

Title

Exploratory analysis of language differences in England and Wales' case law judgements of sexual offences, before and after the #MeToo movement.

Introduction

Sexual offences, most often perpetrated by men against women (World Health Organisation [WHO], (2021), have been chronically mishandled by criminal justice systems. With most sexual crimes never making it to court, due to the fostering of a hostile environment in reporting, policing and judicial processes (e.g., Ministry of Justice [MoJ], 2021), explorative research into what data we have available on such topics is of the utmost importance. This research seeks to investigate the possible differences in the language and themes present in legal judgments of sexual, gender-related offences before and after the #MeToo movement. MeToo, the social justice campaign which promotes awareness of violence against women (Me Too Movement, 2024a), signalled a shift in tolerance for the blatant ignorance and denial of misogyny. Assessing whether measurable changes in the way sexual offences are adjudicated, using data drawn from The National Archives [TNA] (2024a) of England and Wales, can aid in revealing how social movements can potentially enact change within justice systems.

This study utilises different forms of text analysis, namely exploratory descriptive statistics, K-means clustering and sentiment analysis, to delve into this issue. Findings suggest that language within the court judgments show little distinguishable differences between before and after the #MeToo movement. However, research on this topic would benefit from a replication of the following analysis on a much larger sample size, allowing for more robust conclusions over the true impact of the campaign on structural justice mechanisms. This report is structured with a brief look at the history of the MeToo movement and a review of the current literature, followed by research questions, methodology, the results, a discussion over the findings and finally, a conclusion.

Literature

In October 2017, a culture of vicious misogyny and untold stories of sexual violence within Hollywood erupted over social media (Hillstrom, 2018). The #MeToo movement was popularised as a result of an initial X (formerly known as Twitter) exchange, where actress Alyssa Milano (2017) asked "If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet." as a response to the Harvey Weinstein sexual abuse scandal (Hillstrom, 2018). Hundreds of thousands of women across the globe engaged with the tweet (Seales, 2018), and #MeToo became a pinnacle of the start of social change within the modern world. The concept was originally coined in 2006 by Tarana Burke, an intersectional activist and the creator of a non-profit that helps victims of sexual abuse (Garcia, 2017). The grassroots campaign emerged with Burke creating a community to help survivors get justice, provide solidarity and aid in breaking down misguided stereotypes and abhorrent tropes surrounding victim-blaming (Me Too Movement, 2024b).

The MeToo movement not only encouraged victims of sexual offences to speak up, but raised concerns over the current social, political, economic and cultural climate which seeks to maintain the

patriarchy and protect the notion of superiority for white, cisgender men (Nichols, 2019). Within this, judicial processes and court judgements were brought into question as sexual crimes have historically been underreported and justice systems deemed systemically not fit for purpose (Andersson et al., 2024). Kohn (2018) argues that though the #MeToo movement was crucial in generating debate and bringing attention to the failure of our current criminal justice system designs, minor headway was achieved practically. Sexual offences continue to suffer from low prosecution rates, and a culture of misogyny remains within the police and courts (e.g. MoJ, 2021; Turner, 2024).

On the contrary, Prasad (2018) establishes the impact of #MeToo on judicial processes, where United States [US] state legislators worked to ban the use of non-disclosure agreements in sexual offence cases. Additionally, Bowman Williams et al. (2024) found that proposed bills in the US regarding harassment and gender equity reform increased after the popularisation of MeToo, but progress slowed as discourse on the subject dwindled. Langen (2024) conducted text analysis on US state and federal appellate courts to study the impact of the MeToo movement; the findings show a lack of direct causal effect on language due to the justice campaign, but acknowledge a possible influence on the reduction of victim-blaming.

Research questions

Are there discernible differences in language within legal judgments of sexual, gender-related offence case law, before and after the #MeToo movement?

Data and methods

This research uses data from The National Archives, which serves as official repository for an abundance of historical documents from the United Kingdom [UK] government and courts (TNA, 2024b). This includes records of court judgments and decisions from 2003 onwards, made publicly available under the terms of open justice (TNA, 2024a). The archives receive judgements from the Privy Council, UK Supreme Court, Court of Appeal, High Court, Upper Tribunals, Employment Appeal Tribunal and the Family Court but does not routinely receive them from the County, Crown or magistrates' courts (TNA, 2024a). The courts decide which judgements are published, which aids in the preservation of privacy and safeguarding for applicable cases.

