The polarity and diversity of premenstrual testimonies are influenced by question framing in an online experiment

Gabriella S Kountourides (gabriella.kountourides@sjc.ox.ac.uk)
University of Oxford
Alexandra Alvergne
ISEM, Univ Montpellier, CNRS, IRD

Article

Keywords: Premenstrual change, Sentiment analysis, PMS, Framing, Bayesian statistics

Posted Date: May 4th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1619419/v1

Abstract

The discourse around menstrual cycles is often pathologised, leading to negative perceptions of menstruation. The extent to which premenstrual syndrome (PMS) is inherently negative or the result of medicalised framing has long remained unclear.

We address this gap by investigating whether framing the premenstrual experience as being both negative and positive components would enable individuals to report more diverse and positive experiences than are currently reported.

In an online experiment, 2,638 participants were randomly allocated to one of three question phrasings (control: describe your premenstrual experience; treatment one: describe your positive and negative premenstrual experience; treatment two: describe your negative and positive premenstrual experience).

Sentiment analysis was applied to responses to create polarity scores. A two-part Bayesian model was performed to investigate the effect of phrasing in predicting differences in polarity scores. Across all phrasings, responses were negatively skewed.

However, response polarity scores and diversity of symptoms were higher if the question specified both positive and negative experiences compared to the control group. This corroborates previous literature showing that a negative premenstrual phase is a widespread phenomenon but challenges the idea that the premenstrual experience is solely negative.

The findings have implications for how the premenstrual experience is quantified.

Introduction

Worldwide, 48% of fertile women suffer from premenstrual syndrome (PMS), a chronic condition characterised by a constellation of over 200 physical and emotional symptoms [1]. While the experience of the premenstrual phase occurring in the two weeks preceding menstruation is usually thought of as a negative one, women also report positive experiences such as heightened creativity prior to menstruation [2, 3]. However, this perspective is rarely captured by data collection tools or conversations about the premenstrual experience. One possible reason for this omission is that the discourse around menstrual cycles is pathologising and medicalised [4, 5], leading to solely negative perceptions of menstruation. Medicalisation is a social process whereby non-medical conditions become defined and treated as medical problems [6]. Yet, the extent to which PMS is inherently negative and/or the result of the medicalised framing remains unclear. This paper aims to address this gap by investigating whether framing the premenstrual experience as being both negative and positive would allow individuals to report more diverse and positive experiences than are currently reported.

PMS has been described as an inflammatory disease, resulting from a ‘cycle of inflammation’ that underpins the menstrual cycle [7–9]. The cyclical production of reproductive hormones throughout the menstrual cycle leads to cyclical variation in inflammation, which is highest just before and during menstruation [7, 8]. It has been suggested that differences in chronic systemic inflammation could lead directly to the symptoms of PMS [7]. Indeed, inflammatory factors (such as CRP, IL-2, IL-4, IL-10, IL-12) have been shown to be positively associated with PMS symptom severity [10, 11]. As such, inflammatory states seem to have an association with PMS, and therefore negative experiences should not be conflated with disease if they are normal manifestations of the hormonal cycles [8]. These experiences are expected to vary with ecological and environmental factors, such as age and the use of hormonal contraception.

If cyclical inflammation is a cause of premenstrual syndrome, this effect is predicted to disappear for people who do not have cyclical inflammation, such as those whose ovulation is suppressed, and normal hormonal cycling is prevented, for example those on hormonal contraceptives [12]. However, this may vary by contraceptive type, for example combined contraceptive methods suppress ovulation, and are usually administered for a 21-day period, then a 7-day pill-free period, to mimic a 28-day menstrual cycle. This would lead to falling levels of progesterone, inflammation and PMS. Progestin only methods can also partially suppress ovulation, though at a lower rate than combined methods [13].

A young age has been found to be a risk factor for PMS in a study of 114 working women in Bangkok, Thailand (OR 0.61, 95% CI = 0.40–0.94) [14]. Further, the use of hormonal contraceptives should be accounted for when investigating PMS given that steroid hormones influence cyclical immunity and thus the potential for pre-menstrual related inflammation [15]. Biological states do not

exist in silos, in addition to the biological experience, an individual exists in a socio-cultural environment, which can also impact the way they experience the premenstrual phase [16].

Discourse around menstruation varies, both between countries and between people. The way in which PMS is reported and experienced is at least partly culturally mediated, for example through medical discourse and school [17, 18]. This discourse has been argued to contribute to the medicalisation of the premenstrual phase, which impacts how the premenstrual phase is experienced [19, 20]. Some have argued that PMS is a culture-bound syndrome [21]. In this line, a systematic meta-analysis of 17 papers including 18,803 research participants found that the percentage of women reporting PMS ranges from 10% in Switzerland to 98% in Iran, with an average of 48% [1]. Yet, individuals report different views of their menstrual experience, including PMS, even within the same culture [16]. On the other hand, medical texts argue that PMS is not a culture-bound syndrome, however, these texts accept that the experience varies across populations and between individuals [22]. As such, expectation and anticipation of the premenstrual phase could vary in part, by country of residence.

In the absence of a biological marker to clinically diagnose PMS, there now exists over 60 qualitative instruments to collect data about people's menstrual experiences, which predominantly ask about negative effects of the premenstrual phase [23, 24]. However, these psychometric instruments have been criticised for suggesting that women's behaviour is determined solely through hormones, ignoring women's voices and lived experiences [25]. Researchers have highlighted the need for improved tools [26–28], emphasising that excluding positive symptoms leaves an incomplete view of the menstrual cycle [29–32]. Omitting positive symptoms from diagnostic tools not only prevents a comprehensive understanding of the premenstrual phase but may even prime women to consider the premenstrual phase negatively. While studies have found that priming affects women's reports of cycle-related symptoms [33], the positive dimension that women describe the week before their periods is often excluded from research and medical questionnaires. If measures and tools only include negative symptoms, this can lead to a self-fulfilling prophecy of expecting, and then experiencing, negative symptoms.

Reports of positive and negative premenstrual changes can be influenced by the type of knowledge imparted to the person [34, 35]; however, existing studies do not compare both medicalised and positive rhetoric in the same study, and are done on relatively small numbers of participants ($n > 100$). For example, showing women a video describing a link between the menstrual cycle and low mood, increased expectations of negative mood changes, and lowered expectations of positive changes [36]. An American study on 50 women found that women could be primed to think more positively about menstruation despite their past menstrual experiences, while a negative bias in questions can lead to negative questionnaire responses, and perhaps overall negative attitudes [37]. This study found that if women were given a Menstrual Joy Questionnaire (MJQ) rather than a Menstrual Distress Questionnaire (MDQ) they reported more positive attitudes towards menstruation. They also found that the order of positive and negative survey administration had an effect: participants given the MJQ first reported more positive changes [38].

In this paper, we set out to investigate the extent to which framing the experience of the premenstrual phase as both positive and negative alters the polarity (e.g., a positive or negative sentiment), the emotional intensity of the responses (i.e., the magnitude of polarity) and the diversity of text used to describe experiences (see methods). We hypothesised that questions framed to encourage people to think about their positive premenstrual experiences would diminish how negatively they would report their experiences and/or report some positive experiences. To test this hypothesis, we used an online survey to collect text data on the premenstrual experience (PME). Participants were randomly allocated to one of three conditions where they were asked to describe their premenstrual experience using a different framing.

Some were asked about experiences in general without specifying the type of experience, and others were asked about positive and negative experiences with varying orders. We then used text analysis to investigate differences between polarity scores and word diversity across the groups. The use of sentiment analysis allows us to centre women's voices and learn directly from their experiences. As compared to focus groups and interviews, the use of computational methods allows the recruitment of thousands of experiences.

Results

Participants were randomly assigned to one of three groups, a control and two treatments. The control group was assigned a question that is consistent with questions asked by medical professionals, while treatments one and two asked specifically about

positive and negative experiences in different orders (Fig. 1). Text analysis was used to ascribe polarity scores to the response contributed by each individual.

In this sample, we find an effect of question framing on the polarity of the premenstrual testimonies. The polarity (positive, neutral, negative) of recounted experiences is higher (less negative) if the question specified both positive and negative experiences, compared to the control condition, not specifying sentiments in relation to PME. A similar effect is found for the breadth of symptoms described, which increases in treatments as compared to control. There are no differences between treatments manipulating the order of questions.

A total of 4,534 people participated in the study. After removing unfinished responses and those that did not meet the inclusion criteria (menstrual bleed in the past three months; over 18), the final sample size was 2,638 (Fig. 1). Participants were aged 18 to 54 (mean +/- sd = 28), born in 117 different countries, the majority in the UK (n = 713), America (n = 590), Australia (n = 248) and Canada (n = 113), though there were individuals from five continents. Altogether, 17% used hormonal contraception (n = 458), this figure is in line with what we expected compared to the rate of hormonal contraceptive use worldwide (21.5% [39]) (Table 1). The mean polarity of the data was − 0.19 showing that all three groups had a negative score skew (Fig. 2).

The mean polarity score for the control group was lower than in treatment groups (control: mean +/- sd= -0.25 +/- 0.42; treatment one − 0.16 +/- 0.23; treatment two − 0.15 +/- 0.23), Kruskal-Wallis χ^2=23.93, df = 2, $p < 0.001$). Pairwise comparisons revealed significant differences between control and treatment group one, and control and treatment group two ($p < 0.001$) but not between treatments one and two.

(1) Does question framing influence the polarity of premenstrual testimonies?

