What Six Decades of Medical Research Tell Us about Social Movements?

Evidence from Word Embeddings on Menstruation

Viviane Ito

December 2022

Abstract

This article presents an analysis of a corpus of 85,777 abstracts of articles from the PubMed platform, known for its huge database of health articles. By applying Natural Language Processing techniques, the goal is to analyze which associations are made to the term "menstruation" and how they have changed over time. Particularly, the generation of word embeddings is key to capturing the similar terms to "menstruation" and surrounding themes. In addition to the generation of word embeddings, a bootstrap resampling analysis is added to the methodology to ensure robustness. Results point to the changes in meaning over the years, and the association with themes more related to public health and social movements today.
Keywords: Natural Language Processing; Word Embeddings; Quantitative Analysis; Menstruation.

1 Introduction

Menstruation is still seen as an embarrassment by societies, despite feminist activism and constant claims for ending menstrual taboos (Gottlieb, 2020). The publication of "Our Bodies Ourselves" in the 1970s, following the feminists" urge for equality was historical for the popular knowledge of women's health (Bobel, 2008). Five decades later, research emphasizes that nowadays menstruation is regarded as "shameful" or to be "hidden" (Johnston-Robledo & Chrisler, 2020). Views of menstruation as a problem are tied to the normalization of female pain, either as a psychologizing process or as a naturalization of pain (Grace & MacBride-Stewart, 2008). Medical treatment is similar in prejudicial views; for example, endometriosis, a menstruation-related disease, can have a delay of 10 years in diagnostics owing to pain normalization (Ballard et al., 2006; Ito & Pascual, 2022).

The objective of this paper is to analyze how the medical community historically researched menstruation. Since health journals are the major source of knowledge in medical communities, our research focuses on examining abstracts from papers over a period of 6 decades (from 1960 to 2019). To collect the articles in structured data, I built a script that consults PubMed's API to gather the data, in a total of 85,777 articles. I chose to use abstracts as the central text in this paper because they are author-curated expressions of the central topics and discoveries of each article. Each abstract has an average of 200 words. I hypothesize that the 1970s and subsequent feminist demands changed the way that menstruation, a significative aspect of women's health, is referred to.

To capture the changes in the attitudes of health researchers about menstruation, I apply the theoretical framework of word embeddings, whose property is to detect word meanings and similarities. According to the work from Garg et al. (2018), the comparison of word embeddings through time can capture a change in semantics that reflects a change in behavior toward subjects. In their work, the authors examined attitudes concerning gender and race in the US, capturing meaning variation over a century, and assessing specific lexicon for measuring changes.

For defining which lexicon are the focus of my analysis, I considered the most prominent topics of "Our Bodies, Ourselves". since its conception, this book was a cultural mark and it is still a health empowerment symbol. I retrieved the main terms from the book to measure their meanings and similarities in my quantitative analysis. In addition to the word embeddings generation, I ran a bootstrap resampling analysis, to ensure the validity of the similarity metrics. Results point that the meanings of menstruation dislocated from contraception concerns (1970s) and related diseases to a broader spectrum of topics; menstruation is now addressed as a public health problem and begins to be considered not only a women's issue, including transgenders in the discussion (2010s).

The article is organized as follows: Section 2 brings a literature review on how health researchers have studied menstruation, and the application of Computational Linguistics in Social Sciences; Section 3 details the corpus retrieval, and techniques employed for text processing and analysis; Section 4 discusses the results from the word embeddings. Finally, I present the conclusions and the limitations of the study.

2 Theoretical Framework

Word embeddings are representations of words in a low-dimensional continuous space. Deriving from Vectorial Semantics as linguists like Harris (1954) defined, it determines a meaning of a word according to its neighbors on a text. Through this technique, it is possible to represent the context of words. In the last decade, (Mikolov, Le, et al., 2013; Mikolov, Sutskever, et al., 2013) and Pennington et al. (2014) developed respectively Word2Vec and GloVe, algorithms that generate word embeddings. Since then, researchers from multiple areas developed studies using word embeddings as a tool to quantify changes in word meaning. Also, studies have applied word embeddings as a theoretical framework to quantify changes in societal behavior toward issues, using historical corpora to assess such changes.

