liminality poses a challenge for analysis and presentation, and we discuss tradeoffs and ethical practices for this collection. **WARNING:** This paper includes detailed narratives of pregnancy and birth.

CCS Concepts: • **Computing methodologies** → **Natural language processing**; *Discourse, dialogue and pragmatics*.

Additional Key Words and Phrases: Narrative; Birth Stories; Power; Natural Language Processing

ACM Reference Format:

Maria Antoniak, David Mimno, and Karen Levy. 2019. Narrative Paths and Negotiation of Power in Birth Stories. *Proc. ACM Hum.-Comput. Interact.* 3, CSCW, Article 88 (November 2019), 27 pages. <https://doi.org/10.1145/3359190>

1 INTRODUCTION

Birth stories are detailed personal narratives of real experiences giving birth. These stories, which have existed for millennia as written and oral histories passed from person to person, have become popular in online social communities, where they are shared via videos, blog posts, and forum posts. The authors' motivations in posting these personal stories on a public forum are complex: possible motivators include re-working of trauma [4, 93], sharing of information [16], resistance against surveillance [18, 43], and re-negotiating power after a disempowering medical experience [16, 18]. Despite their online popularity, birth stories have not received attention as a computational dataset. We conduct a computer-assisted reading of an online birth stories community and examine how this particular community frames the experience of giving birth and defines norms and exceptions.

Computational modeling of the narrative structure and persona hierarchies expressed across a collection of birth stories can 1) provide a lens into the shared sensemaking through stories in an online health community, 2) suggest a unique dataset and testing ground for researchers studying computational approaches to narrative, and 3) aid medical professionals in understanding the

Authors' address: Maria Antoniak, maa343@cornell.edu; David Mimno, mimno@cornell.edu; Karen Levy, karen.levy@cornell.edu, Cornell University, USA.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Copyright held by the owner/author(s). Publication rights licensed to ACM.

2573-0142/2019/11-ART88 \$15.00

<https://doi.org/10.1145/3359190>

expectations and experiences of pregnant people¹. In this work, we use natural language processing (NLP) techniques to translate between two complementary perspectives of an online community of storytellers: the quantifiable rigor of aggregated statistical patterns, and the richness and specificity of narrative.

Prior work has examined the effects on formerly pregnant individuals of sharing their birth stories [29, 74] and of the sensemaking that occurs in offline communities that share birth stories [16, 58, 88]. We examine the shared norms and expectations that arise through stories in a specific online health community, *r/BabyBumps*, a forum on Reddit. Articulating a story gives it structure, and this structure can help the storyteller assimilate the experience and find meaning [58]. Online communities allow pregnant and postpartum people to not only make sense of their own stories but to make sense of their connected stories as a particular community. Statistical analysis of these stories can reveal the community-held beliefs, expectations, and narrative norms that arise from this online sharing.

As an example of complex discourse, narrative and story have become an area of considerable interest in NLP. For example, there have been increasingly sophisticated attempts to generate coherent stories using neural language models [24, 69]. Previous work focuses on identifying events and relationships between characters in news articles [20] and movie plot summaries [90]. However, it has been difficult to create narrative datasets that have both clearly identifiable structural characteristics as well as realistic levels of creative expression. Constructed datasets often do not reflect the complexities of real narratives [73, 86, 97], while gathered datasets (e.g., from literature, blogs, or news) can include unconstrained numbers of topics, personas, and events.

Medical narratives, and specifically birth stories, fulfill many of the desiderata for computational narrative analyses. Birth stories are highly individual, but all share fundamental similarities due to the fact that they describe the same biological process. From the perspective of computational narratology, the collection provides an opportunity to measure the balance between formalist similarity between stories and creative individuality within stories.

In the internet-using, English-language contexts that make up the collection, we expect childbirth to typically occur in a medical settings, in which birthing is a formalized process. We expect that in birth stories similar events will happen in similar sequences, and indeed, we find strong evidence to support this hypothesis. Unlike other work that identifies character archetypes [5, 7], we expect birth stories to contain a small number of easily identified “personas” such as the first-person narrator, partners, doctors, nurses, and midwives. We find that relatively simple approaches are able to find patterns of personas and events in birth stories, but also to find branching paths that reflect key decisions and complications; we expect this result to extend to other medical narratives which are similarly structured. Establishing what is ordinary can also then help us detect what is unusual and unique about a specific story.

Better analysis of birth stories could also lead to important social and medical benefits. In the U.S., rates of pregnancy-related deaths and complications have steadily risen over the last thirty years [9, 26, 27, 67]. These rates are higher for Hispanic women and non-U.S. citizens and are especially high for black women, whose maternal mortality rates are three to four times higher than those of white women [27, 71]. Postpartum depression affects 6.5-12.9% of people after childbirth, and of these people, about 20% will still suffer from depression a year after delivery [91]. Yet 46% of maternal deaths of black women and 33% of maternal deaths of white women are estimated as preventable [10]. Likewise, the symptoms of postpartum depression can be prevented or mitigated by emotional support during and after pregnancy [16, 91]. While maternal mortality and postpartum depression statistics are alarming, stories can be a more powerful way to learn about the lived

¹We use the terms *pregnant people* and *people who have given birth* throughout the paper.

experiences of pregnant people and improve their treatment, as shown by recent media attention to the stories of births gone wrong [70].

Despite the high potential value of personal medical stories, naturally occurring and publicly available medical datasets are rare. Due to privacy and intellectual property restrictions, medical data involving interactions between patients and clinicians is difficult to obtain for research. Medical language models usually rely on PubMed, the Unified Medical Language System (UMLS) [13], and similar resources composed of biomedical publications [22, 83]. Models trained on these resources better represent academic language than language used by patients. Birth stories are therefore a valuable resource of unsolicited medical narratives told from the patient's point of view. Birth stories could inform medical workers about how to better prepare people for pregnancy, birth, and recovery and help us better understand debilitating conditions like postpartum depression. The medical decisions and outcomes described could inform healthcare choices while the emotions and power dynamics could inform more empathetic nursing care [8, 16].

Our contributions are as follows.

- We gather and label a novel collection of 2,847 naturally occurring birth stories from an online forum and make our labeling pipelines available to the research community.
- We examine the narrative, sentiment, and persona-based patterns contained in the collection of stories and show that while each story is unique, consistent events and personas can be extracted that occur in predictable sequences, proving the utility of the dataset for both computational narrative analysis and the discovering narrative norms and expectations of this online health community.
- We identify outlier stories within the community by measuring unusual topic transitions and correlate these outliers with the authors' labeling and framing of the stories.
- We use a lexicon to automatically measure the power hierarchies of the personas presented in the birth stories; we find evidence of this community's framing of the author as disempowered and of the doula as powerful and boundary-crossing.

2 RELATED WORK

2.1 Pregnancy and the Internet

In addition to traditional medical sources, pregnant people often turn to friends, family, and other new parents for healthcare information [8, 60, 74]. Pregnant people increasingly seek this information on the internet via searches, forums, and social media, and information found online can influence decisions at all stages of pregnancy [61]. This turn to the internet for healthcare information is motivated at least in part by a) dissatisfaction and/or limited access to clinicians and b) physical and social isolation of new parents [41, 47]. The specificity of the internet searches of pregnant people can be clustered into specific gestational periods [40], and postpartum depression can be predicted from patterns of social media use during and after pregnancy [30, 31].

Online health communities (e.g., forums focused on supporting people with a particular health condition) allow patients to give and receive emotional and informational support [98, 99] and to share expressive writing about sometimes traumatic experiences [65]. These communities are an important new source of both information and social support for pregnancy, birth, and pregnancy loss [3, 42, 75]. Compared to social networks or real-life support groups and conversations, disclosure on a public online forum can feel more private due to the anonymity of the users, physical privacy of the users (emotional reactions are hidden), and a sympathetic, knowledgeable audience [42].

Personal tracking and a desire to check one's experiences against those of others could be additional motivators for people to seek medical information via social media [35]. Women often use forms of personal tracking (e.g., menstrual tracking) as part of their healthcare routines. The

amount of detail contained in birth stories (e.g., times between contractions, dosages of medications) suggests that the stories could serve as a diary of information both for the author and the community audience. Among the birth stories we collected, many received comments either reassuring the author or chiming in, with surprise, that they had had similar experiences but had not realized that these were normal. In one notable example, a single birth story in which the author described violently shaking after a c-section inspired a cascade of responses, all saying that they had also experienced uncontrollable shaking but had not realized until this moment that it was normal or happened to anyone else.

2.2 Stories and Datasets

Storytelling and expressive writing about traumatic experiences can have physical, emotional, and social health benefits [72, 76, 78, 79] and have been studied for their role in therapy and the establishment of social norms. In one example, Arigo and Smyth [4] find that expressive writing about eating patterns and personal appearance by college-age women either improves a range of outcomes or reduces the risk of further decline. de Moor et al. [32], on the other hand, find no benefits in breast cancer survivors. Tangherlini [93] argues that the storytelling of paramedics forms the culture and organizational structure of their jobs. Through these stories, the paramedics both work through trauma of difficult experiences and negotiate their place in a hierarchy of workers (doctors, nurses, police officers, other paramedics). Likewise, patients tell stories about their experiences that cast medical professionals into certain roles, which they need to understand if they wish to interact with the patient effectively [92].

The birth stories dataset is one in a growing collection of story and narrative datasets. Narrative analysis research has relied on datasets such as fictional works (novels) [54], news stories [19], fairy tales [55], Wikipedia biographies [6], artificial story datasets [73, 86, 97], and personal stories shared orally or on blogs [44, 77, 84]. The birth stories collection combines the advantages of several of these existing datasets. Unlike artificial story datasets, birth stories include naturally occurring, complex stories, and unlike existing gathered datasets, birth stories are constrained by topic and share common structures and personas. Our dataset of birth stories maintains a balance between data targeted to a particular computational problem and data with real-world origins and impact.

2.3 Birth Stories

Birth stories have existed for millennia, as written and oral histories passed from mother to daughter, friend to friend, and midwife to midwife. The sharing of birth stories can significantly affect the attitudes and beliefs of pregnant people [29, 58, 74]. Birth stories have received much attention from the medical and sociological communities [16, 18, 88], but these studies mostly rely on qualitative methods and solicited stories. For example, Carson et al. [18] use birth stories to explore how young parents frame their experiences giving birth, comparing these personal narratives to societal narratives of stigma and considering how stories can provide opportunities for agency after a surveilled experience (observed by the community, family, partner, and medical professionals). Birth stories have been also explored as a method for knowledge transfer [88]. Callister [16] identifies five cross-cultural motivations for sharing birth stories: integrating the event into the narrator's life; sharing a significant life experience; discussing fears, disappointments, and "missing pieces"; understanding the narrator's own strengths; and connecting to a community of women.

It is only recently that large numbers of birth stories have become available online. We are aware of one other study that has performed a quantitative analysis of online birth stories, using traditional social science content analysis procedures that did not take advantage of computation. Bylund [15] annotate a set of 551 stories with fine-grained decision and sentiment labels as well as noting whether it was the author's first or second birth. These labels are used to determine that

pregnant people are rarely involved in decision-making and this lack of decision-making correlates with negative emotions. Pregnant people were no more likely to be involved in decision-making in their second compared to their first births, but pregnant people with midwives as their clinicians had higher involvement in decision-making.

3 DATA

3.1 Overview

Birth stories are narratives of individual experiences giving birth, often in great medical and emotional detail, publicly posted on blogs, video-sharing websites, and forums such as *Mothering*² or *BellyBelly*.³ Birth stories are very popular: *Mothering* alone contains more than 2,000 stories, and each story can receive thousands of views. The popularity of these stories speaks to the need for a narrative outlet for people who have undergone a surveilled and sometimes traumatic experience and who are now reclaiming their stories by publishing them publicly on the internet and sharing information with other pregnant people.