To investigate the impact of framing on the polarity of responses, we ran a Bayesian multinomial regression model (see methods). Participants in treatment groups one and two had over double the odds of reporting neutral rather than negative experiences as compared to those in the control group (treatment one: OR = 2.25, 95% Credible Interval [CI]:1.55–3.32; treatment two: OR = 2.36, 95% [CI]:1.62–3.47) (Fig. 3, Table 2). The odds of reporting a positive response in treatment groups one and two were not found to be statistically different when compared to the control group (treatment one: OR = 0.97, 95% [CI]:0.77–1.22; treatment two: OR = 1.05, 95% [CI]:0.83–1.32).

Next, we found no significant relationship between the use of hormonal contraceptives and reporting a neutral (OR = 0.93, 95% [CI]:0.62–1.37) or a positive experience (OR = 1.12, 95% [CI]:0.88–1.44). There were also no associations between ages and both neutral (OR 1, 95% [CI]:0.86–1.15) and positive experiences (OR = 0.94 ,95% [CI]:0.85–1.04, Table 2).

An age sensitivity analysis which included only those under the age of 40 to exclude perimenopausal people, gave similar results [SI 1]. We also ran the model grouping different contraceptives (progesterone only, combined, IUD only, no contraception, sterilised) to determine whether individual contraceptives had different results, this was not found to be the case [SI 2].

(2) Does question framing influence the magnitude of polarity (i.e., emotional intensity) in premenstrual testimonies?

Within responses scoring negative overall, treatment groups one and two were associated with less negative scores than those in the control group. Individuals in treatment one were 66% as negative as those in the control group (95% Credible Interval [CI]:0.61–0.72). Individuals in treatment two were 65% as negative as those in the control group ([CI]:0.6–0.71; Table 3; see Fig. 4 for predicted means).

Within responses scoring positive overall, treatment groups one and two were associated with less positive scores compared to the control group. Individuals in treatment one were 78% as positive of those in the control group (95% Credible Interval [CI]:0.66–0.92). Individuals in treatment two were 73% as positive as those in the control group ([CI]:0.62–0.87).

Age was found to influence negative scores across all groups (estimate: 1.06, 95% Credible Interval [CI]:1.02–1.1), with older respondents contributing more negative sentiments. There were no associations between age or the use of hormonal contraception

on the emotional intensity within positive responses.

(3) Does question framing influence word diversity?

We found that lexical density scores (number of unique words / number of total words) were the highest in treatment groups one and two as compared to the control group (Fig. 6). This suggests that the wording in treatment groups one and two elicit more diverse responses, with less repetition when compared with the control. The order of questioning “positive” and “negative” does not make a difference to the diversity of words used, whereas not including “positive” and “negative” in the prompt of the question in treatments does. The most common unigrams (individual words) found across all three groups have much overlap, with ‘breast,’ ‘cramp,’ ‘bloat,’ ‘pain,’ ‘crave.’ However, ‘cry’ and ‘sensitive’ only appear in the control’s list of 10 most frequent words, while “swing” and “acne” only appear in treatment one and two’s most frequent words (Fig. 5).

Finally, to determine whether any differences found were due the question type, rather than survey design (1 box for control, 3 for treatments), we analysed the time it took for participants to answer, and whether the three boxes led to more text. As this could have allowed more unusual responses to emerge, we also examined differences in the number of words per person. We found a statistically significant difference in mean word count by group (Kruskal-Wallis χ^2=278.28, df = 2, $p < 0.001$). Pairwise comparisons revealed with the control group having significantly more words per participant than the treatment groups, but no difference between treatment groups. We did not find any differences in time taken to complete the questionnaire (Kruskal-Wallis χ^2=5.92, df = 2, $p < 0.001$, Table 1).

Table 1
Characteristics of Treatment Groups

	Control Group	Treatment Group 1	Treatment Group 2
Background Characteristics			
Participants per group	863	918	858
Age (mean + sd)	28.1 (7)	28.2 (6.8)	28.2 (6.8)
Residence in the UK	235 (27.2%)	238 (25.9%)	240 (28%)
Hormonal Contraceptive Use	153 (17.7%)	165 (18%)	140 (16.3%)
Area of Residence			
Africa	14 (1.6%)	8 (0.9%)	13 (1.5%)
Asia and the Pacific	70 (8.1%)	63 (6.9%)	49 (5.7%)
Oceania	79 (9.2%)	108 (11.8%)	91 (10.6%)
Europe	374 (43.3%)	361 (39.3%)	360 (42%)
North America	213 (24.7%)	266 (29%)	224 (26.1%)
South America	32 (3.7%)	29 (3.2%)	29 (3.4%)
Polarity Scores			
Score (mean + sd)	-0.3 (0.4)	-0.2 (0.2)	-0.2 (0.2)
Range Score	-4.04 : 2	-0.93 : 2	-1.04 : 2
Total Number of Negative Scores	631 (73.1%)	638 (69.5%)	587 (68.4%)
Total Number of Positive Scores	191 (22.1%)	187 (20.4%)	183 (21.3%)
Total Number of Neutral Scores	40 (4.6%)	93 (10.1%)	88 (10.3%)
Response Characteristics			
Number of Words per response (mean + sd)	31.8 (39.3)	14 (14.5)	15.4 (16.5)
Time taken to complete (mean + sd)	4770.1 (121050.5)	1038.7 (14613.6)	511.9 (1525.8)

Table 2

Output of a multinomial regression model. The response variable is the polarity of responses with 3 categories (positive, neutral and negative). 'Negative' was set as reference category. Bold estimates indicate significant effects.

	Odds Ratio	Est.Error	l-95% CI	u-95% CI
Intercept [Neutral]	**0.07**	**1.32**	**0.04**	**0.13**
Intercept [Positive]	**0.30**	**1.17**	**0.23**	**0.41**
Treatment Group 1: Neutral	**2.25**	**1.21**	**1.55**	**3.32**
Treatment Group 2: Neutral	**2.36**	**1.22**	**1.62**	**3.47**
Standardised Age: Neutral	1.00	1.08	0.86	1.15
Hormonal Contraception (True): Neutral	0.93	1.22	0.62	1.37
Treatment Group 1: Positive	0.97	1.12	0.77	1.22
Treatment Group 2: Positive	1.05	1.13	0.83	1.32
Standardised Age: Positive	0.94	1.05	0.85	1.04
Hormonal Contraception (True): Positive	1.12	1.13	0.88	1.44

Table 3

Model output for the range of negative and positive polarity scores across experimental groups. The estimates are the transformed (exponentiated) estimate for the model. The estimate represents the multiplicative change in polarity score of a treatment compared to control. An individual in treatment one has a multiplicative change in strength of negative score. (a) Negative Model; (b) Positive Model. The control group is the baseline. Bold estimates indicate significant effects.

	Estimate	Est.Error	l-95% CI	u-95% CI
(a) Model on Negative Responses				
Intercept	**0.41**	**1.03**	**0.38**	**0.43**
Treatment Group 1	**0.66**	**1.05**	**0.61**	**0.72**
Treatment Group 2	**0.65**	**1.05**	**0.60**	**0.71**
Standardised Age	**1.06**	**1.02**	**1.02**	**1.10**
Use Hormonal Contraception	0.99	1.05	0.90	1.09
(b) Model on Positive Responses				
Intercept	**0.19**	**1.06**	**0.17**	**0.22**
Treatment Group 1	**0.78**	**1.09**	**0.66**	**0.92**
Treatment Group 2	**0.73**	**1.09**	**0.62**	**0.87**
Standardised Age	1.00	1.04	0.94	1.08
Use Hormonal Contraception	0.90	1.10	0.75	1.08

Discussion

We investigated whether the premenstrual experience (PME) is influenced by question framing using an online experiment. We manipulated the phrasing of the question asking for a description of the premenstrual experience (control: nothing specified;

treatments: framing the question specifically asking about both negative and positive experiences in a different order). Drawing on the analysis of text data collected, the results show that (1) in both control and treatment groups, PME was overall described negatively, (2) framing the question with emotional valence specifying both positive and negative experiences, resulted in higher odds of reporting neutral compared to negative responses, and (3) the experience was described less negatively, and the breadth of symptoms described was higher (more diverse) in treatments compared to the control group. These results have implications for how the premenstrual experience is measured and challenges the idea that the premenstrual experience is solely negative.

The results show that, regardless of the phrasing of the question, the polarity of menstruating individual's accounts of the premenstrual experience are overall negative, but not solely so. Our results are consistent with previous studies showing that menstruating individuals describe their premenstrual phase as a negative period [40, 41]. This suggests that inflammation associated with the menstrual cycle has a large impact on descriptions, leading menstruating individuals to experience the phase in a negative way. However, language and social constructions also play a role in the description of the experience. Negative premenstrual experiences may result from an inflammatory reaction triggered by dropping progesterone levels before menstruation.

In this line, premenstrual syndrome has been regarded as the outcome of a cycle of inflammation [42–44]. Yet, the negative premenstrual experience reported by individuals using hormonal contraception cannot be explained by cyclical inflammation, so other mechanisms might be at play.