Addressing the matter of how word-meaning variation reflects changes in societal behavior, some Computational Social Science works have been developed. Among them, the work from Garg et al. (2018) studies how stereotypes associated with gender and race changed over the course of 100 years; similarly, Kozlowski et al. (2019) assess how the understanding of "class" changed in 100 years. More recently,Card et al. (2022) studied how attitudes toward immigration changed through congress speeches, also analyzing how Republicans and Democrats differ in opinions on this subject.

Regarding gender studies, NLP methods and word meaning variation are broadly applied to detect gender biases in language; Mihalcea and Garimella (2016) studied the particularities in which women and men express themselves differently, working in a discourse disambiguation project. Another line of research is on hate speech and Violence Against Women (VAW), with the development of algorithms that detect and classify texts with misogynistic vocabulary (Fortuna & Nunes, 2018; Fulper et al., 2014; García-Díaz et al., 2021).

In the medical arena, the application of NLP toward word meaning is growing. One example is the study of symptoms: according to Koleck et al. (2019), "pain" and "sleep disturbance" are the most studied symptoms by NLP methods. In PubMed, around 100 NLP works are published annually (Wang et al., 2020). However, works using PubMed's database, considering specific lexicon and word meaning variation, are still rare. One example is from Tighe et al. (2015), where the authors analyze different manners of describing pain. In this paper, I address how the medical approaches to a concept varied through decades. As far as I know, this is a novelty in studies with PubMed's corpora.

Chapter	Terms
Bodies and Identities	Body, identity
Relationships and Sexuality	Relationship, sexuality
Sexual Health and Reproductive Choices	Pregnancy
Childbearing	Pregnancy
Post reproductive Years	Perimenopause
Medical Problems and Navigating the Health Care System	Politics, Activism
Major Forces Affecting Women's Sexuality and Reproductive Health	Politics, Activism

Table 1: OBOS Selected Terms

Contributing to the works mentioned in this Section, the present study aims to apply the study of how the meanings over six decades indicate changes in society, combining the techniques from previous Computational Social Science, gender studies, and the database from PubMed.

3 Methods and Corpus

Data was collected from PubMed's proprietary API, using a Python script to organize information in data frames; the keywords used to retrieve articles were "menstruation" and its variants, "menstruations," "menses," and "menarche." After the collection, the data was preprocessed with tokenization and cleaning processes.

The preprocess resulted in 85,777 articles, with approximately 69 million tokens. For improving the precision of my embedding estimates, I randomly selected 30 bootstrapped samples for each decade and averaged them, as recommended by Antoniak and Mimno (2018). These authors observed that bootstrap is particularly desirable for small corpora. The previously cited article from Kozlowski et al. (2019) also use bootstrap to generate word embeddings. The word embeddings were trained with Word2Vec and GloVe for the sake of robustness.

Besides the techniques for capturing semantical contexts, a specific lexicon was defined based on a revised edition of "Our Bodies, Ourselves" (Norsigian et al., 2011) . This edition contains ideas that span from physiological characteristics to emotional ones; at the end, it addresses women's health and its intersections with public health — vital issues for economic development and equity in society. The chosen words reflect chapters from the book, although the focus is on those exploring social and medical issues.

Medical conditions (cancer, abnormal uterine bleeding, endometriosis), reproductive choices, and parenting were not considered in the study. Such themes

require a vast and specific lexicon, which is not the focus of this article; instead, the aim is to verify if the themes contained in ”Our Bodies, Ourselves“ (OBOS) became more addressed in health studies over the past six decades. For this purpose, the cosine similarity between the chosen words and ”menstruation“ was calculated in the data analysis process.

4 Results

4.1 Overview

In this first part of the analysis, the word embeddings from the whole corpus of 85,777 abstracts were generated by Word2Vec and GloVe. Figure 1 and Figure 2 concentrate on the results reported by term from OBOS, respectively by Word2Vec and GloVe.