Below is a paraphrased and shortened example of a birth story posted to *r/BabyBumps*.

*I finally had my gorgeous baby girl at 41 weeks and 3 days on 3/3/2017!
So I heard from 37 weeks that because of the size of the baby, I probably wouldn't be able to get all the way to 40 weeks and induction might be necessary. Well 39 weeks came and my doctor said that I shouldn't go past 41 weeks.*

...

8:30 AM: AT LAST at 41+1 I went to my appointment and was sent to hospital. I go as fast as I can to labor and delivery knowing that they'll have to schedule an induction. Baby was fine in his current spot but the on call OB decided to start an induction.

...

My partner was really relieved that I decided to get an epidural...I think he was getting nervous! Since he was witnessing me in pain. Honestly the epidural wasn't bad, especially comparing the short term pain to the endless contractions.

...

I kept telling the nurse that I felt some pressure, and she was so surprised when she checked me. The nurse said that I was 10cm and ready to push! She went to get the midwife and I began pushing.

...

He scored 9/10 and immediately latched. Breastfeeding wasn't as strange as I expected and it actually came really naturally. All my fears were unnecessary after all.

The typical range of the story is from the first contraction to the birth of the child. Most of the details concern the stages of birth, the author's feelings and expectations and their contrast with the reality of birthing, and the medical decisions made by or for the author. Some stories begin earlier, describing the pregnancy, symptoms, ultrasounds, birthing classes, and expectations of the pregnant person, and others extend beyond the birth itself and describe time spent in the hospital or the first few weeks adjusting to life with a baby.

²<http://www.mothering.com/forum/166-birth-stories/>

³<http://www.bellybelly.com.au/forums/birth-stories-63/>

Number of stories with more than 500 words	2,847
Average number of words per story	1,311
Number of words in longest story	6,057
Number of unique words	29,952

Table 1. Corpus statistics for the birth stories collected from *r/BabyBumps*.

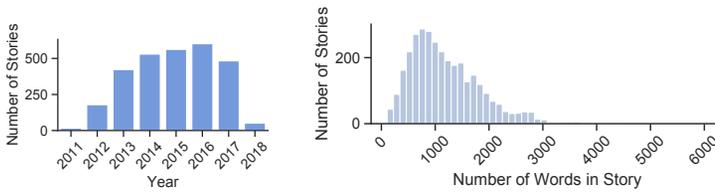


Fig. 1. Number of birth stories per year in the collected dataset and distribution of story lengths (by number of words). Collection began in February, 2011 and ended in March, 2018. The mean story length is 1,311 words.

3.2 Data Collection and Cleaning

We collect 2,847 birth stories from the social website Reddit. While birth stories exist in many venues and forms, we choose to focus on Reddit for its accessibility and well-studied communities. These stories were posted publicly from February 16th, 2011 to February 28th, 2018 on the subreddit (forum) *r/BabyBumps* (all data available up to the date of collection). *r/BabyBumps* is a forum intended to be a “place for pregnant redditors, those who have been pregnant, those who wish to be in the future, and anyone who supports them.” This community includes a wide range of posts related to pregnancy and birth, including humor, requests for advice, rants and venting, recommendations, and journaling posts (e.g., bump and ultrasound photos, summaries of doctor appointments, birth announcements and stories). The community rules explicitly instruct members to post *detailed* birth stories rather than only photos and one line descriptions.

We perform two rounds of filtering: first, we select the posts that contain the n-gram “birth story” in the title, and second, we remove 348 posts that contain fewer than 500 words. We remove these short posts as a second step of data cleaning as many of these shorter posts are either not birth stories or are only parts of birth stories published in installments. We do not include comments, upvotes, or other interactions in our analysis; only the parent posts, containing the stories, are included. Table 1 and Figure 1 show summary statistics for the collected corpus.

This set of stories constitutes a small sample of birth stories posted online, and an even smaller sample of all birth stories, told or untold. Because the stories were posted anonymously to a forum, we do not have demographic data for the authors of the stories. All stories are written in English, the majority appear to take place in western and developed countries, and the authors have access to the internet and to Reddit. As a first research foray into the computational study of birth stories, we expect not that the patterns observed in *r/BabyBumps* will generalize to all other birth stories, but that we will provide evidence of the value and research interest of all birth stories.

3.3 Labeling Process

The stories come with very little structured metadata. We therefore add a small number of codes automatically based on the occurrence of frequent words and phrases the author uses in the title of the post. This approach is inspired by our observation that titles of posts frequently contain terms that are intended to signal information about key aspects of a story to the community. These

descriptions illuminate the community’s understanding and prioritization of events that take place in the birth stories. Paraphrased examples of the post titles can be seen in Table 2. By using the labels assigned by authors, we do not impose our own structure on the community.

We select eight labels (shown in Figure 3) that are both distinct and also frequent enough that we can gather reliable statistics about them. We rank all the unigrams and bigrams used in the titles by frequency, and we map the most frequent n-grams to these eight labels. We expand acronyms and abbreviations (e.g., “pre-e” for preeclampsia, “epi” for epidural, “pp” for postpartum). Figure 3 also shows the number of birth stories assigned to each label and the n-grams used for classification.

Fiona’s Birth Story: The best laid plans go awry? Long read
 Thomas Berry Birth Story (Planned C-Section w/Complications)
 Joseph Jones Birth Story - (planned for natural birth, got an ER c-section)
 Emily Rose’s Birth Story
 My birth story! [home, unmedicated, midwife-assisted water-birth]
 I gave birth! Here’s a very long birth story for anyone interested: pitocin + epi, and a few unexpected things pp
 Damian’s Homebirth Story!
 Alice’s birth story. (long, didn’t go as planned)
 Finally wrote my birth story. Hello to Jonathon!
 Crystal’s birth story... With the funniest ending!
 DJ’s birth story - severe pre-e followed by early induction

Table 2. A random sample of paraphrased story titles. These titles are used for labeling subsets of the stories. Examples shown are paraphrased and names have been changed.

Label	Description	N-Grams	Number of stories
POSITIVE	Positively framed	[positive, <i>less-than positive</i>]	462
NEGATIVE	Negatively framed	[trauma, trigger, negative]	186
UNMEDICATED	Birth without epidural	[no epi, natural, unmedicated, epidural free, no meds, no pain meds, <i>unnatural</i>]	253
MEDICATED	Birth with epidural	[epidural, epi, <i>no epi, epidural free</i>]	192
HOME	Birth takes place at home	[home]	80
HOSPITAL	Birth takes place at hospital	[hospital]	205
FIRST	First birth for the author	[ftm, first time, first pregnancy]	91
SECOND	Second birth for the author	[stm, second]	20
C-SECTION	Birth via cesarean delivery	[cesarian, section, caesar]	328
VAGINAL	Vaginal births	[vaginal, vbac]	121

Table 3. The n-grams drawn from the story titles used to label the stories. N-grams in *italics* are negative examples which are used to disallow stories from being assigned that label.

NEGATIVE labels are usually used as a warning to readers that the story might deviate from the expected sequence of events and contain traumatic experiences. POSITIVE labels do not necessarily indicate the absence of negative events but rather the author’s framing of the story (these are often framed as a “happy ending” despite a difficult birth). MEDICATED and UNMEDICATED labels indicate the author’s decision or the necessity of using pain medication such as the epidural. This decision can be fraught as it is a focal point in the debate between the “natural” birth movement and the medicalized norm. FIRST and SECOND labels indicate whether this was the author’s first or second birth (we did not encounter specific labels for more than two births). C-SECTION and VAGINAL labels indicate whether the birth was completed via Caesarean section (surgery) or was a vaginal delivery. This decision/necessity can also be emotionally difficult as many authors express preferences for vaginal births and express disappointment and shame at needing a c-section. Some of these labels can overlap: for example, a story could be labeled both POSITIVE and UNMEDICATED.

3.4 Data Presentation

In order to balance the utility of the dataset with respect for authors' control of their own stories, we release our processing pipeline (e.g., the set of n-grams used for labeling the stories, the set of n-grams used for labeling personas). To protect the privacy of the authors and the personal medical information included in the stories, we do not release the copies of the data content nor do we include URLs pointing to specific stories. All text examples shown in this paper are paraphrased to further protect the authors' identity [14, 99]. See Section 6 for a fuller discussion of these practical and ethical tradeoffs.

4 STORY PLOTS AND CLINICAL PATHWAYS

4.1 Background: Narrative Analysis

One of our motivations in analyzing birth stories is that we expect to find clear patterns of events that occur in predictable sequences, due to a combination of biological processes and standardized medical protocols in developed countries. At the same time, birth stories can disrupt cultural narratives about "normal" births by describing the author's real experiences of pain, fear, and love as well as the medical realities and outcomes of giving birth. The task of extracting narrative structure for analysis has been formulated in multiple ways. *Plot units* are networks of affect states, where each state models the positive, negative, or neutral affect of a single character [45, 63]. *Scripts* model prototypical sequences of events [20, 80, 89], while *narrative chains* model sequences of events centered on a protagonist [19]. Narrative shapes can also be measured using sentiment scores [38, 85], as can mapping of relationships between personas [5, 54, 64]. We focus on scripts, sentiment, and personas as automatic measurements of both the "typical" and "outlier" narrative structure of birth stories as presented in our online forum setting.

Evaluation of narrative analysis and narrative understanding tasks is difficult as there is often no one correct answer. One approach is the Narrative Cloze or Story Cloze Task, which given all but one missing section of a story asks for the correct insertion (e.g., the ending of the story) [20, 73]. Another approach is to use mixtures of multiple stories and ask to pick apart the sentences from the separate stories [96]. We suggest that the structure of the birth stories collection provides a new testing ground for narrative analysis techniques.

4.2 Methods: Narrative Analysis

4.2.1 Definition of Story Time. The stories range in length from a minimum of 500 words (our selected cutoff) and a maximum of 6,057 words (see Section 3 for a fuller overview of the dataset statistics). Despite this difference in word count, the stories usually begin with the same events (arrival of the due date, water breaking, contractions starting) and end with the same events (birth and weighing of the baby, breastfeeding, leaving the hospital), though a few outliers break off in the middle of the story (e.g., sharing the story in installments). In order to compare these sequences of events across all the stories in the dataset, we divide each story into ten equal sections, and we then use these sections to calculate statistics of interest averaged over all the stories for the corresponding section. We refer to these sequences of normalized sections as *story time*.

4.2.2 Sentiment Over Time. We examine positive and negative emotion patterns across story time through the sentiment analysis tool Valence Aware Dictionary and sEntiment Reasoner (VADER) [52]. VADER is a lexicon and rule-based model designed for social media data, which performs well across domains and suits the mixture of formal and informal language found on *r/BabyBumps*. We use VADER to assign positive, negative, and compound (a normalized sum of the positive and negative scores) sentiment scores to each sentence. (Note that these polarity measurements are distinct from the labeling process described in Section 3.3.) We require that each sentence be at least

5 words long to avoid very short sentences (which will have greater weight due to the averaging) and mistakes due to our sentence segmentation. We divide each story (each set of sentences) into ten equal sections and plot the mean scores for each section.

4.2.3 *Personas Over Time.* Unlike most story collections where comparing character personas across stories is challenging [5, 7], birth stories contain a small and more easily recognized “cast of characters.” These personas are often legally defined professions, such as doctor, midwife, and nurse, making them more consistent and distinctive. Rather than using sequence-based named entity recognition, we develop a simple codebook of terms that refer to these personas. To identify personas and generate candidate terms we use a fast, accurate, and easily accessible parser⁴ [23] to rank all nouns, proper nouns, and pronouns in the collection by their frequency. We then manually collapse the most frequent n-grams into a set of personas, assigning each of these tokens to a normalized version; for example, we assign the n-grams *partner*, *husband*, and *wife* to PARTNER. Although we do not attempt to resolve most pronoun coreference, we treat first-person singular pronouns (I, me) as referring to the AUTHOR persona, and first-person plural pronouns (we, us) as a distinct persona WE.