The analysis shows that age and the use of hormonal contraception did not associate with the polarity of responses. This is surprising because both age and use of hormonal contraceptives are important in mediating cycle differences [45, 46]. The lack of effect on combined contraceptive users could stem from the fact that these types of contraceptives mimic the menstrual cycle and cyclical immunity, and it has also been suggested that these may produce PMS-like symptoms such as irritability, which are confused with PMS [47]. The seven-day placebo pill causes a drop of hormones leading to a withdrawal bleed, which may have been obscured within the general 'hormonal contraceptive' user responses. We had expected progestin only methods to show lower levels of any premenstrual changes, as they do not cause cycling, however, they do impact inflammation and responses to infection. For example, it is interesting to note that in rodents, infection has been facilitated by the administration of progestin contraception [48], so progestin might heighten systemic inflammation. Therefore, those on progestin contraceptives have higher levels of systemic inflammation, which may mask any impact of cyclical inflammation. We also expected differences between IUD users and hormonal contraceptive users. However, while IUD users' ovulation is not affected, any intra-uterine device causes uterine irritation and inflammation [49], which may contribute to systemic inflammation similarly to progestin only methods. While we were surprised by these results, one previous study showed minor differences in PMS severity between those who used and did not use hormonal contraceptives, and encouraged researchers to include hormonal contraceptive users from future studies on the premenstrual phase [50]. It is important to note that we did not have data on how long people had been using their contraceptive methods, and whether they took them regularly. It is also possible that people on hormonal contraceptives were responding with their premenstrual experiences over their lifetime, or for cycles before they took contraception. Or, perhaps people on hormonal contraceptives are conflating negative side effects from contraceptive use with PMS.

The lack of differences between hormonal contraceptive and non contraceptive users could emphasise the importance of the social framing of premenstrual syndrome. It is possible that the negative stereotype of PMS is so pervasive, that people undergoing a withdrawal bleed (from hormonal contraception) rather than a true bleed also believe they are experiencing PMS.

Most respondents did not describe positive experiences, and some even reacted strongly against the idea of having these. In the advertising industry, framing has been shown to affect responses and behaviour [51], however, we did not find this to be the case in all parts of this study. We were surprised to find that the number of neutral responses was impacted by treatment, even though the number of positive and negative responses were not. This suggests that: i) whilst the premenstrual experience is generally a negative one; ii) when menstruating individuals are prompted specifically to consider positive and negative effects, they are more likely to respond with a neutral statement such as "I don't get PMS" or "none". Some research has shown that as the number of Likert-style options increase, the extreme responses decreased [52], suggesting that people opt for a neutral response to avoid the cognitive effort required to give a polarised response[53]. This could explain the reduced spread of positive and negative responses in the treatment groups, given we don't see the same results when there are free text boxes.

Question phrasing influences the number and type of words used to describe the PME. We found that the diversity of language used in treatment groups one and two is higher than in the control group, which had fewer unique words than treatment groups one and two. This suggests more repetition in the responses within the control group. This indicates that the diversity of language used in treatment groups one and two is higher than that in the control group, suggesting that when participants are directly asked to describe different experiences, they have more to say. However, the number of words used does not differ between the groups, rather, they describe different experiences. While some words were recurrent in all three groups, in particular words depicting physical pain (e.g., 'breast,' 'cramp' and 'bloat'), some were unique to the control group such as 'cry' and 'sensitive'. 'Swing' and 'acne' were found to be among the most frequent words used in treatment groups one and two. This suggests that the control elicits negative emotional and psychological symptoms, while the treatments elicited more descriptive bodily symptoms. Note that the number of boxes is unlikely to explain differences in word density: the control group with the single text box had almost double the word count compared to the others. Control responses had longer sentences with more descriptive words, while responses in treatment groups read more often as a list of symptoms. The responses in the treatments being split into 'positive', 'negative' and 'other' may have resulted in a 'list style' response of specific symptoms, while responses in the control group were more verbose, resulting in more general words with fewer distinct words.

Contrary to other studies, we have not found that the order of the wording impacted responses [38, 54]. This could be because participants could see all the words in the same question, regardless of the question order, while other studies administering different orders of positive and negative surveys were done by entirely separate questionnaires [37].

Using sentiment analysis permits the analysis of a large number of responses, many more than would have been feasible if these were being individually analysed by a thematic analysis method [55–57]. Indeed, the large number of respondents led to the collection of the largest and most diverse PMS experiences to date. However, our methodology was limited by the dictionary used in text analysis, which was an adapted dictionary lookup. A tailor-made sentiment dictionary would have been more accurate. However, we are not aware of any annotated training data created for female or reproductive health. Creating a more advanced model for this paper, using deep neural networks was beyond the scope of this study, indeed very few studies have used deep learning to address issues of cross-domain sentiment classification [58]. However, our study has highlighted the need for a new lexicon that encompasses words relevant to female reproductive health.

Another limitation was that we did not have any data on BMI, as obesity has been shown to be a risk factor for PMS [59]. However, due to the large number of respondents which were randomly allocated into the three response groups, any effect due to BMI would have been the same in each group. We opted not to record BMI as sensitive questions can lead to higher survey drop-out rates [60].

While biomedical perspectives may provide mechanistic explanations of PMS, there is no single experience of the premenstrual phase. Using an inductive, feminist approach that places participants experiences and their unique words at the centre of the research could give rise to hypotheses for why such variation in PMS is observed, for example highlighting the presence of neutral responses. It is clear that the premenstrual phase is negative for a majority of people, but this work highlights the existence of neutral and positive dimensions that should also be considered by healthcare providers. The current medical literature emphasises negative cycle changes, while positive ones get very little attention. Discussing the premenstrual phase as "changes" instead of "syndrome" could help break down the concept of PMS, to reduce the negative connotations that it invokes [61, 62].

Moreover, the wording of the question about the premenstrual phase can allow a greater diversity of responses to emerge from participants. Calls for future research should focus on the development of a reproductive health lexicon, with a focus on menstrual cycles. By using this, and ensuring enough models are trained in different languages, we will be able to quickly capture and analyse huge numbers of responses. This in turn will enable researchers to assess different risk factors for a negative premenstrual phase (we suggest inflammatory illnesses and stress levels) while also determining 'risk factors' for positive and neutral premenstrual phases. In this way, our findings have implications for how the premenstrual experience is quantified. Understanding which words are the most frequently used can help researchers build specific tools to capture the premenstrual experience in future, and to establish ways in to improve the premenstrual experiences, for example targeting anti-inflammatory medications at specific cycle points for those with negative premenstrual changes.

Methods

a) Recruitment of participants

The study received ethical clearance from the School of Anthropology and Museum Ethnography Research Ethics Committee (Ref No.: SAME_C1A_19_07). Participants were recruited through tweets and push notifications released by Clue, a free period tracker app based in Germany and released in 2013 by BioWink GmbH. Three tweets were sent out in March 2020 on Twitter by Clue, which were retweeted by both authors. After one month, there had been 200 survey responses. Clue subsequently sent notifications to 30,000 randomly selected app users with their app language set to English, thus were not restricted by geographic location, but rather by user language. Participants followed the link in the tweet/push notification which directed them towards an online survey hosted by the Qualtrics survey platform [63]. In total, 2,751 participants completed the survey, and 1783 dropped out during the survey, leaving a completion rate of 61%. The response rate for push notifications through the Clue app was 8.6% but the overall response rate is likely to be lower, as the response rate for Twitter could not be calculated, since we could not determine how many people saw the tweets. Participants could only complete the survey if they were over 18, had a menstrual bleed in the past three months, and gave informed consent to the use of their data.

b) Survey

The survey included a maximum of 21 questions depending on individual circumstances. The survey collected quantitative data on age, level of partner support, medication, country of birth, vitamin use, and contraceptive use, as well as qualitative data on the premenstrual experience through a free text entry question (SI 3). These data were collected as they have been shown to affect the symptoms and intensity of symptoms that menstruating individuals describe [64–71]. Data collection took place from the 13th March 2020 through to 1st July 2020 and was halted once no more push notifications were scheduled, and no new responses had been recorded for three days.

c) Experimental design

To test for the effect of framing on the polarity of responses and diversity of words used in responses, participants were randomly assigned to one of three groups, a control and two treatments. Each group was shown a different wording for the question asking to describe one's premenstrual experience. The control group was assigned a question that is consistent with questions asked by medical professionals: "*Please tell us about any changes (e.g., physical, social, psychological, emotional, behavioural) you experience the week before your period*", which was answered in a single, generic text box entry. Treatment one was assigned the question "*Please tell us about any negative and/or positive changes (e.g., physical, social, psychological, emotional, behavioural) you experience the week before your period*" which was answered in three boxes: 'negative', 'positive' and 'other'. Treatment two was assigned the question "*Please tell us about any positive and/or negative changes (e.g., physical, social, psychological, emotional, behavioural) you experience the week before your period*", which was answered in three boxes, 'positive', 'negative', and 'other'. Thus treatments one and two were assigned a question that would specify the type of change (negative or positive) experienced, but in a different sequence (negative first for treatment one, and positive first for treatment two), differing only in the order of symptoms they asked for. At no point in the survey did the words 'PMS' or 'pre-menstrual syndrome' appear, to avoid biasing responses.

d) Data Management

Responses were downloaded from the survey platform 'Qualtrics' onto a password-protected computer. The '*qualtRics'* package [72] was used to move survey results into R using the 'Qualtrics' API. To maintain security, an encrypted *.Renviron* file was used to store environment variables. Any personal information collected by the platform, such as IP addresses were removed, and each person was assigned a unique ID.

e) Text Analysis

Data Cleaning: Text processing was performed, as texts written in questionnaires are often unedited, typos and abbreviations often occur [73]. Similarly, the same words can be represented differently, such as 'PMS', 'premenstrual syndrome' and 'pre-menstrual syndrome'. To address this, a cleaning function was created (the script is open access and can be found on GitHub (https://bit.ly/3saG0ie). The function created was adapted from code written in the '*textclean*' package [74].

i) Extracting the polarity of responses

Polarity scores: Polarity scores were calculated to answer the following two questions: (i) do the number of positive, negative and neutral responses vary by group? (ii) does the range of scores vary by group? To create a polarity score for each response, we performed a sentiment analysis, a lexicon-based approach that uses text mining and computational linguistics using the *sentimentr* package [75]. By comparing words in the text to a pre-defined dictionary of words (lexicon), an algorithm assigns positive, negative and neutral tags to sentences. Negators, shifters and intensifiers (words such as "not" and "really") are accounted for in the final polarity score [76]. A challenge in sentiment analysis is the lack of datasets that have been used to train a model which can adapt to specific domains [77]. We are not aware of any annotated datasets about menstruation that could be used to train a model to accurately classify responses as positive or negative. As such, we adapted the default lexicon in the *sentimentr* package [75] by creating a custom dictionary, whereby the polarity of relevant words was edited (SI 4).