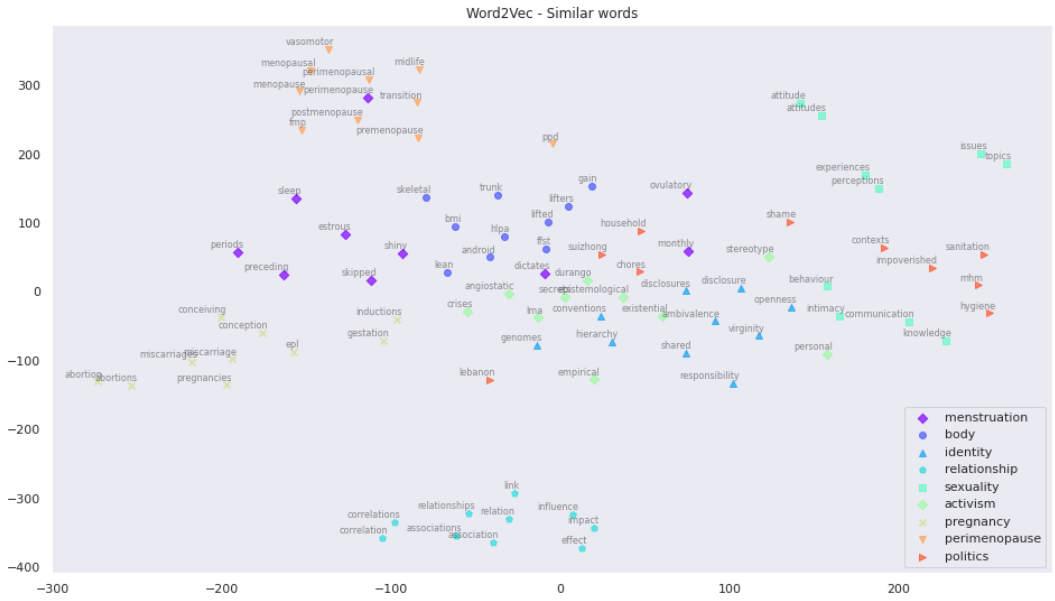


Figure 1: The 10 Most Similar Words From OBOS Terms – Word2Vec

Concerning medical dimensions, four terms are clearly located in this segment. Beginning with ”menstruation,“ the most similar terms usually have a relation with the cycle, regularity, and symptoms (such as ”monthly“, ’skipped”, and “ovulatory”). Considering “body”, “BMI” (body mass index) and “lean” indicate their association with weight. “Perimenopause” and “pregnancy” are other terms that are undeniably on the medical dimension: the representations of these terms concentrate on the left side of the chart.

In relation to social and cultural dimensions, "identity" is linked to aspects that are divergent. On one hand, there are terms such as "conventions", "openness", and "hierarchy" that seem rooted in social aspects. On the other hand, the term "genomes" indicates a medical dimension to the word: identity being in the genes. For another ambiguous term, "relationship", the word embeddings captured relations with terms such as "association" and "correlation". Social and cultural dimensions (such as friendships) were not present on the map. "Sexuality" is related to "attitudes", "experiences", but also with "communication" and "knowledge"; also, "activism" is related to knowledge in "epistemological".

The most revealing feature is "politics", which is similar to many terms that include social and cultural experiences. "Impoverished", "sanitation", "household", and "hygiene" are examples of terms related to public health; an important cultural aspect is "shame": the association of this negative feeling with menstruation is constantly regarded as the reason why women do not want to share their experiences with their periods. "Chores" locate women and menstruation at home, with domestic obligations.

Moving forward, word embeddings were trained with GloVe algorithm and plotted in a chart (Figure 2) with the same layout for comparison.

Figure 2: The 10 Most Similar Words From OBOS Terms — GloVe

Although some plotted terms on the map differ from one model to another, some contexts are similar. Differently from Word2Vec, the medical-related terms are located specifically on the top part of the chart, whereas the cultural and social dimensions are mostly at the bottom part. "Menstruation", "pregnancy", and

"menopause" are related mostly to medical related terms. One difference is that the GloVe method also captured some substances such as "atisara", which is used in ayurvedic medicine.

In "identity", "attractiveness" and "masculinity" are two aspects not in Word2Vec and that might indicate cultural dimensions of menstruation (usually associated with femininity). "Sexuality" is associated with "communication", similarly to the result of Word2Vec. Also, "politics" indicates a relation with "poverty", that has the same root as "impoverished", detected by Word2Vec.