4.2.4 *Sequences of Events.* We find that simple methods are sufficient to identify readily interpretable events and event sequences, or scripts. We train a latent Dirichlet allocation (LDA) [12] topic model with 50 topics on the birth stories collection, using 100 word chunks as the training documents. We then divide each story into 10 equal segments and plot the distributions of the topics over the resulting normalized story time.

To establish additional validity, we can compare these topics to descriptions of the birth process from health care providers such as the Mayo Clinic.⁵ We calculate the probability of transitioning between topics by finding the most probable topics for each segment of text, counting the number of times each pair of topics occurs in neighboring text segments, and normalizing by the total number of transitions. We compare the learned topic transitions to the Mayo Clinic’s document, which describes a “normal” birth path with few deviations.

4.2.5 *Community Outliers.* Using the learned transition probabilities between topics, we identify outlier stories containing less probable transitions. We rank the stories according to the summed log probabilities of the transitions of the last five 100 word chunks (the same size used for the topic model training). We limit the number of transitions in each story to five to control for the varying lengths of the stories, and we choose the last five transitions under the hypothesis that the story endings are more likely to display variation in sequences of events than the story beginnings (e.g., stories that end in emergency surgeries or unexpected trips to the hospital). For each bigram in the post titles (where authors frequently include tags or labels that frame and situate the story within the community), we average the log probabilities for all stories that include that ngram in the post title. This allows us to measure the correlation between the outlier stories and the authors’ framing of the stories.

4.3 Results: Narrative Analysis

The discovered sequences of events, persona patterns, and story likelihoods describe the narrative norms and expectations of the *r/BabyBumps* community. These norms are informed by the biological process of birthing and the medical organization and standardizing of procedures; but the specific sequence of events that the authors in *r/BabyBumps* choose to highlight also arise from this online health community’s particular expectations and priorities.

⁴<http://www.spacy.io>

⁵<https://www.mayoclinic.org/healthy-lifestyle/labor-and-delivery/in-depth/stages-of-labor/art-20046545>

4.3.1 Sentiment Over Time. Table 4 shows a small sample of the most positive and most negative sentences in the collection. The most positive sentences focus on the moments after birth, praise for the pregnant person’s partner and their help, and complimenting the baby. In contrast, the most negative sentences focus on the pain of contractions, worry that the birth plan might change (e.g., labor slowing), fear for the baby’s health, and painful side effects postpartum. Again, note that the positive polarity measurements are distinct from the positive labeling described in Section 3.3.

Sentiment	Sentence
0.99	It was the hardest day I’ve ever experienced but I’m so happy for the help I received from midwife and doctor, my family and of course my wonderful wonderful husband who I thought I loved more than was possible but now I love even more
0.95	1pm: Anesthesiologist arrives, I feel slightly worried but my amazing nurse explains everything perfectly and i realize it isn’t that scary as i expected
0.93	Even though we only had a small amount of time together, she would up being the best at keeping both of us focused on the positive side of the contractions and she also helped with breathing and surviving them
-0.98	I’m not going to refer to contractions as "waves" or any of that feel-good BS; it was horrible pain, constant pain, pain that didn’t change but was just fiery burning pain, as if my body was turning itself inside out
-0.98	Also, the monitors on my stomach were annoying and distracting me when I was trying to get through the contractions, but they wouldn’t change them, and the blood pressure cuff really hurt me worse than even the contractions, which the doctor decided meant that the pitocin should be increased but which actually meant that it just hurt
-0.98	She hated the whole experience; she hated the clothing, she hated the pillow, she hated the gel in her eyes, she hated being fed, and she hated being held

Table 4. The **most positive** and **most negative** sentences in the birth stories collection. Sentiment scores are produced using VADER, a lexical and rule-based system. The most positive sentences tend to focus on the author’s support team while the most negative sentences tend to focus on pain and medical emergencies. (All examples are paraphrased for privacy.)

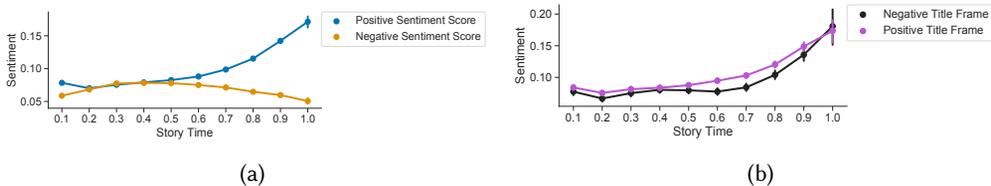


Fig. 2. (a) Positive and negative sentiment over story time, measured using the VADER sentiment tool. Higher scores indicate greater sentiment strength. (b) The compound sentiment score over story time for stories labeled POSITIVE or NEGATIVE by the author in the post title.

Figure 2 shows the mean positive and negative sentiment scores over story time. We observe that positive sentiment dips in the middle of the stories and then climbs towards the end of the stories, while negative sentiment builds to a gentle peak in the middle of the stories and then recedes. These patterns match prior expectations that the most painful and difficult parts of childbirth are near the middle of the story (contractions heightening, pushing the baby, emerging complications such as c-section).

4.3.2 Personas Over Time. Apart from the narrator, we consistently find a PARTNER and other FAMILY members. In addition we choose to identify a distinct first-person plural persona WE. Medical personnel include a NURSE, physicians (DOCTOR and ANESTHESIOLOGIST), and a MIDWIFE. As shown in Table 5, the AUTHOR is the most frequent, with an average of 74 mentions per story. Not all of

these personas appear in every story: not all births are attended by a midwife or doula, and there is one story in the collection that is told exclusively from the second person point of view of both the partner and pregnant person. Of these medical personnel, the DOCTOR is the most frequent, both in total n-gram count across the collection and in the number of stories containing at least one mention. The MIDWIFE is referenced more often than the FAMILY and ANESTHESIOLOGIST, but the FAMILY appears more consistently across the stories and the ANESTHESIOLOGIST appears equally as often, indicating the small but consistent role they play. Nearly double the percentage of stories in our sample include mentions of a midwife as compared to Bylund [15].

Persona	N-Grams	Total Mentions	Stories Containing Mentions	Average Mentions per Story
AUTHOR	I, me, myself	210,795	2,846	74.0
We	we, us, ourselves	24,757	2,764	8.7
BABY	baby, son, daughter	14,309	2,668	5.0
DOCTOR	doctor, dr, doc, ob, obgyn, gynecologist, physician	10,025	2,262	3.5
PARTNER	partner, husband, wife	8,998	2,006	3.2
NURSE	nurse	7,080	2,012	2.5
MIDWIFE	midwife	4,069	886	1.4
FAMILY	mom, dad, mother, father, brother, sister	3,490	1,365	1.2
ANESTHESIOLOGIST	anesthesiologist	1,398	876	0.5
DOULA	doula	896	256	0.3

Table 5. Personas identified in the birth stories collection and the n-grams used to classify the personas.

In some cases births are also attended by a DOULA, a non-medical assistant or “birth coach.” The support of a doula can significantly reduce costs and improve birth outcomes, particularly for disadvantaged groups [46, 59, 95], yet the possibility of hiring a doula is not guaranteed. The doula has been described as an ambiguous, unexpected persona in the U.S. cultural birth narrative who acts as a liaison or “bridge” between other personas [51]. Doulas and midwives are typically viewed as having lost power during the last century’s medicalization of childbirth and prioritization of medical staff [48, 51, 62]. Births supported by a doula are rare, but there has been a recent increase from 3% of births in the U.S. in 2006 to 6% in 2012⁶ [33]. We observe that the DOULA is the least frequent persona in the stories, with only 0.3 mentions per story.

Figure 4 shows the frequencies of the personas over story time. These plots show string matches with the lists mapped to each persona type (see Table 5). As expected, the frequencies of medical personnel like the NURSE and ANESTHESIOLOGIST peak near the middle of the stories while mentions of the BABY peak near the end of the stories. Interestingly, there is a decrease in mentions of WE near the middle of the story (even though AUTHOR mentions remain constant). This could indicate that the partner and family members become less useful as supporters at the height of contractions, when the medical personas become more important. A peak in mentions of the DOULA near the beginning of the story could emphasize the unique nature of the doula as compared to the other medical personas, which peak later in the stories.

4.3.3 Sequences of Events. The strengths of individual topics over time tracks with our expectations of the stories, confirming that the dataset contains enough shared structure to be easily detectable. For example, we can see in Figure 3 that water breaking happens near the beginning of the story; a topic about contractions starting and/or pitocin (a medication that induces labor) being administered peaks around 30% into the story; there is little sleep in the middle of the story; photos are shared sometimes at the beginning of the story but usually near the end of the story. We note that the

⁶<https://transform.childbirthconnection.org/reports/listeningtomothers/trends/>

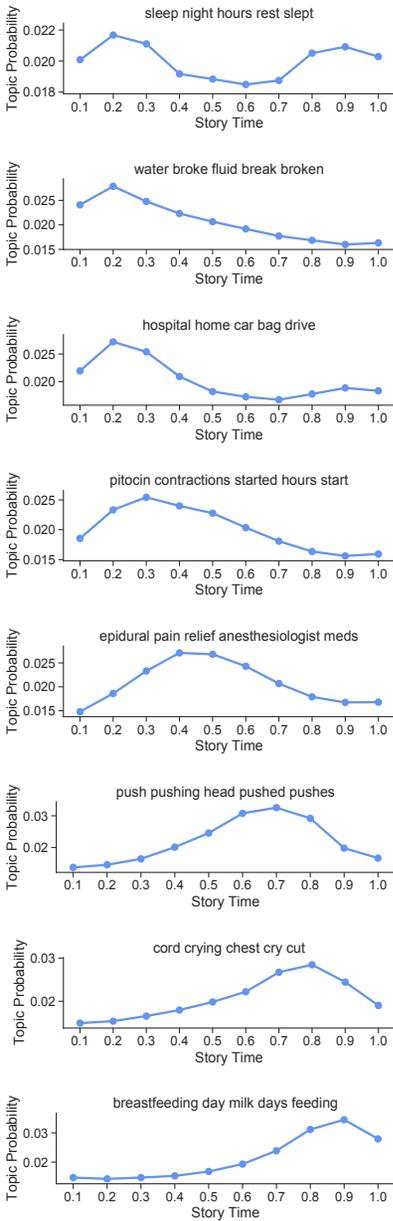


Fig. 3. A selection of topics over time. Plots are labeled with the five highest probability words for each topic. Results show the probability for each topic at 10% intervals of story time, averaged across all stories.

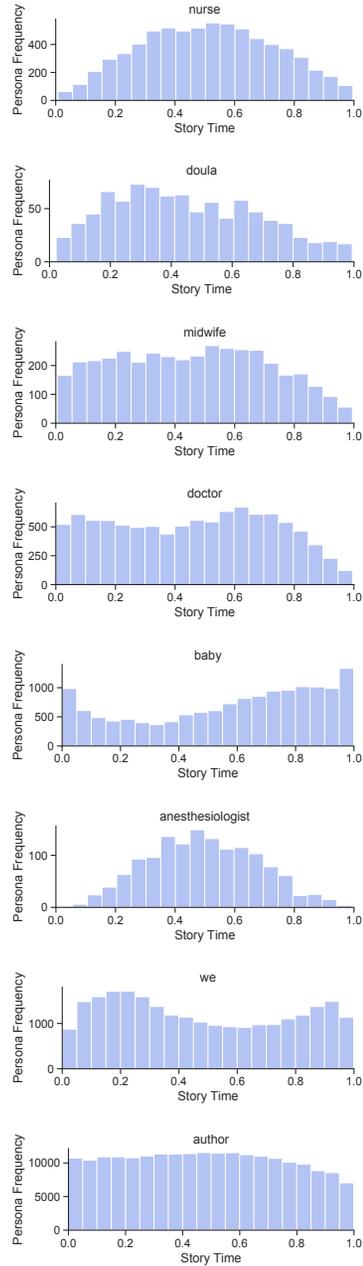


Fig. 4. Histograms showing the frequencies of persona mentions over story time. Some entities (e.g., *author*) are consistently more frequent than rare entities (e.g., *doula*). Some frequency patterns are expected while others are surprising (e.g., frequency of *we* decreases near the middle of the stories).

model was trained with no information about word position within stories, but topics nevertheless exhibit strong temporal clustering.