Word Diversity: The text was analysed to answer the following three questions (i) which single words (unigrams) are the most frequent and important in each group? (ii) what are the similarities in words between control and treatments? (iii) Are unique words found in any of the groups?

First, further data processing was performed by tokenising the text to break sentences into a one-word-per-row format. Highly common words that do not carry meaning such as "and", "thus" (stopwords) were removed (SI 5), and applied to the text, removing 22,997 words, leaving 32,111. Words were then lemmatised, when various forms of a word are grouped together so that they can be analysed as a single word (e.g., hunger, hungry), using the '*textstem*' package [78]. Specific steps can be seen in (SI 6).

Second, word frequency was used to determine which singular words were used most frequently across the three groups. For each group, the lexical density corresponding to the ratio of unique words relative to the total number of words in each group was calculated, where higher scores mean more unique words and less repetition. All text was considered together, irrespective of which box participants wrote it.

f) Statistical Analysis

Missing Data Imputation: Descriptive statistics were calculated to show distributions for each variable (group, age, use of hormonal contraceptives). There were 118 missing observations for age and 255 for residence. To avoid power reduction and bias by excluding incomplete cases [79], multiple imputation was performed using chained random forest in the missRanger [80] package. This uses all other variables as co-variables and a random forest plot to predict each variable, iterating it until the average out-of-bag prediction error of the models no longer improves. The final model for each question was run five times, each on one of the imputed datasets, estimates pooled. The models were run on both complete cases and imputed datasets to check that results were similar [SI 7].

Analysis: To investigate if the experimental treatment influences the polarity of responses, we ran a two-part model [81] to account for the occurrence of zeros (neutral responses) in a categorical model (model 1), and the non-zero polarity scores (positive and negative) using a continuous distribution (model 2 + 3).

In model 1, we categorized polarity as positive (> 0), negative (0<) and neutral (0) and used a multinomial logistic regression with minimally informative priors, with the 'negative' polarity set as reference. In model 2 we focused on the range of negative polarity scores, which were inversed for the purpose of using a continuous probability distribution. In model 3 we focused on the range of positive polarity scores. In all models age was standardised, and in models 2 and 3, polarity was modelled using a gamma distribution. We compared the effect of control vs. treatments (investigating any effect of framing the premenstrual experience as positive and negative), as well as treatment one vs. treatment two (investigating any effect of the order of terms with opposed emotional valence).

Models were fitted using Stan computational framework [82] made accessible by the *brms* v2.21.2 package which fits Bayesian models using the Hamiltonian Monte Carlo (HMC) algorithm [83]. Estimate uncertainty was summarised using the 95% credible intervals (CI). Different models were compared using the leave-one-out cross-validation (LOO-CV) method. All models were checked through graphical posterior predictive checking, where graphical displays of posterior predictive distribution data were simulated and compared to our data [84].

Model Comparison The full model was compared to a null model to investigate the relative fit [SI 8]. An approximate leave-one-out cross-validation (LOO-CV) was run [85]. The difference in the two model's expected predictive accuracy was examined using the ELPD-LOO (expected log pointwise predictive density leave-one-out cross-validation), and the better fitting model was selected [SI 8]. Post predictive checks can be seen in [SI 9].

Declarations

Acknowledgements

The authors thank Alex Mielke for helpful advice on the statistical analysis, Austin Argentieri for advice on multiple imputation, and Aki Vehtari for comments on [SI 7+8].

Author Contributions

GK and AA Conceived and designed the analysis with input from Clue. GK Collected the data and conducted data cleaning and analysis. GK wrote the initial draft and prepared the figures and tables. GK and AA finalised the manuscript.

Data Availability Statement

The datasets generated during and/or analysed during the current study are available in the [Github] repository, [https://bit.ly/3saG0ie].

Additional Information

The author(s) declare no competing interests.

References

1. Direkvand-Moghadam A, Sayehmiri K, Delpisheh A, Kaikhavandi S. Epidemiology of premenstrual syndrome (PMS)-a systematic review and meta-analysis study. Journal of Clinical and Diagnostic Research: JCDR. 2014;8: 106–109. doi:10.7860/JCDR/2014/8024.4021
2. Campagne DM, Campagne G. The premenstrual syndrome revisited. European Journal of Obstetrics Gynecology and Reproductive Biology. 2007;130: 4–17. doi:10.1016/j.ejogrb.2006.06.020
3. King M, Ussher JM. It's not all bad: Women's construction and lived experience of positive premenstrual change. Feminism & Psychology. 2013;23: 399–417. doi:10.1177/0959353512440351
4. Ussher JM. Premenstrual syndrome: Reconciling disciplinary divides through the adoption of a material-discursive epistemological standpoint. Annual Review of Sex Research. 1996;7: 218–251. doi:10.1080/10532528.1996.10559914
5. Ussher JM, Perz J. PMS as a process of negotiation: Women's experience and management of premenstrual distress. Psychology & Health. 2013;28: 909–927. doi:10.1080/08870446.2013.765004
6. Rose N. Medicine, History and the Present. In: Jones C, Porter R, editors. London: Routledge; 1994.
7. Bertone-Johnson ER. Chronic Inflammation and Premenstrual Syndrome: A Missing Link Found? Journal of Women's Health. 2016;25: 857–858. doi:10.1089/jwh.2016.5937
8. Alvergne A, Högqvist Tabor V. Is female health cyclical? Evolutionary perspectives on menstruation. Trends in Ecology and Evolution. 2018;33: 399–414. doi:10.1016/j.tree.2018.03.006
9. Mattina GF, Van Lieshout RJ, Steiner M. Inflammation, depression and cardiovascular disease in women: The role of the immune system across critical reproductive events. Therapeutic Advances in Cardiovascular Disease. 2019;13: 1753944719851950. doi:10.1177/1753944719851950
10. Puder JJ, Blum CA, Mueller B, De Geyter C, Dye L, Keller U. Menstrual cycle symptoms are associated with changes in low-grade inflammation. European Journal of Clinical Investigation. 2006;36: 58–64. doi:10.1111/j.1365-2362.2006.01591.x
11. Bertone-Johnson ER, Ronnenberg AG, Houghton SC, Nobles C, Zagarins SE, Takashima-Uebelhoer BB, et al. Association of inflammation markers with menstrual symptom severity and premenstrual syndrome in young women. Human Reproduction.