According to both models, menstruation, pregnancy and perimenopause are deeply connected with "medical" aspects, almost opposed to public health terms. In both charts, the location of more "medicalized" terms is clearly apart from "social" terms. This idea converges with social studies, that claim that medical views of women's health are based on "medicalization". Purdy (2001) considers that the benefit of medicalization is the market possibilities that open from this phenomenon. Another view is the "pathologizing" aspect of medical studies for women's health. In this case, studies are focused only on diseases related to the theme; this happened with premenstrual syndrome, that can be classified as a psychiatric condition in some cases (Ussher et al., 2007).

Although this overview offered a broad perspective on the subject, a more in-depth analysis is required for understanding each decade. Since Word2Vec and GloVe captured similar contexts, I proceeded to the next round of analysis with Word2Vec and bootstrap analysis.

4.2 Decades

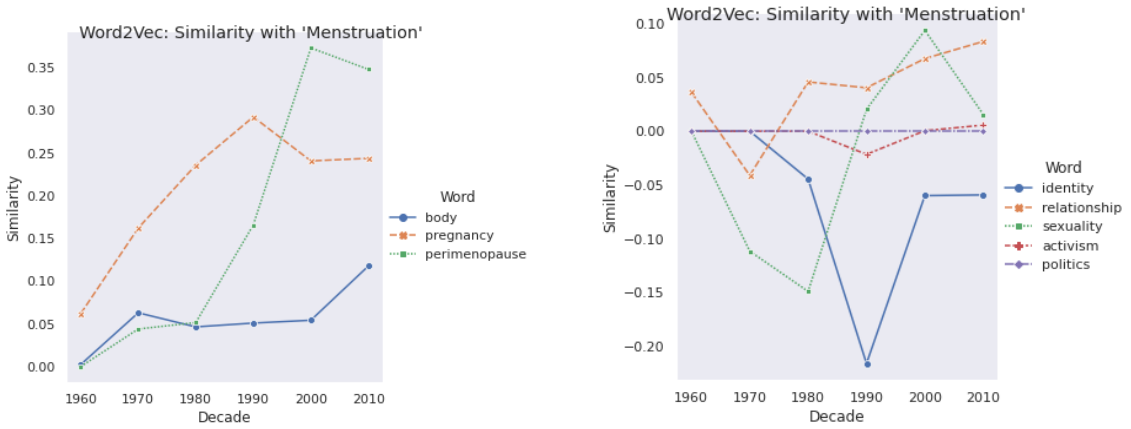
For the analysis of the decades, first, I isolated the 5 more similar terms to "menstruation" as can be verified in Table 2. Menstruation patterns, such as cycle length, incidence, and regularity, are the most observed similar terms even across decades. As in the overview analysis, menstruation seems to still be thought of a set of symptoms.

When trying another approach, I analyzed the selected terms from OBOS, splitting them in two charts: the ones that are more related to "medical" views ("body", "pregnancy", and "perimenopause") are in (a). On one hand, "pregnancy" rises from 1960s to 1990s, registering a reduction in the 2000s. On the other hand, "perimenopause" has the more expressive growth from 1990s to 2000s.

Terms related to social issues are plotted in (b). All terms have low similarity to "menstruation", except from "sexuality" that is slightly more similar in the 2000s.

1960	1970	1980	1990	2000	2010
ovulatory	cycle	latency	first	ovulatory	ovulatory
periods	woman	cycle	last	cycle	skipped
bleeding	days	woman	period	subsequent	ovulation
monthly	treatment	experienced	abortion	length	perimenopause
sleep	cycles	cycles	incidence	woman	regularity

Table 2: Five Most Similar Terms to "Menstruation"



(a) Medical Related

(b) Social Related

Figure 3: Similarity with Menstruation through Decades

A further selection of social and political terms, based on the chart from Subsection 4.1 resulted in more insighful charts. When plotting "masculinity", "hygiene", "shame", and "communication", there is a movement upwards from the last couple of decades. Therefore, the importance of testing with several specific terms.

5 Discussion

From the study, there is a clear division of understanding menstruation as "medical" or "social" issue as captured by both models of word embeddings. Interest in "social" aspects rise in the last decade; including "masculinity"; which might be the inclusion of trans men as people that menstruate. As Bobel et al. (2020) argue, there is a misconception that only healthy young women menstruate; however, other groups such as older women and transgenders also have their periods.