Although we do not necessarily expect the authors’ accounts of childbirth to track with medical descriptions, we find strong correspondences. The topic “bed back tub ball sitting” is reflected in descriptions of Stage 1 active labor: “consider these ways to promote comfort during active labor: Change positions, Roll on a large rubber ball (birthing ball), Take a warm shower or bath.” The topic “push pushing head pushed pushes” also tracks with the Mayo Clinic description of Stage 2 (birthing) “Push! Your health care provider will ask you to bear down during each contraction or tell you when to push. Or you might be asked to push when you feel the need.”

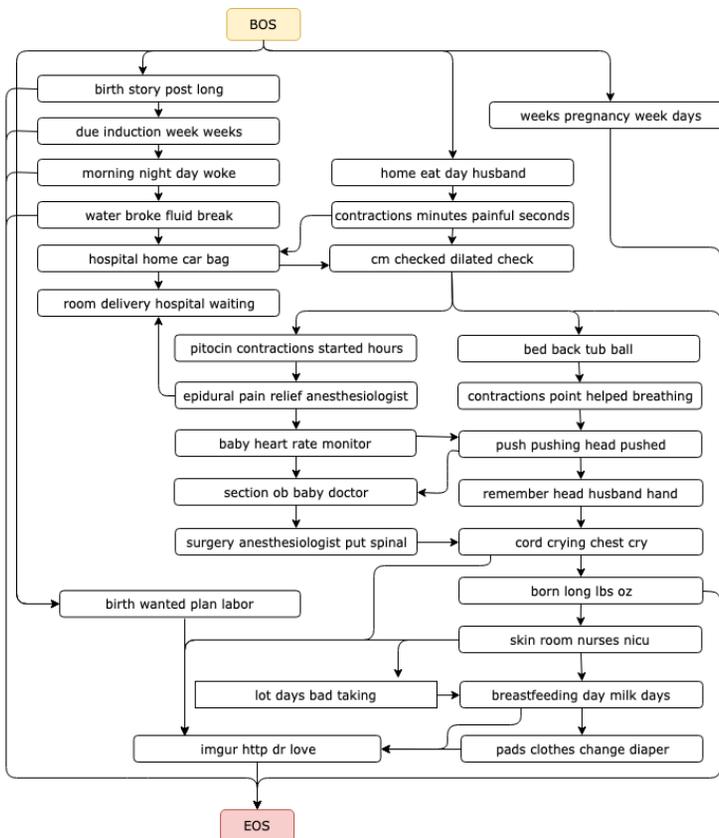


Fig. 5. Flowchart of the most probable topic transitions (above 0.2%). We removed one orphan node without a parent path leading to the beginning of story (BOS) state.

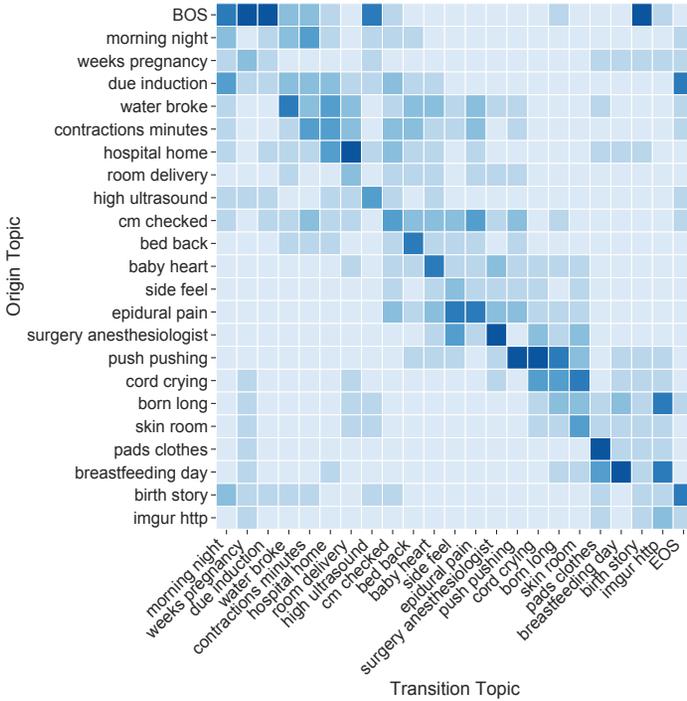


Fig. 6. Transitions between topics and the beginning of story (BOS) and end of story (EOS) states. Only the topics with the highest probability transitions are shown in this figure. Labels are the two most probable words for each topic. The *Origin Topic* is the most probable topic of the current section of story text. The *Transition Topic* is the most probable topic for the next section of story text.

Figures 5 and 6 demonstrate that while there is common structure across the stories shared in the birth stories community, there are also dividing paths and exceptions to most rules. Figure 5 shows some of these paths; for example, some people require pitocin, a medication that can induce labor, while others begin labor without this intervention. Figure 5 also includes events that are important to the pregnant person but not “medically important”; for example, the logistics of driving to the hospital and waiting in the lobby before receiving a room can be stressful and require planning on the part of the author, but these events are not included in the Mayo Clinic information. Likewise, the Mayo Clinic description contains events such as the delivery of the placenta that are less frequently identified in birth stories.

Highest Probability	Bigram	Lowest Probability	Bigram
-34.15	positive induction	-35.88	post partum
-34.38	positive hospital	-35.90	unplanned c
-34.51	story ftm	-35.93	baby birth
-34.53	positive medicated	-35.99	traumatic birth
-34.63	med free	-36.02	story baby
-34.67	story long	-36.04	positive birth
-34.67	story finally	-36.09	birth story
-34.71	hospital birth	-36.14	pp advice
-34.82	story weeks	-36.15	home birth
-34.83	weeks pp	-36.18	natural birth
-34.86	hour labor	-36.19	induction epidural
-34.92	failed induction	-36.23	emergency c
-34.94	vaginal delivery	-36.30	belated birth
-34.99	story hospital	-36.41	positive unmedicated
-35.03	vaginal birth	-36.42	story plus
-35.07	baby tax	-36.43	slightly traumatic
-35.07	super positive	-36.47	c section
-35.08	super long	-36.49	happy ending
-35.22	mostly positive	-36.85	trigger warning
-35.24	late birth	-36.86	long birth

Table 6. The bigrams drawn from the post titles associated with the most and least probable stories. Probabilities represent the averages of the summed log probabilities of the topic transitions in a story (see Section 4.2.4 for a full explanation). Lower scores indicate stories with more unusual topic transitions (sequences of events).

4.3.4 Community Outliers. Despite the strong evidence of shared narrative structures, not all stories fit into these community norms. In Table 6, we show how outlier stories (stories whose sequences of events are unusual in this community) are framed by the authors. These results indicate a) the types of births that this community normalizes and b) the outlier stories whose topic sequences are unusual for this particular community of storytellers.

As expected, stories with low probability topic transitions are more likely to have titles that include bigrams like *emergency c* and *unplanned c* that indicate unplanned surgeries. We also find that low probability stories are associated with bigrams like *trigger warning* and *slightly traumatic* that indicate negative experiences. We find that the titles of stories with high probability topic transitions are more likely to include bigrams like *positive induction*, *positive hospital*, *positive medicated* which indicated medicalized births, while stories with low probability topic transitions are more likely to have titles that include bigrams like *home birth*, *natural birth*, and *positive unmedicated* which indicate home and/or unmedicated births. Interestingly, the bigram *happy ending* is strongly associated with low probability sequences of topics; authors in this community often choose to emphasize the story’s “happy ending” despite unexpected and/or negative events.

These results could be seen as supporting the findings in Bylund [15] that negative emotions in birth stories are associated with a lack of control, or in this case, unexpected and unplanned events. The emphasis on “happy endings” could imply a re-framing of the birth experience in which the author controls the frame if not the sequence of events.

5 NEGOTIATION OF POWER RELATIONSHIPS BETWEEN PERSONAS

5.1 Background: Power and Lack of Power

The motivations guiding the writing and sharing of birth stories in an online community are varied, but we identify three branches related to the power of the author. These include the feeling of disempowerment during birth, circumventing of surveillance, and re-writing societal narratives.

One possible motivation for authors to share their birth stories online is that they have felt disempowered during their medical experiences. The technical knowledge differential between

physicians and patients and the hierarchical structure of medicine can result in an asymmetrical power relationship and the dismissal of patients' testimonies as unreliable [17]. Prior work has shown that women are often not trusted when describing their own pain; women's verbal reports are discounted by doctors as "constructed or exaggerated" and they are more likely than men to be inadequately treated for their pain [21, 25, 34, 36, 49]. Pain is subjective and can only be measured via self-reporting; to distrust pain reports implies a distrust of and arbitration over people's experiences and stories. The ability to process the trauma of birth often requires that the pregnant person fill in "missing pieces" of their experience to regain control [2]. Callister [16] identifies feelings of inadequacy and disappointment as motivators for sharing birth stories.

We also consider birth stories as a way for pregnant people to disrupt surveillance of their choices and bodies. Tangherlini [93]) describes how paramedics use stories as an alternative place to frame and negotiate power hierarchies away from the surveillance of their supervisors and dispatchers. Pregnant people are monitored both personally (by family and friends) and medically (by doctors, nurses, midwives, and other medical staff), and the experience of society's "gaze" extends to even the most intimate moments of birth [18, 43]. Pregnant people report feeling "drained" from this experience of being watched [8]. The birth stories community might offer a way of circumventing this surveillance by allowing for the creation of personal narratives that re-frame the power and agency of the author in comparison to the doctors, nurses, family members, and other personas.

Re-writing the story could also help the pregnant person to escape proscribed societal narratives about birth that assign blame when birthing goes wrong. Whether these societal narratives originate from the natural birth movement (home births, the idea that women are "made" to give birth and therefore capable without intervention) or the medical establishment (viewing pregnancy as a pathological state that requires intervention), both can warp personal views of childbirth and over-emphasize the amount of control the pregnant person and the medical establishment have to ensure a "happy ending" [11, 18, 58, 62]. From this perspective, it could make sense for the author of a birth story to de-emphasize their own power.

Automatic analysis of relationships between personas could provide evidence of how the birth stories community views the power of different actors in the stories and confirm our hypotheses about their motivations. Prior work has tried to measure power and/or reconstruct power hierarchies from natural language. Prabhakaran et al. [82] train a classifier on dialogs to predict whether a person has situational power, while Jockers and Kirilloff [57] learn associations between verbs and genders in novels and examine the resulting sets of verbs for their level of *agency*, a concept related to power. Danescu-Niculescu-Mizil et al. [28] measure power via language coordination, while Prabhakaran et al. [81] use poll rankings as a proxy for power labels in political debates. Sap et al. [87] annotate a set of verbs with power labels and use these labels to measure framing differences between genders in movie scripts.