2014;29: 1987–1994. doi:10.1093/humrep/deu170

12. Jasienska G, Bribiescas RG, Furberg AS, Helle S, Núñez-de la Mora A. Human reproduction and health: An evolutionary perspective. The Lancet. 2017;390: 510–520. doi:10.1016/S0140-6736(17)30573-1
13. Rivera R, Yacobson I, Grimes D. The mechanism of action of hormonal contraceptives and intrauterine contraceptive devices. American Journal of Obstetrics and Gynecology. 1999;181: 1263–1269. doi:10.1016/S0002-9378(99)70120-1
14. Thipsiri Prungsin, Surasak Taneepanichskul. Prevalence and quality of life (QOL) with premenstrual syndrome (PMS) among the working women in reproductive age group in Bangkok, thailand. Journal of Health Research. 2016;30: S139–S145. doi:10.14456/JHR.2016.77
15. Yonkers KA, O'Brien PS, Eriksson E. Premenstrual syndrome. The Lancet. 2008;5: 1200–10. doi:10.1016/S0140-6736(08)60527-9
16. Chandra PS, Chandra PS. Cultural Variations of Premenstrual Experience. Int J Soc Psychiatry. 1989;35: 343–349. doi:10.1177/002076408903500407
17. Walker AF, De Souza MC, Vickers MF, Abeyasekera S, Collins ML, Trinca LA. Magnesium Supplementation Alleviates Premenstrual Symptoms of Fluid Retention. Journal of Women's Health. 1998;7: 1157–1165. doi:10.1089/jwh.1998.7.1157
18. Heard KV, Chrisler JC, Kimes LA, Siegel HN. Psychometric evaluation of the menstrual joy questionnaire. Psychological Reports. 1999;84: 135–136. doi:10.2466/pr0.1999.84.1.135
19. Lopez LM, Kaptein AA, Helmerhorst FM. Oral contraceptives containing drospirenone for premenstrual syndrome. Cochrane Database of Systematic Reviews. 2012. doi:10.1002/14651858.CD006586.pub4
20. Green LJ, O'Brien PMS, Panay N, Craig N. Management of premenstrual syndrome. BJOG. 2017;48: e73–e105. doi:10.1111/1471-0528.14260
21. Johnson SR. The epidemiology and social impact of premenstrual symptoms. Clinical Obstetrics and Gynecology. 1987;30: 367. Available: https://journals.lww.com/clinicalobgyn/Citation/1987/06000/The_Epidemiology_and_Social_Impact_of_Premenstrual.17.aspx
22. Green LJ, O'Brien PMS, Panay N, Craig N. Management of premenstrual syndrome. BJOG. 2017;48: e73–e105. doi:10.1111/1471-0528.14260
23. Knaapen L, Weisz G. The biomedical standardization of premenstrual syndrome. Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences. 2008;39: 120–134. doi:10.1016/j.shpsc.2007.12.009
24. Romans S, Clarkson R, Einstein G, Petrovic M, Stewart D. Mood and the Menstrual Cycle: A Review of Prospective Data Studies. Gender Medicine. 2012;9: 361–384. doi:10.1016/j.genm.2012.07.003
25. Taylor D. From "It's All in Your Head" to "Taking Back the Month": Premenstrual Syndrome (PMS) Research and the Contributions of the Society for Menstrual Cycle Research. Sex Roles. 2006;54: 377. doi:10.1007/s11199-006-9009-z
26. Steiner M, Haskett RF, Carroll BJ. Premenstrual tension syndrome: The development of research diagnostic criteria and new rating scales. Acta Psychiatrica Scandinavica. 1980;62: 177–190. doi:10.1111/j.1600-0447.1980.tb00605.x
27. Steiner M, Streiner DL, Steinberg S, Stewart D, Carter D, Berger C, et al. The measurement of premenstrual mood symptoms. Journal of Affective Disorders. 1999;53: 269–273. doi:10.1016/S0165-0327(98)00121-9
28. Haywood A, Slade P, King H. Assessing the assessment measures for menstrual cycle symptoms: A guide for researchers and clinicians. Journal of Psychosomatic Research. 2002;52: 223–237. doi:10.1016/S0022-3999(02)00297-0
29. Choi PYL, McKeown S. What are young undergraduate women's qualitative experiences of the menstrual cycle? Journal of Psychosomatic Obstetrics & Gynecology. 1997;18: 259–265. doi:10.3109/01674829709080697
30. Logue CM, Moos RH. Positive perimenstrual changes: Toward a new perspective on the menstrual cycle. Journal of Psychosomatic Research. 1988;32: 31–40. doi:10.1016/0022-3999(88)90086-4
31. Stewart DE. Positive changes in the premenstrual period. Acta Psychiatrica Scandinavica. 1989;79: 400–405. doi:https://doi.org/10.1111/j.1600-0447.1989.tb10276.x
32. Chrisler JC, Johnston IK, Champagne NM, Preston KE. Mentsrual Joy The Construct and Its Consequences. Psychology of Women Quarterly. 1994;18: 375–387. doi:https://doi.org/10.1111/j.1471-6402.1994.tb00461.x

33. Rose JG, Chrisler JC, Couture S. Young Women's Attitudes Toward Continuous Use of Oral Contraceptives: The Effect of Priming Positive Attitudes Toward Menstruation on Women's Willingness to Suppress Menstruation. Health Care for Women International. 2008;29: 688–701. doi:10.1080/07399330802188925
34. Marván ML, Escobedo C. Premenstrual symptomatology: Role of prior knowledge about premenstrual syndrome. Psychosomatic medicine. 1999;61: 163167.
35. Aubeeluck A, Maguire M. The menstrual joy questionnaire items alone can positively prime reporting of menstrual attitudes and symptoms. Psychology of Women Quarterly. 2002;26: 160162.
36. Olasov B, Jackson J. Effects of expectancies on women's reports of moods during the menstrual cycle. Psychosomatic Medicine. 1987.
37. Chrisler JC, Johnston IK, Champagne NM, Preston KE. Menstrual Joy: The Construct and Its Consequences. Psychology of Women Quarterly. 1994;18: 375–387. doi:10.1111/j.1471-6402.1994.tb00461.x
38. Kues JN, Janda C, Krzikalla C, Andersson G, Weise C. The effect of manipulated information about premenstrual changes on the report of positive and negative premenstrual changes. Women & Health. 2018;58: 16–37. doi:10.1080/03630242.2016.1263274
39. United Nations. Contraceptive Use by Method 2019: Department of Economic; Social Affairs, Population Division; 2019 Dec. doi:10.18356/1bd58a10-en
40. Stubbs ML, Costos D. Negative Attitudes Toward Menstruation: Implications for Disconnection Within Girls and Between Women. Women & Therapy. 2004;27: 37–54. doi:10.1300/J015v27n03_04
41. Ussher JM, Perz J. "I feel fat and ugly and hate myself": Self-objectification through negative constructions of premenstrual embodiment. Feminism & Psychology. 2020;30: 185–205. doi:10.1177/0959353519900196
42. Bertone-Johnson ER, Ronnenberg AG, Houghton SC, Nobles C, Zagarins SE, Takashima-Uebelhoer BB, et al. Association of inflammation markers with menstrual symptom severity and premenstrual syndrome in young women. Human Reproduction. 2014;29: 1987–1994. doi:10.1093/humrep/deu170
43. Gold EB, Wells C, Rasor MO. The Association of Inflammation with Premenstrual Symptoms. Journal of Women's Health. 2016;25: 865–874. doi:10.1089/jwh.2015.5529
44. Alvergne A, Vlajic Wheeler M, Högqvist Tabor V. Do sexually transmitted infections exacerbate negative premenstrual symptoms? Insights from digital health. Evolution, Medicine, and Public Health. 2018;2018: 138–150. doi:10.1093/emph/eoy018
45. Matsumoto S. Statistical studies on menstruation; a criticism on the definition of normal menstruation. Gunma Journal of Medical Sciences. 1962;11: 294–318. Available: https://ci.nii.ac.jp/naid/10020467343/
46. Vitzthum VJ. The ecology and evolutionary endocrinology of reproduction in the human female. American Journal of Physical Anthropology. 2009;140: 95–136. doi:10.1002/ajpa.21195
47. Rapkin A. A review of treatment of premenstrual syndrome & premenstrual dysphoric disorder. Psychoneuroendocrinology. 2003;28: 39–53. doi:10.1016/S0306-4530(03)00096-9
48. Kaushic C, Ashkar AA, Reid LA, Rosenthal KL. Progesterone increases susceptibility and decreases immune responses to genital herpes infection. Journal of Virology. 2003;77: 4558–4565. doi:10.1128/JVI.77.8.4558-4565.2003
49. Ortiz ME, Croxatto HB. Copper-T intrauterine device and levonorgestrel intrauterine system: biological bases of their mechanism of action. Contraception. 2007;75: S16–S30. doi:10.1016/j.contraception.2007.01.020
50. Yonkers KA, Cameron B, Gueorguieva R, Altemus M, Kornstein SG. The influence of cyclic hormonal contraception on expression of premenstrual syndrome. Journal of Women's Health. 2017;26: 321–328. doi:10.1089/jwh.2016.5941
51. Chang C. Ad framing effects for consumption products: An affect priming process. Psychology & Marketing. 2008;25: 24–46. doi:10.1002/mar.20199
52. Weijters B, Cabooter E, Schillewaert N. The effect of rating scale format on response styles: The number of response categories and response category labels. International Journal of Research in Marketing. 2010;27: 236–247. doi:10.1016/j.ijresmar.2010.02.004
53. Krosnick JA, Holbrook AL, Berent MK, Carson RT, Hanemann WM, Kopp RJ, et al. The Impact of "No Opinion" Response Options on Data Quality. Public Opinion Quarterly. 2001;66: 371–403. doi:10.1086/341394