TNA website provides a filtering function and a search query for requested terms; this study used the phrases "rape", "sexual assault" and "sexual offence" to identify suitable judgements. By ensuring all sampled documents included these terms, the sample size was reduced to a manageable level, focusing on cases directly relevant to the research, rather than those that only briefly mentioned sexual offences. 15th October 2017 was used as the defining split point for the dates to categorise before and after the #MeToo social media movement. Purposive sampling was employed to select court judgements which best suited the research objectives; only cases that related to female victims were chosen, and all cases including children were excluded. Judgements from the Court of Appeal were also omitted to avoid redundancy and overrepresentation of cases that might not provide additional insights relative to the other courts. Though there is a focus on original documents, some appeals remained in the sample as they reached the highest courts. With the significance of these judgements, it was logical to retain them within the analysis. An example of such included an appeal

which concerned the Met Office police texts within the Sarah Everard case (Cobban v Director of Public Prosecutions [2024] EWHC 1908), which has a clear relevance to the MeToo movement due to structural misogyny and violence against women (Horeck et al., 2023). The sample consisted of 18 documents altogether, with 9 judgements for both pre and post-#MeToo dates. The equal weighting of cases in each period was a coincidence but fared well for a comparable distribution. However, the pre-#MeToo category included judgments ranging from 2004-2017, whereas the post-#MeToo group had a reduced scope of 2017-2024.

The methods strategy predominantly comprised exploratory and descriptive text analysis. This included the inspection of frequent terms and specific phrases relevant to the #MeToo movement, using n-grams analysis and term frequency-inverse document frequency [TF-IDF]. Specifically, this aided in the distinguishability of ‘important’ words relative to a judgement within a collection of the documents. Unsupervised learning, specifically K-means clustering, was employed to investigate whether the documents would be sorted into groups that were analogous to the pre and post-#MeToo categories. Principal component analysis was utilised to reduce the dimensionality of the data for plotting, as the dataset included a vast number of terms. Additionally, sentiment analysis was undertaken for the final means of comparisons between the two periods.

All text data was pre-processed and transformed prior to the aforementioned analysis, which included the creation of corpora, document-term matrices [DTM] and cleaning (making all text lowercase, removing stop words, problematic characters, non-alphanumeric and sparse terms).

Results

When conducting an exploration of the most frequently used terms in court case judgments related to sexual and gender-based offences, the pre-#MeToo period revealed that the top five words ‘case’, ‘police’, ‘claimant’, ‘court’, and ‘evidence’ were used more than 450 times each, as evidenced in Figure 1. The post-#MeToo documents, on the other hand, showed ‘court’, ‘case’, ‘section’, ‘act’ and ‘rights’ as the top terms, which had similar rates of frequency.

To inspect the documents further, topic-specific terms such as ‘blame’, ‘character’, ‘consent’, ‘power’ and ‘violent’ were searched for within each of the judgements. Figure 2 shows ‘consent’ to have the largest frequency disparities between the pre and post-#MeToo groups. Before the social media campaign, ‘consent’ appeared 161 times within the judgements compared to just 22 afterwards. The use of the phrase ‘good character’ was also reviewed; it occurred 12 times in five different documents pre-#MeToo, however was only present once post-#MeToo.

TF-IDF analysis was employed to highlight the significance of particular words within each judgement in comparison to the entire collection of documents. However, this analysis did not yield substantial insights into the broader implications of the #MeToo movement. Notably, gendered terms did not emerge as particularly prominent (see Appendix 1, 2).

With regard to unsupervised learning, K-means clustering revealed that the documents did not follow the post and pre-#MeToo distribution. All judgements were numbered from 1-18 (1-9 were pre-#MeToo and 10-18 were post). When sorted into two clusters, all documents except for two (9

and 17) were categorised together. Performing K-means clustering again, with the optimal number of clusters drawn from an elbow curve (see Appendix 3), resulted in Figure 3. 9 and 17 remained in a separate second cluster, and 2, 6, 11 and 15 emerged as a third cluster. Therefore showing no meaningful associations to the time periods based on the dates of the popularisation of the #MeToo movement.

Finally, the sentiment analysis results show variability with both positive and negative extremes within both periods. However, pre-#MeToo showed a wider range from -55.10 to 71.55, with more negative term associations. Prominent negative terms include ‘risk’, ‘assault’, and ‘rape’, indicating a focus on serious and adverse issues. In contrast, post-#MeToo texts exhibit more moderate sentiment values, from -39.40 to 49.35, with increased emphasis on positive terms like ‘right’ and ‘protection’. Figure 4 and 5, respectively, visualise such differences. Looking at the overall scores of emotional valence, a key difference appears where ‘disgust’ ranks higher within the post-#MeToo sample (see Appendix 4, 5).