5.2 Methods: Measurements of Power

To examine the power dynamics between medical professionals, other personas, and the author of the story, we employ a lexicon of power frames curated by Sap et al. [87]. We adopt their definition of power as *authority* and a *high level of control*, as these correspond well with the hierarchies in the medical institution and the preoccupation with choice that is frequent in our set of stories. This lexicon includes 2,155 verbs, each of which is mapped to a directional power label (*agent*, *theme*, *equal*). These labels indicate whether the subject of the verb has more power, equal power, or less power than the object of the verb. For example, the verb *accuse* is labeled with the *agent* having power, while the verb *concede* is labeled with the *theme* having power. This lexicon was designed to measure gender bias in a dataset of movie scripts, but we find that the majority (62%) of the labeled verbs occur in our dataset, despite its small size and specific focus, and we examine the

individual verbs most associated with each persona type to ensure that no one outlier is incorrectly overweighting the power scores. The following sentences are paraphrased examples of verbs labeled with power, where in (1) the NURSE has a high power score and the AUTHOR has a lower power score and in (2) the DOCTOR has a high power score.

- (1) *The only time I got upset was when the nurse **accused** me of not feeding my child.*
- (2) *The **doctor** **broke** my water.*

Following Sap et al. [87], we lemmatize the verbs in the lexicon and parse the birth stories using the spaCy parser [50], and we match verbs where our personas of interest are the agent or theme to the verbs in the frames lexicon (see Section 4.2.3 for a description of the persona labeling process). If the power label positively corresponds to the persona (e.g., verb labeled with *agent*, persona is the subject of the verb), we add one to the persona’s score, and if the label negatively corresponds to the persona (e.g., verb labeled *agent*, persona is the direct object of the verb), we subtract one from the persona’s score. Finally, we normalize the power scores by the number of times the persona or dyad occurs as subjects and direct objects in the corpus, and average the results over 20 bootstrap samples of the birth stories dataset. Not all personas occur in all stories (see Table 5 for persona frequencies), so the power scores represent averages of scores from only those stories in which a persona occurs.

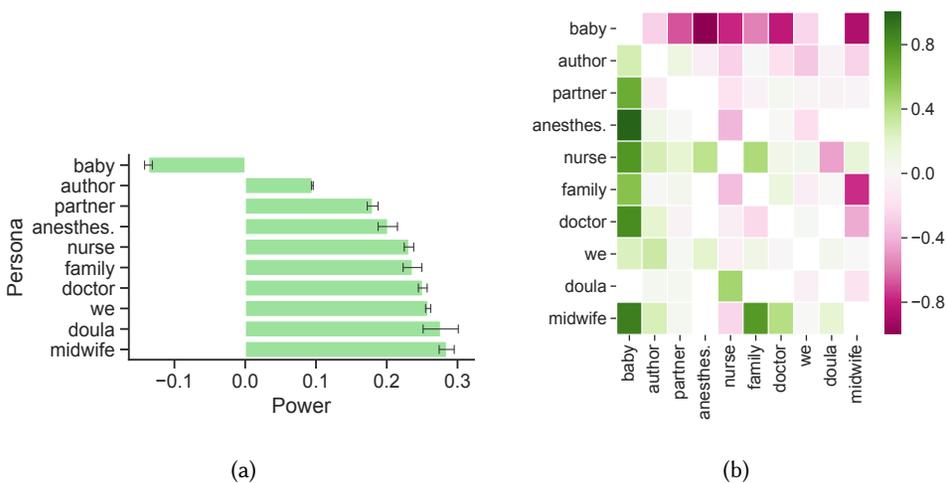


Fig. 7. (a) Power scores for each persona. Error bars show standard deviation over 20 bootstrap samples of the collection. (b) Estimated power of personas (rows) over other personas (columns). The NURSE is consistently framed as more powerful than the other personas, except for the DOULA.

5.3 Results: Measurements of Power

Power is a frequent topic in the *r/BabyBumps* community. These discussions often center on the author’s decision to make decisions about pain medications and birth settings. We highlight a few paraphrased examples below.

- (1) *I managed to not take pitocin, and I’m glad that I felt some of the discomfort of pushing. It was kinda **empowering** and I’m proud of what I did!*
- (2) *Nothing about the birth was what I expected, but I felt **empowered** to have had a natural birth and mostly without the assistance of my doctor!*

- (3) *The nurse wanted to monitor the baby but honestly, it really felt so **empowering** to say no to something. She backed down after I said that and actually was a great nurse.*

These examples highlight the common contexts in which empowerment is discussed in this community. Example (3) describes the author’s perception that most birthing decisions are made by personas other than the author. However, while the examples shown are all positive (in the sense that the author feels empowered), these moments represent explicit framing by the author, post-birth, of what they view as empowering. To explore subtler attitudes about power, we turn to automatic measurement of power hierarchies between personas.

As shown in Figure 7, the AUTHOR consistently frames themselves as having the least power except for the BABY, even though the authors are the narrators and framers of the stories. The clinicians (MIDWIFE, DOULA, DOCTOR, NURSE, ANESTHESIOLOGIST) are more often written about using high power framing. This could provide support for the theory that the authors have felt disempowered during their medical experiences, and that this lack of control motivated the writing of the stories, though it does not support a “re-writing” in which the authors have more power. These results could also support the view that the authors are resisting a societal narrative in which mothers are blamed for the supposed control they have over the events of the birth.

The heatmap in Figure 7 illuminates the power relationships between specific pairs of personas; a persona could conceivably have a high power score overall but a lower score when only the interactions with another persona are counted. As expected, the verbs used with the BABY have consistently lower power scores than with any of the other personas (except for the DOULA, whose relationship with the BABY appears neutral). Again, verbs involving the AUTHOR consistently frame them as the less powerful of the pair, except when the matched person is the BABY. Surprisingly, it is the NURSE and MIDWIFE, not the DOCTOR, who are framed as dominating most relationships. Only the DOULA occurs with verbs that indicate more power than the NURSE; this supports the doula’s hypothesized special place in the hierarchy of personas. The MIDWIFE is also a powerful persona, but unlike the DOULA, she is framed as having less power than the NURSE.

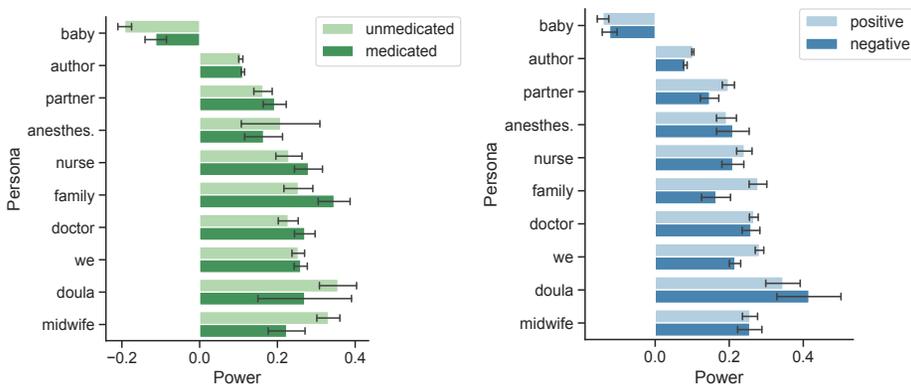


Fig. 8. Power scores for each persona across different labeled sections of the collection. The PARTNER, AUTHOR, FAMILY, and WE all display greater power in POSITIVE stories, while the MIDWIFE displays greater power in UNMEDICATED stories.

Figure 8 shows how these power measurements can shift for labeled cross-sections of the dataset (see Section 3.3 for a description of the labeling process). We find that the MIDWIFE has a higher power score in UNMEDICATED stories, while the FAMILY has a higher power score in MEDICATED

stories. The AUTHOR, FAMILY, PARTNER, and WE have more power in POSITIVE stories. This could indicate that when the author and their support group are offered more power, stories are presented more positively; however, it could also indicate that when the birth goes well, the author chooses to frame the story in a style that emphasizes their support group. These results support the finding in [15] that positive stories are more likely to involve decision making on the part of the authors.

	Author	Baby	Doctor	Nurse	Midwife	Doula
get_subj	0.0404	do_subj - 0.0129	call_obj - 0.0760	check_subj - 0.0499	call_obj - 0.0876	call_obj - 0.1412
know_subj	0.0219	want_obj - 0.0113	check_subj - 0.0271	call_obj - 0.0382	check_subj - 0.0499	suggest_subj - 0.0357
start_subj	0.0215	get_subj - 0.0110	decide_subj - 0.0221	ask_obj - 0.0272	suggest_subj - 0.0203	get_subj - 0.0306
do_subj	0.0199	start_subj - 0.0081	get_subj - 0.0159	give_subj - 0.0171	ask_obj - 0.0168	hold_subj - 0.0144
push_subj	0.0123	drop_subj - 0.0072	do_subj - 0.0145	get_subj - 0.0149	give_subj - 0.0152	keep_subj - 0.0128
make_subj	0.0070	make_subj - 0.0063	show_subj - 0.0116	keep_subj - 0.0147	get_subj - 0.0146	show_subj - 0.0122
wake_subj	0.0069	turn_subj - 0.0055	break_subj - 0.0105	hold_subj - 0.0116	decide_subj - 0.0143	remind_subj - 0.0117
keep_subj	0.0065	show_subj - 0.0027	give_subj - 0.0091	start_subj - 0.0113	do_subj - 0.0125	give_subj - 0.0101
decide_subj	0.0059	enjoy_obj - 0.0022	start_subj - 0.0091	do_subj - 0.0092	break_subj - 0.0103	start_subj - 0.0090
end_subj	0.0038	decide_subj - 0.0021	make_subj - 0.0088	bring_subj - 0.0091	show_subj - 0.0081	make_subj - 0.0080
get_obj	-0.0028	give_obj - 0.0053	wait_subj - -0.0014	leave_obj - -0.0016	announce_subj - -0.0018	massage_subj - -0.0022
believe_subj	-0.0031	catch_obj - -0.0057	like_subj - -0.0014	believe_subj - -0.0018	hear_subj - -0.0021	cheer_subj - -0.0024
wait_subj	-0.0034	need_subj - -0.0072	mention_subj - -0.0026	bring_obj - -0.0018	ring_obj - -0.0026	fill_subj - -0.0025
lose_subj	-0.0035	reach_obj - -0.0076	reach_subj - -0.0029	mention_subj - -0.0023	reach_subj - -0.0036	recognize_subj - -0.0027
hear_subj	-0.0057	bring_obj - -0.0089	explain_subj - -0.0036	explain_subj - -0.0048	offer_subj - -0.0042	alarm_obj - -0.0028
call_subj	-0.0069	put_obj - -0.0113	offer_subj - -0.0039	offer_subj - -0.0052	explain_subj - -0.0043	warn_obj - -0.0028
ask_subj	-0.0097	deliver_obj - -0.0196	call_subj - -0.0089	want_subj - -0.0060	want_subj - -0.0096	join_subj - -0.0057
need_subj	-0.0117	push_obj - -0.0240	ask_subj - -0.0133	call_subj - -0.0174	get_obj - -0.0103	get_obj - -0.0141
check_obj	-0.0137	hold_obj - -0.0306	get_obj - -0.0136	get_obj - -0.0195	call_subj - -0.0123	ask_subj - -0.0171
want_subj	-0.0266	get_obj - -0.0369	want_subj - -0.0152	ask_subj - -0.0238	ask_subj - -0.0218	hire_obj - -0.0438

Fig. 9. Most frequent verbs from the power lexicon associated with each persona in the birth stories corpus. Green indicates a positive power contribution, while pink indicates a negative power contribution. The cell values indicate the proportion of persona mentions with the given verb and power relationship.