54. Romans SE, Kreindler D, Asllani E, Einstein G, Laredo S, Levitt A, et al. Mood and the menstrual cycle. Psychotherapy and Psychosomatics. 2013;82: 53–60. doi:10.1159/000339370
55. Mays N, Pope C. Assessing quality in qualitative research. BMJ. 2000;320: 50–52. doi:10.1136/bmj.320.7226.50
56. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural Topic Models for Open-Ended Survey Responses. American Journal of Political Science. 2014;58: 1064–1082. doi:10.1111/ajps.12103
57. Lee JD, Kolodge K. Exploring Trust in Self-Driving Vehicles Through Text Analysis. Human Factors. 2020;62: 260–277. doi:10.1177/0018720819872672
58. Moslmi TA, Omar N, Abdullah, S, Albared M. Approaches to cross-domain sentiment analysis: A systematic literature review. IEEE Access. 2017;5: 16173–16192.
59. Masho SW, Adera T, South-Paul J. Obesity as a risk factor for premenstrual syndrome. Journal of Psychosomatic Obstetrics & Gynecology. 2005;26: 33–39. doi:10.1080/01443610400023049
60. Yan T. Consequences of asking sensitive questions in surveys. Annual Review of Statistics and Its Application. 2021;8: 109–127. doi:10.1146/annurev-statistics-040720-033353
61. Bancroft J. The menstrual cycle and the well being of women. Social Science & Medicine. 1995;41: 785–791. doi:10.1016/0277-9536(95)00045-9
62. Richardson JTE. The premenstrual syndrome: A brief history. Social Science & Medicine. 1995;41: 761–767. doi:10.1016/0277-9536(95)00042-6
63. Qualtrics. Provo, Utah USA: Qualtrics; 2005.
64. Yu M, Zhu X, Li J, Oakley D, Reame NE. Perimenstrual symptoms among Chinese women in an urban area of China. Health Care for Women International. 1996;17: 161–172. doi:10.1080/07399339609516230
65. Bertone-Johnson ER, Hankinson SE, Bendich A, Johnson SR, Willett WC, Manson JE. Calcium and Vitamin D Intake and Risk of Incident Premenstrual Syndrome. Arch Intern Med. 2005;165: 1246–1252. doi:10.1001/archinte.165.11.1246
66. Doyle Caroline, Ewald HASwain, Ewald PW. Premenstrual Syndrome: An Evolutionary Perspective on Its Causes and Treatment. Perspectives in Biology and Medicine. 2007;50: 181–202. doi:10.1353/pbm.2007.0015
67. Ussher JM, Perz J. Empathy, Egalitarianism and Emotion Work in the Relational Negotiation of PMS: The Experience of Women in Lesbian Relationships. Feminism & Psychology. 2008;18: 87–111. doi:10.1177/0959353507084954
68. Brown J, Majoribanks J, O'Brien P, Wyatt K. Selective serotonin reuptake inhibitors for premenstrual syndrome. 2009; 124.
69. Weiss G, Goldsmith LT, Taylor RN, Bellet D, Taylor HS. Inflammation in reproductive disorders. Reproductive Sciences. 2009. pp. 216–229. doi:10.1177/1933719108330087
70. Ussher JM, Perz J. PMS as a process of negotiation: Women's experience and management of premenstrual distress. Psychology & Health. 2013;28: 909–927. doi:10.1080/08870446.2013.765004
71. Lorenz TK, Demas GE, Heiman JR. Interaction of menstrual cycle phase and sexual activity predicts mucosal and systemic humoral immunity in healthy women. Physiology & behavior. 2015;152: 92–8. doi:10.1016/j.physbeh.2015.09.018
72. Ginn J, Silge J. qualtRics. 2020. Available: https://docs.ropensci.org/qualtRics/
73. Wan Y, Gao Q. An Ensemble Sentiment Classification System of Twitter Data for Airline Services Analysis. 2015 IEEE International Conference on Data Mining Workshop (ICDMW). 2015. pp. 1318–1325. doi:10.1109/ICDMW.2015.7
74. Rinker TW. Textclean: Text Cleaning Tools. Buffalo, New York; 2018.
75. Rinker TW. Sentimentr: Calculate Text Polarity. Buffalo, New York; 2019.
76. Mostafa MM. More than words: Social networks' text mining for consumer brand sentiments. Expert Systems with Applications. 2013;40: 4241–4251. doi:10.1016/j.eswa.2013.01.019
77. Moslmi TA, Omar N, Abdullah S, Albared M. Approaches to cross-domain sentiment analysis: A systematic literature review. IEEE Access. 2017;5: 16173–16192.
78. Rinker TW. Dictionaries and Word Lists for the 'qdap' Package. Buffalo, New York; 2018.
79. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338: b2393. doi:10.1136/bmj.b2393
80. Mayer M. missRanger: Fast imputation of missing values. 2021. Available: https://CRAN.R-project.org/package=missRanger

81. Neelon B, O'Malley AJ, Smith VA. Modeling zero-modified count and semicontinuous data in health services research Part 1: background and overview. Statistics in Medicine. 2016;35: 5070–5093. doi:https://doi.org/10.1002/sim.7050
82. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: A Probabilistic Programming Language. Journal of Statistical Software. 2017;76: 1–32. doi:10.18637/jss.v076.i01
83. Bürkner P-C. brms: An R Package for Bayesian Multilevel Models Using Stan. Journal of Statistical Software. 2017;80: 1–28. doi:10.18637/jss.v080.i01
84. Gabry J, Simpson D, Vehtari A, Betancourt M, Gelman A. Visualization in Bayesian workflow. Journal of the Royal Statistical Society: Series A (Statistics in Society). 2019;182: 389–402. doi:https://doi.org/10.1111/rssa.12378
85. Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Statistics and Computing. 2017;27: 1413–1432. doi:10.1007/s11222-016-9696-4
86. Ellis A. Learn multilevel models: An introduction to brms. 2021. Available: https://awellis.github.io/learnmultilevelmodels/walkthrough-brms.html/

Figures

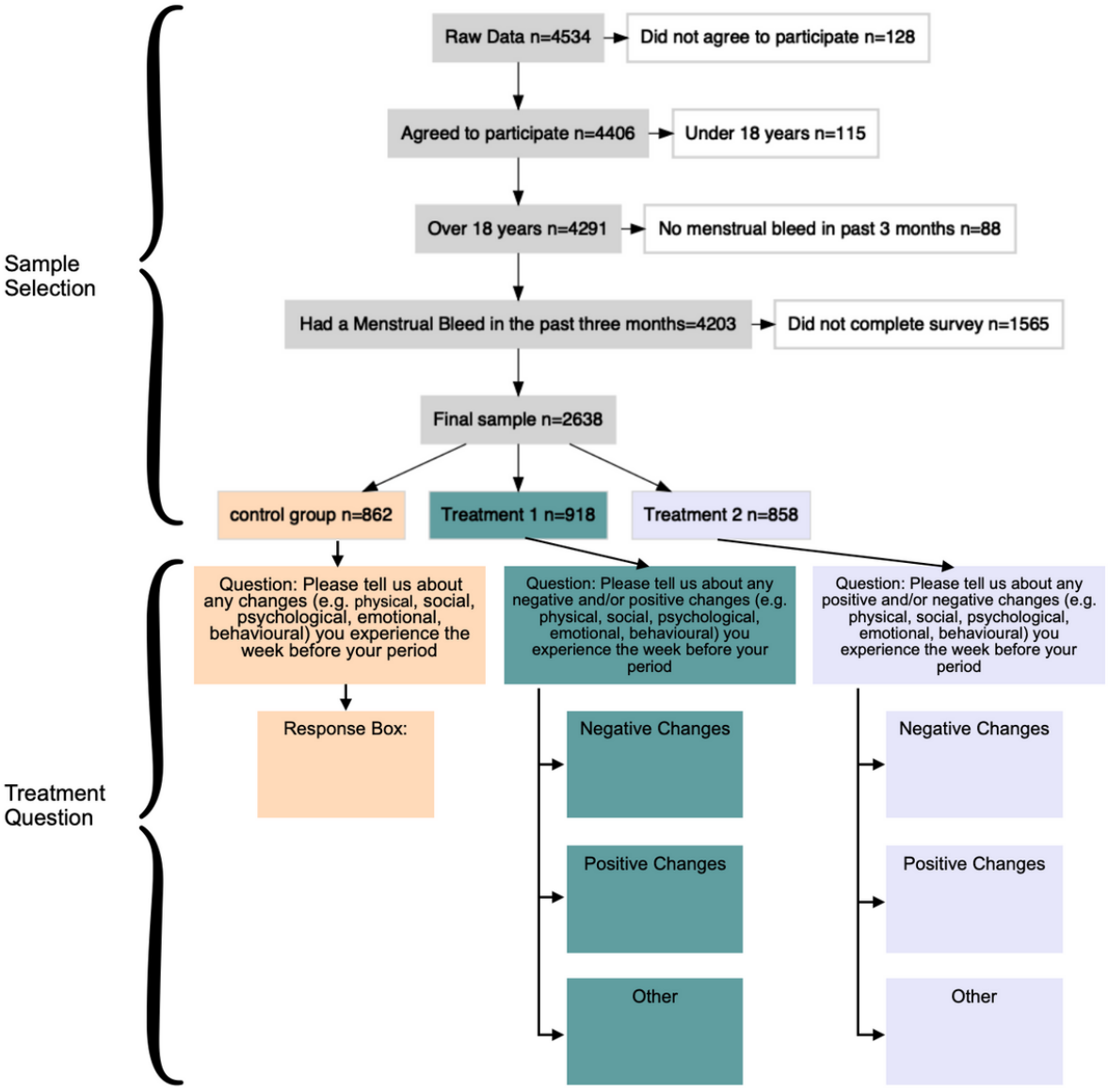

Figure 1

Flow chart of sample selection and treatment allocation.