Movements from 1970 claimed changes in health but menstruation is still studied by "clinical" aspects if taking similar words into consideration through the decades. Nonetheless, inside the medical community, movements of broadening

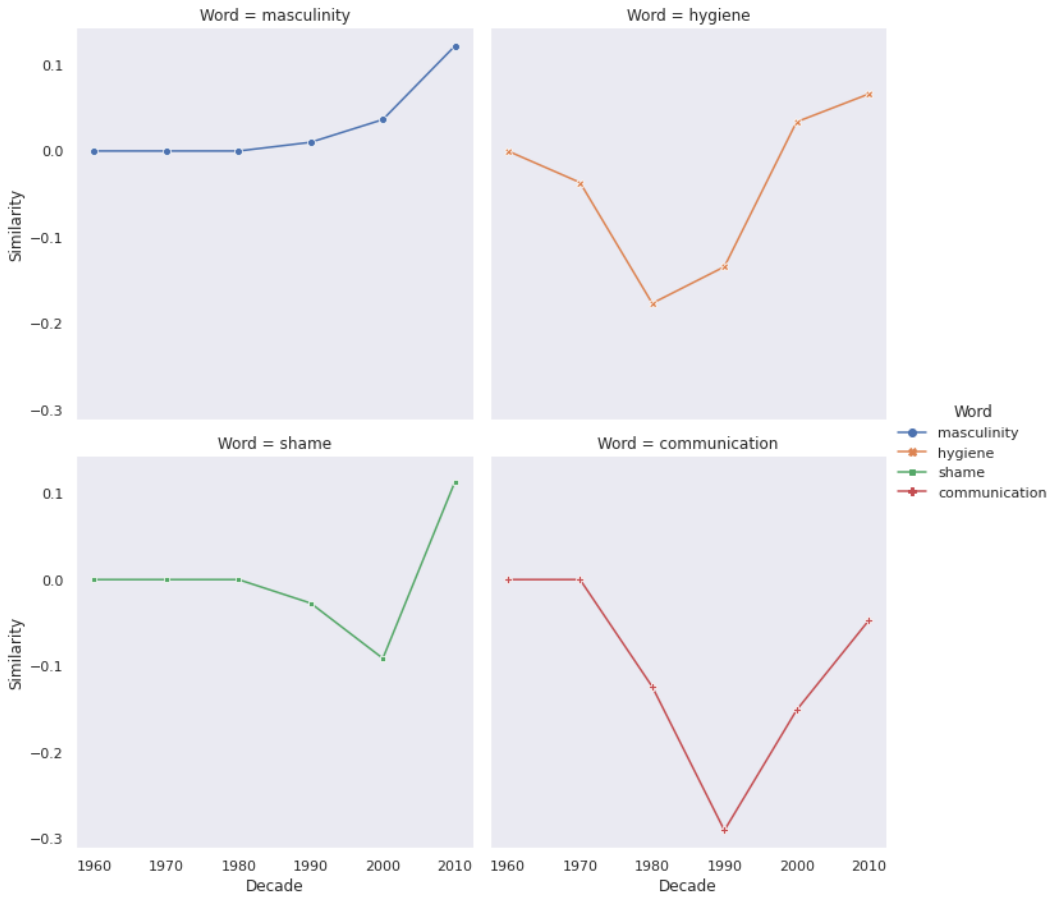


Figure 4: Similarity with Menstruation through Decades - Social and Political Related

women's studies, in a more patient-centric approach are in development. Recently, Critchley et al. (2020) reunited a team of specialists for a balance on menstruation studies in the health field. Some gaps addressed in the present study were taken into account in the authors' assessment, which might be implemented in this decade.

As limitations, I would like to emphasize that the lexicon analysis is still superficial; the human lexicon is too rich, diverse, and complex for that. Other various lexicon analysis can be performed in future studies using the same database. For example, a specific analysis of a set of diseases would be a great complement to this article. Another limitation was from hardware: I computed 30 bootstrap samples from each decade due to restrictions on my computational resources.

References

Antoniak, M., & Mimno, D. (2018). Evaluating the stability of embedding-based word similarities. *Transactions of the Association for Computational Linguistics*, *6*, 107–119. https://doi.org/10.1162/tacl_a_00008

Ballard, K., Lowton, K., & Wright, J. (2006). What's the delay? a qualitative study of women's experiences of reaching a diagnosis of endometriosis. *Fertility and sterility*, *86*(5), 1296–1301.