Figure 9 demonstrates the coverage of the lexicon; we find a wide range of words contribute positively and negatively to the power scores. Positive scores indicate verbs that contributed positively to the power score (e.g., the persona appeared as the subject of a verb labeled with the *agent* power relationship), while negative scores indicated verbs that contributed negatively to the power score (e.g., the persona appeared as the direct object of a verb labeled with the *agent* power relationship). We observe that the verbs most contributing to the BABY's negative power score are *get*, *hold*, *push*, *deliver*, *put*, *bring*, and *feed*, whereas the verbs most contributing to the DOCTOR's positive power score are *calls*, *checks*, and *decides*.

Like the DOCTOR and NURSE, the DOULA's positive score is owed to verbs like *call*, *get*, and *show*. We can see from these ranked verbs that the framing of the DOULA's role is centered around support and encouragement while the other medical professionals make decisions and perform actions on the pregnant person (e.g., *check*, *break [water]*). The DOULA also often appears as the object of the verb *hire*, which indicates the DOULA's inside/outside position in the medical organization as patient advocate, birthing coach, and optional, sometimes uncredentialed supplement to the medical professionals.

We can use these results to re-explore the set of birth stories and discover specific texts where power hierarchies are enacted.

- (1) *I had planned on a natural birth, had **hired** a doula, read books, all of those things. And here I was on a random Monday just going about my business when the baby **decided** to arrive. It took me a while to process all of it.*
- (2) *The nurse said that if I wasn't more dilated by 12PM that she was going to start me on pitocin. I was scared about this happening, I had heard a lot about the cascading interventions that often*

happen in hospitals. That's why I hired a doula, so that I could have another point of view and someone to guide me so I could have this birth naturally.

Despite the low power score of the verb *hire*, the themes of power and control re-emerge in these examples. Example (1) highlights the actions the author has taken to control the events, and the disruption of this control by the baby. Both examples highlight the author's perspective of the doula as a powerful persona who can circumvent the medical hierarchy.

6 DISCUSSION

6.1 Lessons Learned

We view this study as a “close reading” via a computational lens of a specific online community. This computational reading allows us to discover patterns and outliers within our target community and also to prove the feasibility of similar research on other communities that share medical stories. We do not claim that the patterns discovered here will hold for all other collections of birth stories; instead, we claim that such analysis can a) provide specific, statistical evidence of patterns suggested or observed in prior work and b) prompt further research in similar communities.

Narrative analysis is a challenging task in natural language processing, partly because of the difficulty of creating datasets that are realistic but not too complex for current models. We suggest that birth stories hit a sweet spot between formulaic, artificial datasets and complex, organic datasets; they are organically created (written spontaneously by the authors), share narrative structure and are constrained by topic despite each story being unique, are plentiful enough to act as training data for machine learning models, and suggest real-world motivations for their analysis. We successfully identify sentiment, topic and person-based patterns that demonstrate the recoverable narrative qualities of birth stories.

By uncovering this shared structure, we discovered events not described in medical literature that are nevertheless important to the authors in this community. We also use the learned topic patterns to discover outlier stories whose sequences of events do not match those of the community's expectations, and we examine the framing of these stories. The authors of these stories emphasize unexpected events, negative or triggering experiences, and “happy endings” in their titling and framing of the posts. These results suggest that a lack of control is associated with negative emotions, as found in Bylund [15], and that in this community, reframing of these unexpected events around “happy endings” is common.

Our analysis of power reveals a shared understanding of power hierarchies in *r/BabyBumps*. On average, the medical professionals are most often assigned as the agent of verbs that imply power, while the author is most often assigned as either the theme of these verbs or as the agent of verbs that imply lack of power. In this community, on average, the author is acted upon while others take action. Doulas dominate these power metrics, with average scores higher than doctors or nurses, which suggest that the doulas, as patient advocates, could lend power to the patient's voice. These results are supported by prior work indicating that pregnant people often feel disempowered during the birthing process.

Many of our methods rely on averaging across bootstrapped sets of stories in order to recover overarching themes in this community. We recognize that the “average” story is not representative of every story in this community. We explore the question of outlier stories in Section 4.3.4 where we use the learned pattern of events to identify stories whose sequences of events are unusual. The authors often label these stories as *unexpected*, which confirms our interpretation of these outliers as stories that fall outside the expectations of the community. In this sense, discovering archetypes through averaging is what allows us to discover by comparison stories in the minority.

For medical professionals, these results could help inform care decisions and priorities. While doctors, nurses, and other medical professionals observe hundreds of births every month, they do not observe these births through the pregnant person's eyes, and their interactions with the pregnant person are mediated by their power differential (the medical professional's technical knowledge, their place in the hierarchy of the hospital setting, their gender, race, age, and education level, etc.). Birth stories allow these professionals to view a routine procedure through the fresh eyes of a person who is experiencing pregnancy and birth for perhaps the first time and to discover areas that could be improved for the patient, like the events (e.g., traveling to and arriving at the hospital) that are prominent in birth stories but not usually highlighted in medical documentation of births. Postpartum depression can be alleviated through attention to the patient's emotional needs and feelings of agency during the birth [16, 91], and our work has highlighted that these needs sometimes go unmet within this community.

While the experiences of all pregnant people are valuable, we particularly highlight the importance of listening to underrepresented perspectives. While we were not able to control for race, education level, or other demographic variables in this study, we hope that our results show that such computational analysis of birth stories is feasible and valuable. We have demonstrated that birth stories can highlight experiences missing from the dominant medical narratives of birth.

6.2 Ethical Considerations

The dataset of unsolicited birth stories is a valuable resource for analysis of an online health community, and we encourage further work on this and other medical narrative datasets which prioritize the patient's voice. Many birth stories recount experiences in which the pregnant person felt that they were not empowered, and drawing attention to these voices can benefit the broader community of pregnant people. However, we caution practitioners to handle this data with care. Research value, even for the community of pregnant people, and reproducibility must be balanced against the specific ethical concerns surrounding publicly shared medical data [1, 56, 94]. We identify a series of tensions inspired by both prior work on ethical use of online medical data and the three guiding principles of the Belmont Report: *respect for persons*, *beneficence*, and *justice* [39].

Sharing a complete dataset of the collected birth stories would allow other researchers to directly reproduce and evaluate our results as well as train their own computational models. Reproduction is particularly important for this collection as the self-selection of the authors can contribute to biases in the dataset [56]. Furthermore, publication of the data would respect the contributions of the authors as "amateur artists" who should be given credit for their creative work [14].

However, copying and storing the birth stories removes the data from the control of the authors, and this would undermine the study's *respect for persons*. While the stories posted to Reddit are already public, copying that data prevents the authors from removing or editing their stories. Birth stories contain extremely personal medical, interpersonal, and emotional information not just about the author but also about the baby, who cannot consent to this public sharing. While the authors posted their stories to a public venue, our own exploration of their motivations (e.g., resistance against surveillance, regaining power) does not include providing material for researchers; indeed, studies on other social websites [37] show that most users are unaware that their public data is being used for research and would revoke permission if possible.

One possible method to support both reproducibility and privacy is to release the URLs of the data points, rather than the content (as in the Twitter API terms, which allow sharing of Tweet IDs but not the Tweet content). This maintains the user's ability to delete the content at any time and remove it from the dataset, while also allowing researchers to directly compare their results. Another tactic that allows for data analysis without compromising the privacy of the authors is to only share *paraphrased* examples from the dataset [14, 99]. This both conceals private and/or

identifying information and prevents the audience from searching for exact string matches in order to identify the source story.

Due to the very sensitive nature of the dataset, we choose to not release either the dataset or the URLs and instead prioritize the authors' protection. To support the replication of our results on other birth stories datasets, we release our labeling pipeline (e.g., n-grams used for labeling, pre-processing steps). We follow the recommendations in Bruckman [14] and Yang et al. [99] to mask the stories by only providing paraphrases rather than exact text snippets in all of the examples highlighted in this paper, which minimizes the possible identification of and harm to the authors. These potential harms are weighed against our hope that the results of our study will aid in the understanding of modern experiences of pregnancy and birth and spur further research that centers the pregnant person's voice. We also plan to share the final paper and a public-facing blogpost with the *r/BabyBumps* community to support the principle of *respect for persons*. We strongly recommend that researchers using this data follow these same privacy-motivated guidelines.

6.3 Limitations

Our collection of birth stories was constrained to a single online forum, *r/BabyBumps*, and some of our results might not extend to other data sources. The unsolicited nature of these birth stories is valuable, as it removes some potential biases, such as framing of the story to meet the researcher's expectations, which has been a concern in studies of oral birth stories [18]. However, the self-selection of the authors can lead to a sample of stories that does not accurately represent the true population of births [56]. For example, natural (unmedicated and/or home) births are over-represented in our dataset; the slowly rising national average of home births is about 1.6% [66, 68], while we observe at least double that percentage in our labeled data. None of the stories in our dataset describe miscarriage or pregnancy loss, though these events are common; it is estimated that 10-20% of pregnancies end in miscarriage⁷.

Likewise, our lack of demographic data does not allow us to control for race, socioeconomic status, previous births, marital status, prenatal care, or comorbidities (e.g., depression) and medication. Each of these features can influence the story trajectory [27, 53, 71]. Therefore, medical patterns observed in this data need to be replicated before wider conclusions can be drawn. We hope our results can prompt new discussions and emphasize the importance of the pregnant person's voice.

Not all the stories in our dataset fit cleanly into the narrative norms described in our results. Some stories begin much earlier, with descriptions of the entire pregnancy, and others end much later, with descriptions of life and complications postpartum. Indeed, as discussed in Section 4.3.3, we find that many stories follow "diverging paths", e.g., when things go wrong or the author makes decisions that run counter to the community norms. While this means that our dataset is not as clean as an artificial set of stories, the clear topical and persona-based patterns that we discover indicate that this narrative structure is shared across many of the stories.

6.4 Future Work

First, expanding this computational study with author interviews and participant feedback could substantially help to validate our findings. Moving from narratives to narrower sentiment concepts, we might focus on descriptions of pain, which feature heavily in birth stories. We might examine how the authors relate this pain to other experiences (e.g., "the first contractions felt like light menstrual cramps") and how these descriptions progress and change through the arc of the story. We might also investigate what information is left out of these stories. Although the stories are often explicit, revealing intimate details of the birth, some aspects are still withheld, possibly for

⁷<https://www.mayoclinic.org/diseases-conditions/pregnancy-loss-miscarriage/symptoms-causes/syc-20354298>

privacy, to hide embarrassing details, or to avoid painful memories. Similar behavior is observed in the comments following these stories. An “anything goes” attitude prevails, and yet some topics are still left unmentioned. Beyond birth stories, we can extend our analysis to information seeking and the types of questions asked in pregnancy and healthcare forums. Do they ask more objective questions (e.g., *What does it mean for your water to break?*) or subjective questions (e.g., *What does an epidural feel like?*)? There are also clear extensions of this work in other narratives, such as trauma narratives and blogs of chronic illnesses.

7 CONCLUSION

We performed a first computational analysis of a large dataset of birth stories from an online forum, and we discovered clear narrative patterns of sentiment, personas, and event sequences in these stories. This dataset is a valuable resource both for exploration of the narrative norms and expectations arising in a particular online health community and for exploring the patient’s point of view during a sometimes traumatic medical event. We showed that branching narrative pathways exist across these stories and that there is no single “normal” birth story, and we explored the power framing of the personas appearing in the stories, finding that the community frames the author as disempowered and the doula as powerful. We made practical recommendations for the ethical use of this and similar datasets derived from publicly posted stories, and we hope that our analysis will spur further computational research into medical narratives.

8 ACKNOWLEDGEMENTS

Special thanks to Katherine Antoniak, Sharifa Sultana, and our anonymous reviewers for their thorough and constructive comments. This work was supported by NSF #1526155, #1652536, and the Alfred P. Sloan Foundation.