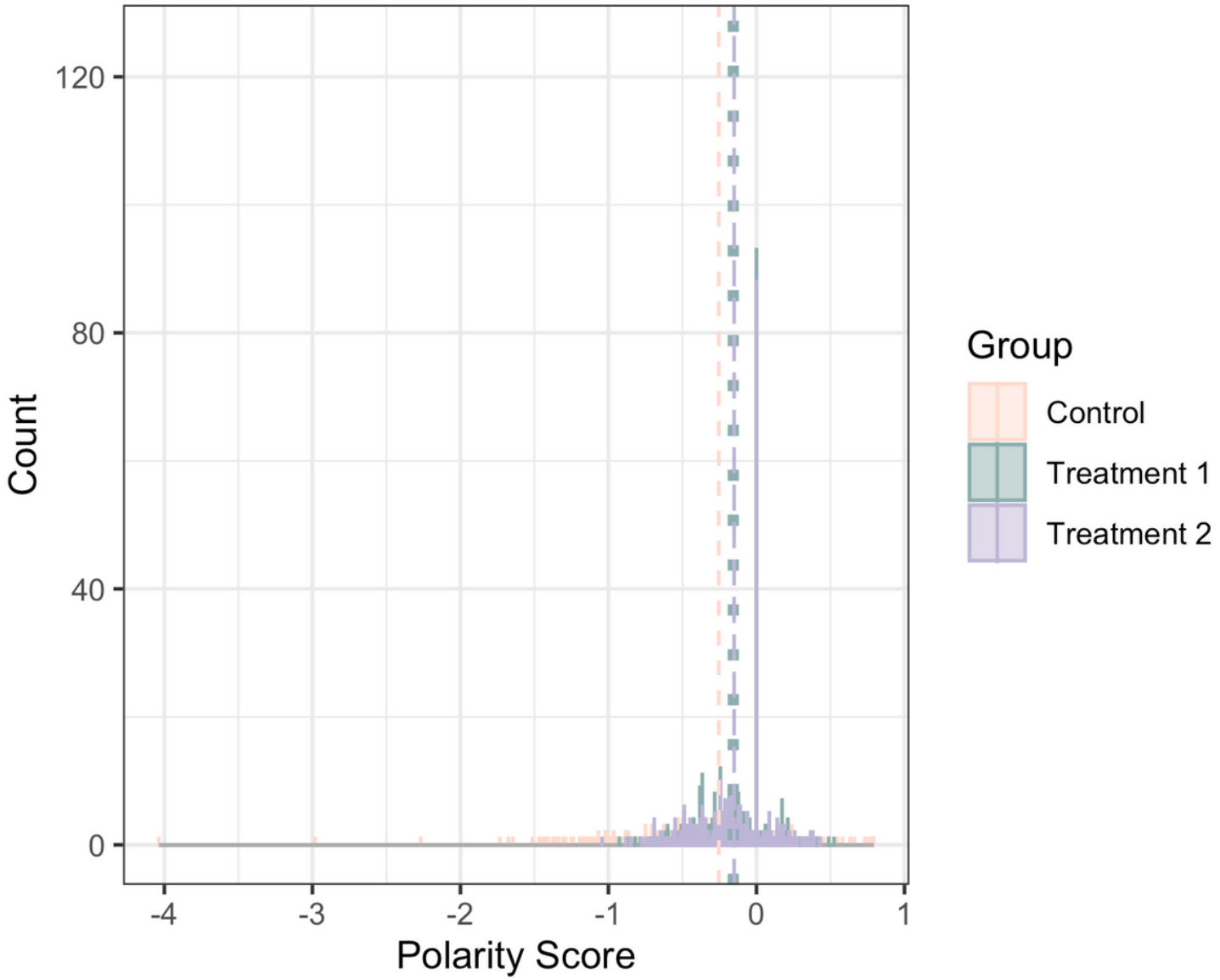

Figure 2

Framing effect on the polarity of premenstrual experiences (PME). There is a negative skew for all groups, with the control group showing the most extreme negative responses. The most common score in treatment one and two is 0. The mean for the control group is seen in the dashed pink line, dotted green for treatment one and long dashed purple for treatment two. The most common counts for treatment one and two are at 0, due to the large number of 'neutral' responses, thus the lines overlap.

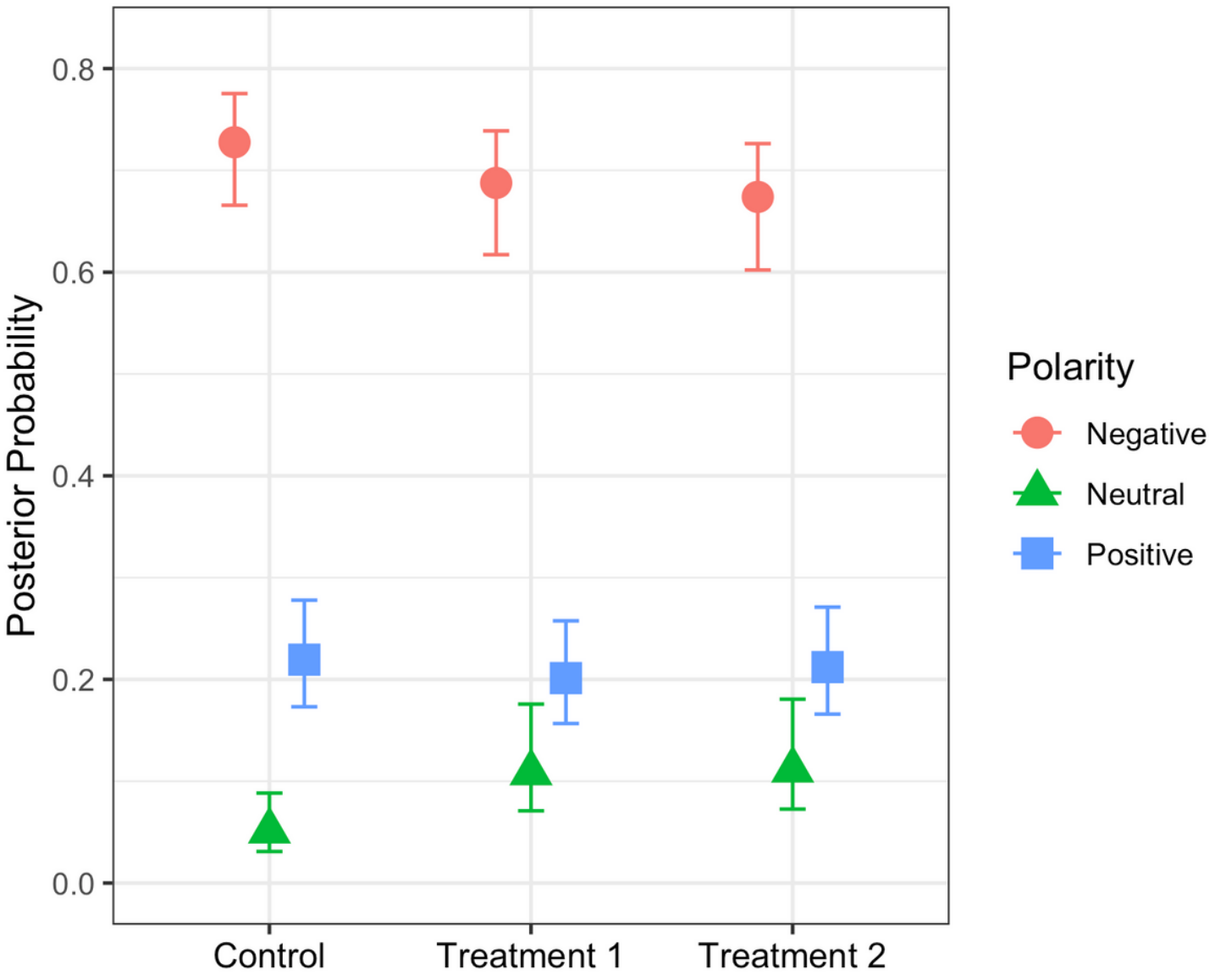

Figure 3

Posterior probability of the polarity response, conditional on all other predictors. Error bars indicate 95% confidence intervals. Red circles: negative experiences; green triangles: neutral experiences; blue square: positive experiences. Responses are overly negative across treatment groups, but participants in treatment groups one and two contributed more neutral responses.

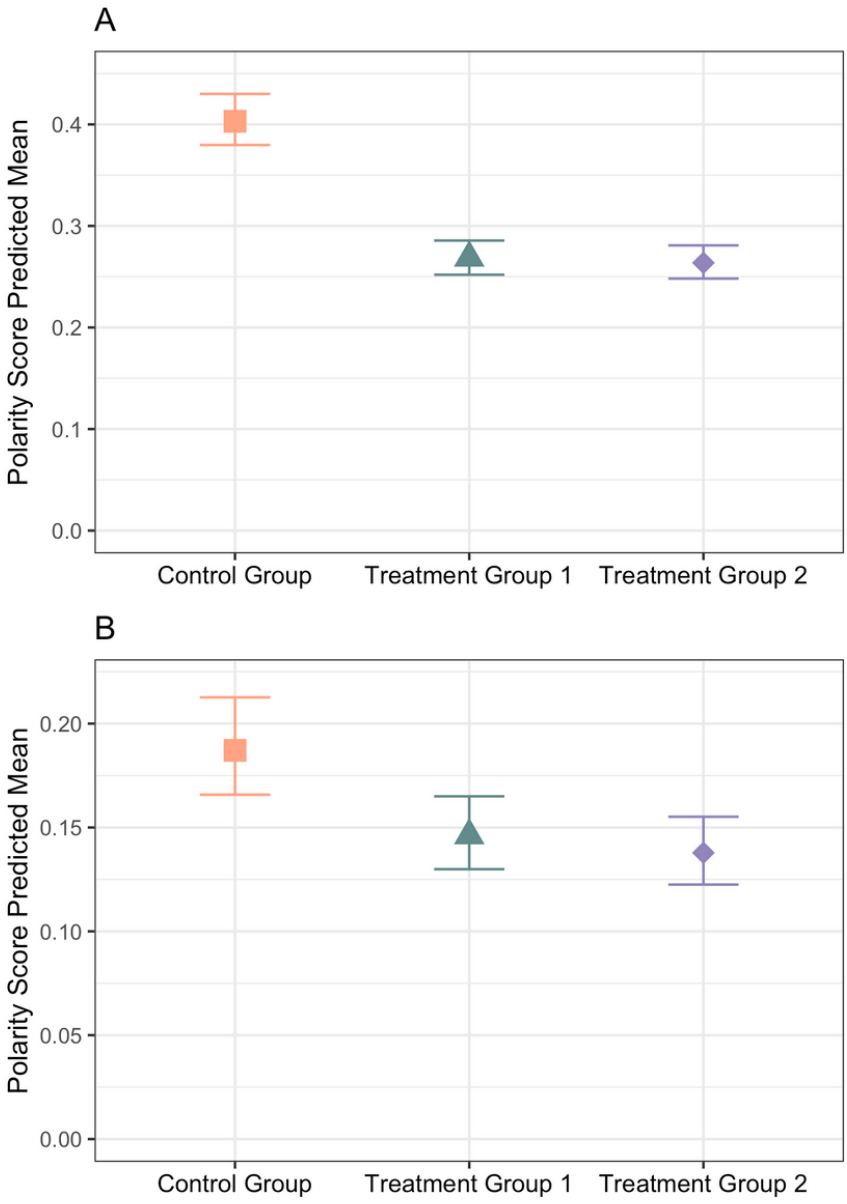

Figure 4

Predicted means of emotional intensity score and its 95% Credible Intervals (A) negative polarity scores; (B) positive polarity scores. The small 95% credible intervals which do not overlap between control and treatments one and two in facet (A) suggests differences in the range of negative scores across experimental groups. Larger 95% credible intervals in treatments one and two in facet (B) suggests less unreliable differences in the range of positive responses.