Bobel, C. (2008). From convenience to hazard: A short history of the emergence of the menstrual activism movement, 1971–1992. *Health care for women international*, *29*(7), 738–754.

Bobel, C., Winkler, I. T., Fahs, B., Hasson, K. A., Kissling, E. A., & Roberts, T.-A. (2020). The palgrave handbook of critical menstruation studies.

Card, D., Chang, S., Becker, C., Mendelsohn, J., Voigt, R., Boustan, L., Abramitzky, R., & Jurafsky, D. (2022). Computational analysis of 140 years of us political speeches reveals more positive but increasingly polarized framing of immigration. *Proceedings of the National Academy of Sciences*, *119*(31), e2120510119.

Critchley, H. O., Babayev, E., Bulun, S. E., Clark, S., Garcia-Grau, I., Gregersen, P. K., Kilcoyne, A., Kim, J.-Y. J., Lavender, M., Marsh, E. E., et al. (2020). Menstruation: Science and society. *American journal of obstetrics and gynecology*, *223*(5), 624–664.

Fortuna, P., & Nunes, S. (2018). A survey on automatic detection of hate speech in text. *ACM Computing Surveys*, *51*(4). https://doi.org/10.1145/3232676

Fulper, R., Ciampaglia, G. L., Ferrara, E., Ahn, Y., Flammini, A., Menczer, F., Lewis, B., & Rowe, K. (2014). Misogynistic language on twitter and sexual violence. *Proceedings of the ACM Web Science Workshop on Computational Approaches to Social Modeling (ChASM)*, 57–64.

García-Díaz, J. A., Cánovas-García, M., Colomo-Palacios, R., & Valencia-García, R. (2021). Detecting misogyny in Spanish tweets. An approach based on linguistics features and word embeddings. *Future Generation Computer Systems*, *114*, 506–518. https://doi.org/10.1016/j.future.2020.08.032

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, *115*(16), E3635–E3644.

Gottlieb, A. (2020). Menstrual taboos: Moving beyond the curse. *The Palgrave handbook of critical menstruation studies*, 143–162.

Grace, V. M., & MacBride-Stewart, S. (2008). "how to say it": Women's descriptions of pelvic pain. *Women & Health*, *46*(4), 81–98.

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2-3), 146–162.

Ito, V., & Pascual, M. (2022). Exploring chronic diseases experiences: A corpus-based approach.

Johnston-Robledo, I., & Chrisler, J. C. (2020). The menstrual mark: Menstruation as social stigma. *The Palgrave Handbook of Critical Menstruation Studies*, 181–199.

Koleck, T. A., Dreisbach, C., Bourne, P. E., & Bakken, S. (2019). Natural language processing of symptoms documented in free-text narratives of electronic health records: A systematic review. *Journal of the American Medical Informatics Association*, *26*(4), 364–379.

Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, *84*(5), 905–949.

Mihalcea, R., & Garimella, A. (2016). What men say, what women hear: Finding gender-specific meaning shades. *IEEE Intelligent Systems*, *31*(4), 62–67.

Mikolov, T., Le, Q. V., & Sutskever, I. (2013). Exploiting similarities among languages for machine translation. *arXiv preprint arXiv:1309.4168*.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *arXiv preprint arXiv:1310.4546*.

Norsigian, J., et al. (2011). *Our bodies, ourselves*. Simon; Schuster.

Pennington, J., Socher, R., & Manning, C. D. (2014). Glove: Global vectors for word representation. *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, 1532–1543.

Purdy, L. (2001). Medicalization, medical necessity, and feminist medicine. *Bioethics*, *15*(3), 248–261.

Tighe, P. J., Goldsmith, R. C., Gravenstein, M., Bernard, H. R., & Fillingim, R. B. (2015). The painful tweet: Text, sentiment, and community structure analyses of tweets pertaining to pain. *Journal of medical Internet research*, *17*(4), e3769.

Ussher, J. M., Perz, J., & Mooney-Somers, J. (2007). The experience and positioning of affect in the context of intersubjectivity: The case of premenstrual syndrome.