REFERENCES

- [1] Jacob Abbott, Haley MacLeod, Novia Nurain, Gustave Ekobe, and Sameer Patil. 2019. Local Standards for Anonymization Practices in Health, Wellness, Accessibility, and Aging Research at CHI. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 462, 14 pages. <https://doi.org/10.1145/3290605.3300692>
- [2] Dyanne D Affonso. 1977. “Missing pieces”—A study of postpartum feelings. *Birth* 4, 4 (1977), 159–164.
- [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, 158.
- [4] Danielle Arigo and Joshua M Smyth. 2012. The benefits of expressive writing on sleep difficulty and appearance concerns for college women. *Psychology & Health* 27, 2 (2012), 210–226.
- [5] David Bamman, Brendan O’Connor, and Noah A. Smith. 2013. Learning Latent Personas of Film Characters. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Sofia, Bulgaria, 352–361. <https://www.aclweb.org/anthology/P13-1035>
- [6] David Bamman and Noah A. Smith. 2014. Unsupervised Discovery of Biographical Structure from Text. *Transactions of the Association for Computational Linguistics* 2 (2014), 363–376. https://doi.org/10.1162/tacl_a_00189
- [7] David Bamman, Ted Underwood, and Noah A. Smith. 2014. A Bayesian Mixed Effects Model of Literary Character. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Association for Computational Linguistics, Baltimore, Maryland, 370–379. <https://doi.org/10.3115/v1/P14-1035>
- [8] Lesley Barclay, Louise Everitt, Frances Rogan, Virginia Schmied, and Aileen Wyllie. 1997. Becoming a mother—an analysis of women’s experience of early motherhood. *Journal of Advanced Nursing* 25, 4 (1997), 719–728.
- [9] CJ Berg, WM Callaghan, C Syverson, and Z Henderson. 2010. Pregnancy-related mortality in the United States, 1998 to 2005. *Obstetrics & Gynecology* 116, 6 (2010), 1302.
- [10] Cynthia J Berg, Margaret A Harper, Samuel M Atkinson, Elizabeth A Bell, Haywood L Brown, Marvin L Hage, Avick G Mitra, Kenneth J Moise, and William M Callaghan. 2005. Preventability of pregnancy-related deaths: results of a state-wide review. *Obstetrics & Gynecology* 106, 6 (2005), 1228–1234.

- [11] Sarah Blackwood. 2016. Monstrous Births: Pushing Back Against Empowerment in Childbirth. <https://thehairpin.com/monstrous-births-3d666cda5030?gi=7cbf948dfd9>. (2016).
- [12] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent Dirichlet allocation. *JMLR* 3, Jan (2003), 993–1022.
- [13] Olivier Bodenreider. 2004. The unified medical language system (UMLS): integrating biomedical terminology. *Nucleic Acids Research* 32, suppl_1 (2004), D267–D270.
- [14] Amy Bruckman. 2002. Studying the amateur artist: A perspective on disguising data collected in human subjects research on the Internet. *Ethics and Information Technology* 4, 3 (2002), 217–231.
- [15] Carma L Bylund. 2005. Mothers’ involvement in decision making during the birthing process: a quantitative analysis of women’s online birth stories. *Health Communication* 18, 1 (2005), 23–39.
- [16] Lynn Clark Callister. 2004. Making meaning: Women’s birth narratives. *Journal of Obstetric, Gynecologic, & Neonatal Nursing* 33, 4 (2004), 508–518.
- [17] Havi Hannah Carel and Ian James Kidd. 2014. Epistemic injustice in healthcare: a philosophical analysis. *Medicine, Health Care, and Philosophy* 17 4 (2014), 529–40.
- [18] Anna Carson, Cathy Chabot, Devon Greyson, Kate Shannon, Putu Duff, and Jean Shoveller. 2017. A narrative analysis of the birth stories of early-age mothers. *Sociology of Health & Illness* 39, 6 (2017), 816–831.
- [19] Nathanael Chambers and Dan Jurafsky. 2008. Unsupervised Learning of Narrative Event Chains. In *Proceedings of ACL-08: HLT*. Association for Computational Linguistics, Columbus, Ohio, 789–797. <https://www.aclweb.org/anthology/P08-1090>
- [20] Nathanael Chambers and Dan Jurafsky. 2009. Unsupervised Learning of Narrative Schemas and their Participants. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP*. Association for Computational Linguistics, Suntec, Singapore, 602–610. <https://www.aclweb.org/anthology/P09-1068>
- [21] Esther H Chen, Frances S Shofer, Anthony J Dean, Judd E Hollander, William G Baxt, Jennifer L Robey, Keara L Sease, and Angela M Mills. 2008. Gender disparity in analgesic treatment of emergency department patients with acute abdominal pain. *Academic Emergency Medicine* 15, 5 (2008), 414–418.
- [22] Billy Chiu, Gamal Crichton, Anna Korhonen, and Sampo Pyysalo. 2016. How to Train good Word Embeddings for Biomedical NLP. In *Proceedings of the 15th Workshop on Biomedical Natural Language Processing*. Association for Computational Linguistics, Berlin, Germany, 166–174. <https://doi.org/10.18653/v1/W16-2922>
- [23] Jinho D. Choi, Joel Tetreault, and Amanda Stent. 2015. It Depends: Dependency Parser Comparison Using A Web-based Evaluation Tool. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing (Volume 1: Long Papers)*. Association for Computational Linguistics, Beijing, China, 387–396. <https://doi.org/10.3115/v1/P15-1038>
- [24] Elizabeth Clark, Yangfeng Ji, and Noah A. Smith. 2018. Neural Text Generation in Stories Using Entity Representations as Context. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*. Association for Computational Linguistics, New Orleans, Louisiana, 2250–2260. <https://doi.org/10.18653/v1/N18-1204>
- [25] Lindsey L Cohen, Jean Cobb, and Sarah R Martin. 2014. Gender biases in adult ratings of pediatric pain. *Children’s Health Care* 43, 2 (2014), 87–95.
- [26] Andreea A Creanga, Cynthia J. Berg, Jean Y. Ko, Sherry L Farr, Van T Tong, Francesca Bruce, and William M. Callaghan. 2014. Maternal mortality and morbidity in the United States: where are we now? *Journal of Women’s Health* 23 1 (2014), 3–9.
- [27] Andreea A Creanga, Cynthia J Berg, Carla Syverson, Kristi Seed, F Carol Bruce, and William M Callaghan. 2015. Pregnancy-related mortality in the United States, 2006–2010. *Obstet. Gynecol.* 125, 1 (Jan. 2015), 5–12.
- [28] Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of Power: Language Effects and Power Differences in Social Interaction. In *Proceedings of the 21st International Conference on World Wide Web (WWW ’12)*. ACM, New York, NY, USA, 699–708. <https://doi.org/10.1145/2187836.2187931>
- [29] Robbie E Davis-Floyd. 2004. *Birth as an American rite of passage*. Univ of California Press.
- [30] Munmun De Choudhury, Scott Counts, and Eric Horvitz. 2013. Predicting Postpartum Changes in Emotion and Behavior via Social Media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’13)*. ACM, New York, NY, USA, 3267–3276. <https://doi.org/10.1145/2470654.2466447>
- [31] Munmun De Choudhury, Scott Counts, Eric J. Horvitz, and Aaron Hoff. 2014. Characterizing and Predicting Postpartum Depression from Shared Facebook Data. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work & #38; Social Computing (CSCW ’14)*. ACM, New York, NY, USA, 626–638. <https://doi.org/10.1145/2531602.2531675>
- [32] Janet S de Moor, Lemuel Moyé, David Low, Edgardo Rivera, S Eva Singletary, Rachel T Fouladi, and Lorenzo Cohen. 2008. Expressive writing as a presurgical stress management intervention for breast cancer patients. *Journal of the Society for Integrative Oncology* 6, 2 (2008).

- [33] Eugene R Declercq, Carol Sakala, Maureen P Corry, Sandra Applebaum, and Ariel Herrlich. 2013. Listening to mothers III: Pregnancy and birth; Report of the third national US survey of women's childbearing experiences. *New York, NY: Childbirth Connection* (2013).
- [34] Brian D Earp, Joshua T Monrad, Marianne LaFrance, John A. Bargh, Lindsey L. Cohen, and Jennifer Anne Richeson. 2019. Gender Bias in Pediatric Pain Assessment. *Journal of Pediatric Psychology* 44 4 (2019), 403–414.
- [35] Daniel A. Epstein, Nicole B. Lee, Jennifer H. Kang, Elena Agapie, Jessica Schroeder, Laura R. Pina, James Fogarty, Julie A. Kientz, and Sean Munson. 2017. Examining Menstrual Tracking to Inform the Design of Personal Informatics Tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. ACM, New York, NY, USA, 6876–6888. <https://doi.org/10.1145/3025453.3025635>
- [36] Joe Fassler. 2015. How Doctors Take Women's Pain Less Seriously. <https://www.theatlantic.com/health/archive/2015/10/emergency-room-wait-times-sexism/410515/>. (2015).
- [37] Casey Fiesler and Nicholas Proferes. 2018. "Participant" Perceptions of Twitter Research Ethics. *Social Media+ Society* 4, 1 (2018), 2056305118763366.
- [38] Michael Flor and Swapna Somasundaran. 2017. Sentiment Analysis and Lexical Cohesion for the Story Cloze Task. In *Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics*. Association for Computational Linguistics, Valencia, Spain, 62–67. <https://doi.org/10.18653/v1/W17-0909>
- [39] National Commission for the Protection of Human Subjects of Biomedical and Md Behavioral Research, Bethesda. 1978. *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Superintendent of Documents.
- [40] Adam Fourney, Ryen W. White, and Eric Horvitz. 2015. Exploring Time-Dependent Concerns About Pregnancy and Childbirth from Search Logs. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 737–746. <https://doi.org/10.1145/2702123.2702427>
- [41] Lorna Gibson and Vicki L. Hanson. 2013. Digital Motherhood: How Does Technology Help New Mothers?. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*. ACM, New York, NY, USA, 313–322. <https://doi.org/10.1145/2470654.2470700>
- [42] Katherine J. Gold, Martha E. Boggs, Emeline Mugisha, and Christie Lancaster Palladino. 2012. Internet message boards for pregnancy loss: who's on-line and why? *Women's health issues : official publication of the Jacobs Institute of Women's Health* 22 1 (2012), e67–72.
- [43] Byron J Good. 1993. *Medicine, rationality and experience: an anthropological perspective*. Cambridge University Press.
- [44] Andrew Gordon and Reid Swanson. 2009. Identifying personal stories in millions of weblog entries. In *Third International Conference on Weblogs and Social Media, Data Challenge Workshop, San Jose, CA*, Vol. 46.
- [45] Amit Goyal, Ellen Riloff, and Hal Daumé III. 2010. Automatically Producing Plot Unit Representations for Narrative Text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Cambridge, MA, 77–86. <https://www.aclweb.org/anthology/D10-1008>
- [46] Kenneth J Gruber, Susan H Cupito, and Christina F Dobson. 2013. Impact of doulas on healthy birth outcomes. *The Journal of Perinatal Education* 22, 1 (2013), 49–58.
- [47] Xinning Gui, Yu Chen, Yubo Kou, Katie Pine, and Yunan Chen. 2017. Investigating Support Seeking from Peers for Pregnancy in Online Health Communities. *Proc. ACM Hum.-Comput. Interact.* 1, CSCW, Article 50 (Dec. 2017), 19 pages. <https://doi.org/10.1145/3134685>
- [48] Robert A Hahn. 1996. *Sickness and healing: An anthropological perspective*. Yale University Press.
- [49] Diane E Hoffmann and Anita J Tarzian. 2001. The Girl Who Cried Pain: A Bias Against Women in the Treatment of Pain. *Journal of Law, Medicine and Ethics* 29 (2001), 13–27. <https://doi.org/10.1111/j.1748-720X.2001.tb00037.x>
- [50] Matthew Honnibal and Mark Johnson. 2015. An Improved Non-monotonic Transition System for Dependency Parsing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Lisbon, Portugal, 1373–1378. <https://doi.org/10.18653/v1/D15-1162>
- [51] Haley Kranstuber Horstman, Jenn Anderson, and Rebecca A Kuehl. 2017. Communicatively making sense of doulas within the US master birth narrative: Doulas as liminal characters. *Health Communication* 32, 12 (2017), 1510–1519.
- [52] Clayton J Hutto and Eric Gilbert. 2014. VADER: A parsimonious rule-based model for sentiment analysis of social media text. In *Eighth International AAI Conference on Weblogs and Social Media*.
- [53] Jeannette R Ickovics, Valerie Earnshaw, Jessica B Lewis, Trace S Kershaw, Urania Magriples, Emily Stasko, Sharon Schindler Rising, Andrea Cassells, Shayna Cunningham, Peter Bernstein, et al. 2016. Cluster randomized controlled trial of group prenatal care: perinatal outcomes among adolescents in New York City health centers. *American Journal of Public Health* 106, 2 (2016), 359–365.
- [54] Mohit Iyyer, Anupam Guha, Snigdha Chaturvedi, Jordan Boyd-Graber, and Hal Daumé III. 2016. Feuding Families and Former Friends: Unsupervised Learning for Dynamic Fictional Relationships. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, San Diego, California, 1534–1544. <https://doi.org/10.18653/v1/N16-1180>