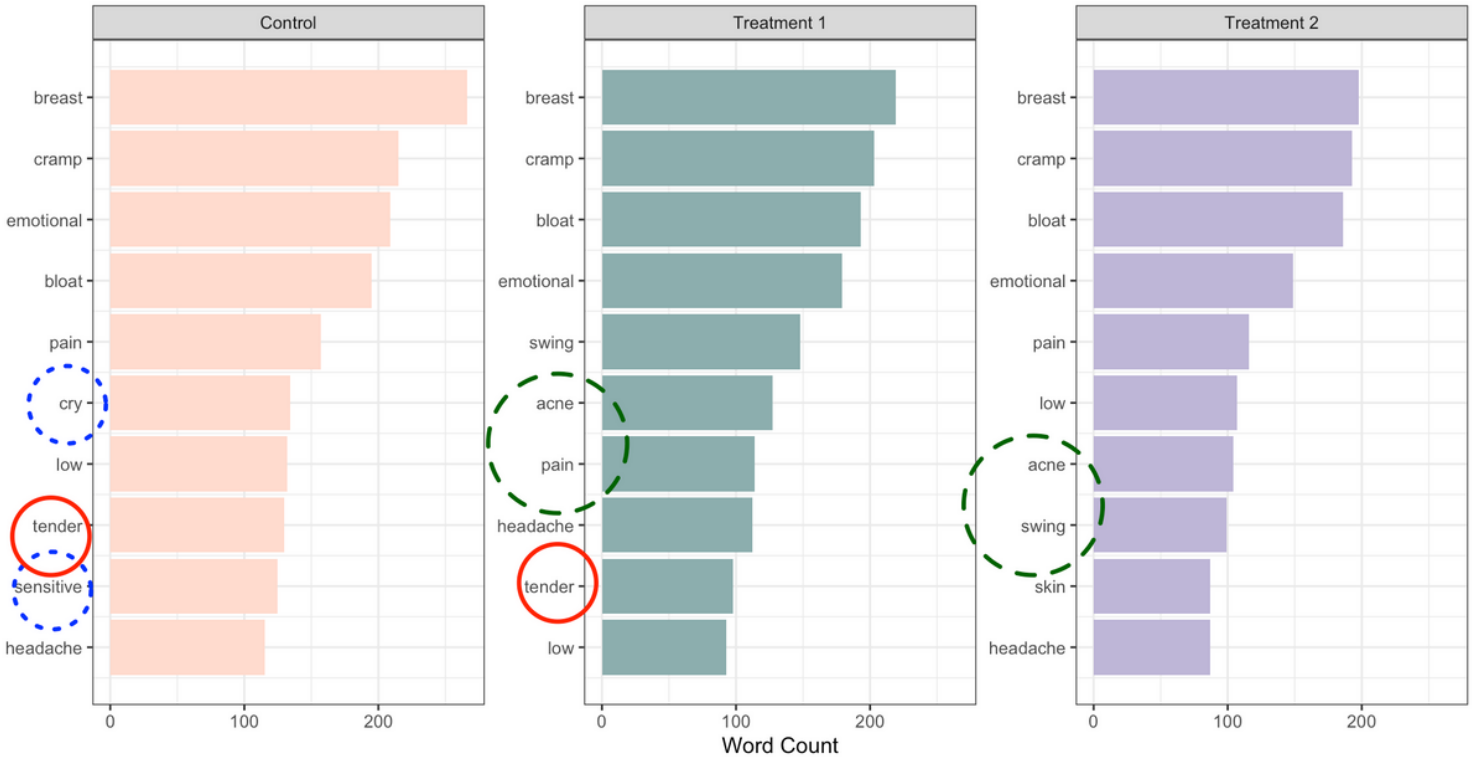

Figure 5

The ten most frequent words found in each group. Many words such as ‘breast’, ‘cramp’ and ‘bloat’ are found across all three groups, while the order of these varies slightly between the groups. Dotted blue circles denote the most frequent words only found in the control group (‘cry’ and ‘sensitive’). Solid red circles denote the most frequent words found in control and treatment one (‘tender’) while the dashed green circles denote the most common words only found in treatment one and two (‘swing’ and ‘acne’).

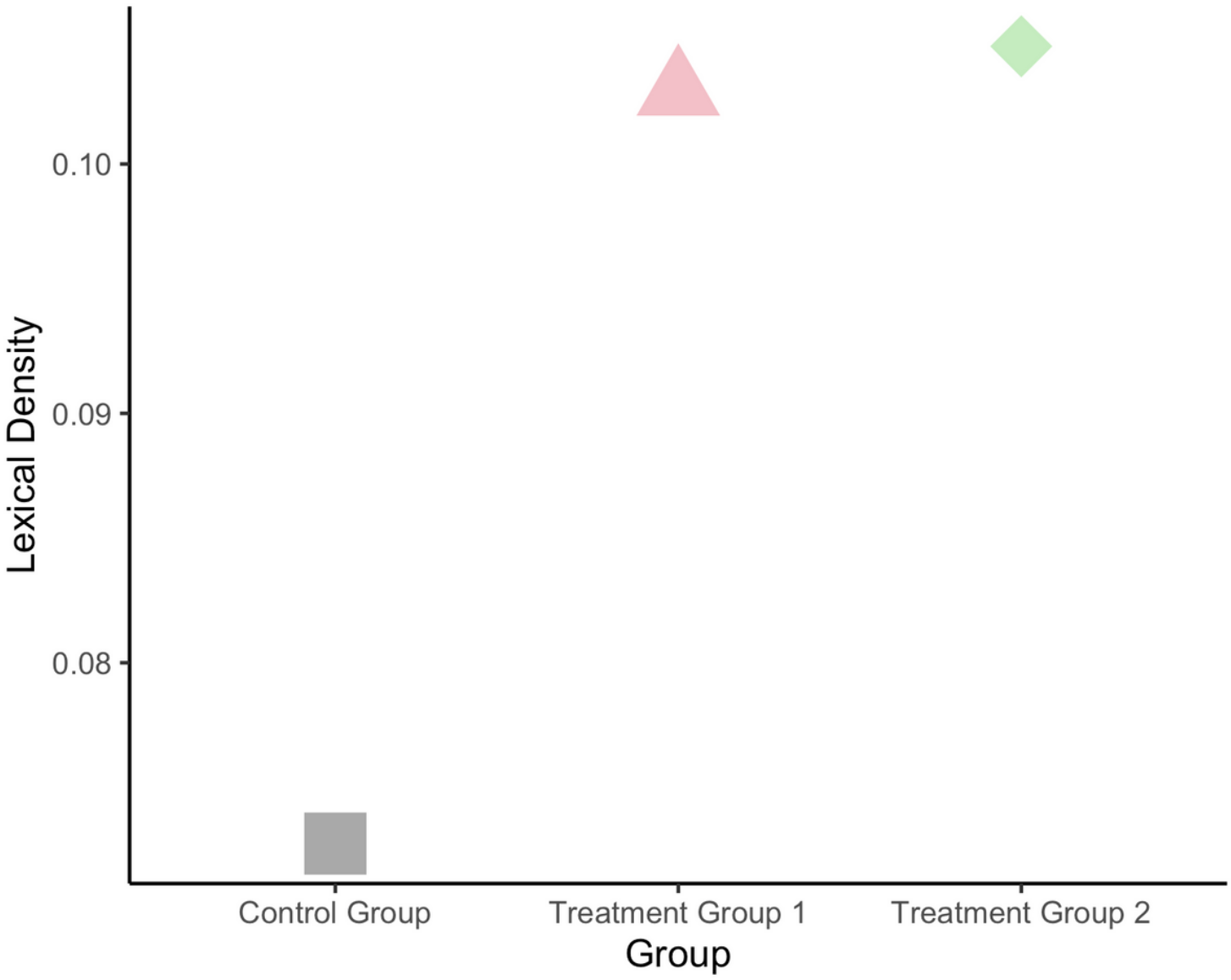

Figure 6

Lexical density across treatment groups. Lexical density corresponds to the ratio of unique words relative to the total number of words in each group. Treatment groups one and two contain significantly more unique words than the control group (Kruskal-Wallis χ^2=8.73, df=2, p <0.05).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryInfo.pdf