- [55] Bram Jans, Steven Bethard, Ivan Vulić, and Marie Francine Moens. 2012. Skip N-grams and Ranking Functions for Predicting Script Events. In *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*. Association for Computational Linguistics, Avignon, France, 336–344. <https://www.aclweb.org/anthology/E12-1034>
- [56] A. Cecile J.W. Janssens and Peter Kraft. 2012. Research conducted using data obtained through online communities: ethical implications of methodological limitations. *PLoS Medicine* 9, 10 (2012), e1001328.
- [57] Matthew Jockers and Gabi Kirilloff. 2018. Understanding gender and character agency in the 19th century novel. (2018).
- [58] Lesley Kay, Soo Downe, Gill Thomson, and Kenny Finlayson. 2017. Engaging with birth stories in pregnancy: a hermeneutic phenomenological study of women’s experiences across two generations. *BMC Pregnancy and Childbirth* 17, 1 (2017), 283.
- [59] Katy Backes Kozhimannil, Rachel R Hardeman, Laura B Attanasio, Cori Blauer-Peterson, and Michelle O’Brien. 2013. Doula care, birth outcomes, and costs among Medicaid beneficiaries. *American Journal of Public Health* 103, 4 (2013), e113–e121.
- [60] Jennifer L. Kraschnewski, Cynthia H. Chuang, Erika S. Poole, Tamara Peyton, Ian Blubaugh, Jaimey Pauli, Alyssa Feher, and Madhu Reddy. 2014. Paging Dr. Google: Does technology fill the gap created by the prenatal care visit structure qualitative focus group study with pregnant women. *Journal of Medical Internet Research* 16, 6 (2014).
- [61] Briège M Lagan, Marlene Sinclair, and W George Kernohan. 2010. Internet use in pregnancy informs women’s decision making: a web-based survey. *Birth* 37, 2 (2010), 106–115.
- [62] Linda L Layne. 2003. Unhappy endings: a feminist reappraisal of the women’s health movement from the vantage of pregnancy loss. *Social Science & Medicine* 56, 9 (2003), 1881–1891.
- [63] Wendy G. Lehnert. 1981. Plot Units and Narrative Summarization. *Cognitive Science* 5 (1981), 293–331.
- [64] Stephanie Lukin, Kevin Bowden, Casey Barackman, and Marilyn Walker. 2016. PersonaBank: A Corpus of Personal Narratives and Their Story Intention Graphs. (May 2016), 1026–1033. <https://www.aclweb.org/anthology/L16-1163>
- [65] 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. *PACMHCI* 1 (2017), 73:1–73:24.
- [66] Marian F MacDorman and Eugene Declercq. 2018. Trends and state variations in out-of-hospital births in the United States, 2004–2017. *Birth* (2018).
- [67] Marian F MacDorman, Eugene Declercq, Howard Cabral, and Christine Morton. 2016. Recent Increases in the U.S. Maternal Mortality Rate: Disentangling Trends From Measurement Issues. *Obstet. Gynecol.* 128, 3 (Sept. 2016), 447–455.
- [68] Marian F MacDorman, TJ Mathews, and Eugene R Declercq. 2012. Home births in the United States, 1990–2009. (2012).
- [69] Lara J. Martin, Prithviraj Ammanabrolu, Xinyu Wang, William Hancock, Shruti Singh, Brent Harrison, and Mark O Riedl. 2018. Event representations for automated story generation with deep neural nets. In *Thirty-Second AAAI Conference on Artificial Intelligence*.
- [70] Nina Martin, Emma Cillekens, and Alessandra Freitas. 2017. Lost Mothers. *ProPublica* (2017).
- [71] Nina Martin and Renee Montagne. 2017. Nothing Protects Black Women From Dying in Pregnancy and Childbirth. *ProPublica* (2017).
- [72] Erin L Merz, Rina S Fox, and Vanessa L Malcarne. 2014. Expressive writing interventions in cancer patients: A systematic review. *Health Psychology Review* 8, 3 (2014), 339–361.
- [73] Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James Allen. 2016. A Corpus and Cloze Evaluation for Deeper Understanding of Commonsense Stories. In *Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, San Diego, California, 839–849. <https://doi.org/10.18653/v1/N16-1098>
- [74] Sarah Munro, Jude Kornelsen, and Eileen Hutton. 2009. Decision making in patient-initiated elective cesarean delivery: the influence of birth stories. *Journal of Midwifery & Women’s Health* 54, 5 (2009), 373–379.
- [75] Nikki Newhouse. 2016. bump2bump: Online Peer Support in First-Time Pregnancy. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 239–243.
- [76] Pok-Ja Oh and Soo Hyun Kim. 2016. The Effects of Expressive Writing Interventions for Patients With Cancer: A Meta-Analysis.. In *Oncology Nursing Forum*, Vol. 43.
- [77] Jessica Ouyang and Kathy McKeown. 2014. Towards Automatic Detection of Narrative Structure. In *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*. European Languages Resources Association (ELRA), Reykjavik, Iceland, 4624–4631. http://www.lrec-conf.org/proceedings/lrec2014/pdf/1154_Paper.pdf
- [78] James W Pennebaker. 1997. Writing about emotional experiences as a therapeutic process. *Psychological Science* 8, 3 (1997), 162–166.

- [79] James W Pennebaker and Sandra K Beall. 1986. Confronting a traumatic event: toward an understanding of inhibition and disease. *Journal of Abnormal Psychology* 95, 3 (1986), 274.
- [80] Karl Pichotta and Raymond J Mooney. 2016. Learning statistical scripts with LSTM recurrent neural networks. In *Thirtieth AAAI Conference on Artificial Intelligence*.
- [81] Vinodkumar Prabhakaran, Ajita John, and Dorée D. Seligmann. 2013. Who Had the Upper Hand? Ranking Participants of Interactions Based on Their Relative Power. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*. Asian Federation of Natural Language Processing, Nagoya, Japan, 365–373. <https://www.aclweb.org/anthology/I13-1042>
- [82] Vinodkumar Prabhakaran, Owen Rambow, and Mona Diab. 2012. Who's (Really) the Boss? Perception of Situational Power in Written Interactions. In *Proceedings of COLING 2012*. The COLING 2012 Organizing Committee, Mumbai, India, 2259–2274. <https://www.aclweb.org/anthology/C12-1138>
- [83] Sampo Pyysalo, Filip Ginter, Hans Moen, Tapio Salakoski, and Sophia Ananiadou. 2013. Distributional semantics resources for biomedical text processing. *Languages in Biology and Medicine* (2013), 39–44.
- [84] Elahe Rahimtoroghi, Jiaqi Wu, Ruimin Wang, Pranav Anand, and Marilyn Walker. 2017. Modelling Protagonist Goals and Desires in First-Person Narrative. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*. 360–369.
- [85] Andrew J. Reagan, Lewis Mitchell, Dilan Kiley, Christopher M. Danforth, and Peter Sheridan Dodds. 2016. The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science* 5 (2016), 1–12.
- [86] Matthew Richardson, Christopher J.C. Burges, and Erin Renshaw. 2013. MCTest: A Challenge Dataset for the Open-Domain Machine Comprehension of Text. In *Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Seattle, Washington, USA, 193–203. <https://www.aclweb.org/anthology/D13-1020>
- [87] Maarten Sap, Marcella Cindy Prasettio, Ari Holtzman, Hannah Rashkin, and Yejin Choi. 2017. Connotation Frames of Power and Agency in Modern Films. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Copenhagen, Denmark, 2329–2334. <https://doi.org/10.18653/v1/D17-1247>
- [88] Jane Staton Savage. 2001. Birth stories: A way of knowing in childbirth education. *The Journal of Perinatal Education* 10, 2 (2001), 3–7.
- [89] Roger C Schank and Robert P Abelson. 1977. Scripts, plans, goals, and understanding: An inquiry into human knowledge structures. (1977).
- [90] Shashank Srivastava, Snigdha Chaturvedi, and Tom Mitchell. 2016. Inferring Interpersonal Relations in Narrative Summaries. (2016). <https://www.aaai.org/ocs/index.php/AAAI/AAAI16/paper/view/12173>
- [91] Donna E Stewart and Simone Vigod. 2016. Postpartum Depression. *N. Engl. J. Med.* 375, 22 (Dec. 2016), 2177–2186.
- [92] Roger Suss. 2014. The hero with a thousand faces visits the doctor. *Canadian family physician Medecin de famille canadien* 60 7 (2014), 656.
- [93] Timothy R Tangherlini. 2000. Heroes and lies: Storytelling tactics among paramedics. *Folklore* 111, 1 (2000), 43–66.
- [94] Effy Vayena, Marcel Salathé, Lawrence C Madoff, and John S Brownstein. 2015. Ethical challenges of big data in public health.
- [95] Susan C Vonderheid, Rieko Kishi, Kathleen F Norr, and Carrie Klima. 2011. Group prenatal care and doula care for pregnant women. In *Reducing racial/ethnic disparities in reproductive and perinatal outcomes*. Springer, 369–399.
- [96] Su Wang, Eric Holgate, Greg Durrett, and Katrin Erk. 2018. Picking Apart Story Salads. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, Brussels, Belgium, 1455–1465. <https://doi.org/10.18653/v1/D18-1175>
- [97] Jason Weston, Antoine Bordes, Sumit Chopra, and Tomas Mikolov. 2015. Towards AI-Complete Question Answering: A Set of Prerequisite Toy Tasks. *CoRR abs/1502.05698* (2015).
- [98] Diyi Yang, Robert E. Kraut, Tenbroeck Smith, Elijah Mayfield, and Dan Jurafsky. 2019. Seekers, Providers, Welcomers, and Storytellers: Modeling Social Roles in Online Health Communities. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 344, 14 pages. <https://doi.org/10.1145/3290605.3300574>
- [99] Diyi Yang, Zheng Yao, Joseph Seering, and Robert Kraut. 2019. The Channel Matters: Self-disclosure, Reciprocity and Social Support in Online Cancer Support Groups. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM, New York, NY, USA, Article 31, 15 pages. <https://doi.org/10.1145/3290605.3300261>

Received April 2019; revised June 2019; accepted August 2019