Discussion of findings

Interestingly, the most frequent terms used in court case judgments did show an onus on ‘evidence’ before the #MeToo movement. This focus aligns with the emphasis on evidentiary support in legal proceedings prior to the social justice campaign, which advocated for believing women who came forward with their stories (Gilmore, 2023). However, drawing conclusions based solely on term frequency can be misleading. Variations in the data may stem from differences in the types of cases included in each sample. For instance, the post-#MeToo period showed an increased prevalence of terms related to case law, such as ‘section’, ‘act’, and ‘rights’. Therefore, the changes may simply reflect evolving judicial priorities or case types rather than a straightforward change in focus.

Discussions surrounding consent are paramount to the MeToo movement, with an emphasis on genuine consent having to be informed, voluntary, and unequivocal, which challenges previous norms (Hillstrom, 2018). The campaign seeks to address and rectify power imbalances and sexual misconduct through a clearer understanding of consent (Hillstrom, 2018). Therefore, the relative absence of the term post-#MeToo suggests the movement had less of an influence on judicial considerations in this regard. On the contrary, the use of ‘good character’ (often used as a mitigating factor at sentencing for the perpetrator) decreased substantially after the popularisation of #MeToo within the sampled documents. This suggests that the broader culture of victim-blaming may have been improved post-#MeToo, where improper justifications for alleviations of sentences carry less weight during court judgements.

Though the frequent terms show some variation between the pre and post-#MeToo documents, the results of the unsupervised clustering negate substantial differences. The unsupervised categorisations showed no relation to the previous distinct groups based on before and after the #MeToo movement. This implies that while there are observable shifts in the language used, these changes may not be sufficiently pronounced or consistent to be captured by the clustering method used.

The sentiment analysis indicates a slight shift in discourse: pre-#MeToo texts show a broader sentiment range with prominent negative terms, highlighting the serious cases dealt with in court. The post-#MeToo texts exhibit a more moderate range, with increased positive terms such as 'right'. This deviation could reflect a broader change in focus towards empowerment, but is more likely due to the types of cases included, as mentioned previously. Sentiment analysis is limited by its reliance on word-based scoring, which may not capture contextual nuances. Thus, various interpretations can be drawn from such results.

Overall, the results from this study imply the #MeToo movement had little impact on judicial language within case law but may suggest a move away from victim-blaming notions. These findings are largely consistent with those presented in Langen (2024).

Limitations

Confined by time and resources, this study suffers from a small sample size which limits the ability to draw especially meaningful results. The exploration is somewhat naive and surface-level considering the complexities of such a significant and consequential topic, but is an important contribution nonetheless, as it serves as a baseline for future analysis. With an improved search function on The National Archives website to filter all judgements relating to sexual and gender-based offence cases and an appropriate web-scraper, language and theme differences may be more apparent within the pre and post-#MeToo eras. An increased sample size would also allow for the exclusive use of judgements closer to the initial popularisation date of #MeToo (i.e. a period of 2014 to 2020), which provides a firmer grounding for robust inferences.

As with the majority of causal research, it is important to note that correlation is not a determining factor for causation. Isolating the independent effect of #MeToo on court judgements is impossible within our obscure social world, and is especially difficult with the enactment of new judicial processes. In addition to this, the sole use of a date to determine the impact of a social change movement must be considered a large assumption for analogous reasons.

Conclusions

In conclusion, while there are observable shifts in language, such as the reduction in phrases such as 'good character' and the emergence of more moderate sentiments post-#MeToo, most changes appear to be more reflective of an evolving judicial focus and case types rather than profound alterations in legal discourse. Hence, while some linguistic variations exist, they may not be sufficiently pronounced to indicate a substantial transformation in judicial language for sexual, gender-based case law, in England and Wales, due to the #MeToo movement.

To explore this research objective further, an analysis of language used in police reports, and the magistrates' and Crown courts, would be especially useful. Future research would also benefit from global comparisons to wholly understand the impact of social change movements on justice mechanisms.

References

- Andersson, U., Karlsson, L., McGlynn, C. and Wegerstad, L. (2024). After #MeToo: Law, Justice and Sexual Violence: Introduction to the Special Issue. *Feminist Legal Studies*. 32, pp. 137-141.
- Bowman Williams, J., Tippett, E. C. and Ramdin, A. (2024). Mind the Gap (s): Mitigating Harassment in a Post# MeToo Workplace. *Southern California Law Review*. 98(5), pp. 1-59.
- Garcia, S. E. (2017). The Woman Who Created #MeToo Long Before Hashtags. *The New York Times*. [Online]. 20 October. [Accessed 4 August 2024]. Available from: <https://www.nytimes.com/>
- Gilmore, L. (2023). *The #MeToo effect: What happens when we believe women*. Columbia: University Press.
- Hillstrom, L. C. (2018). *The #metoo movement*. California: Bloomsbury Publishing.
- Horeck, T., Ringrose, J., Milne, B. and Mendes, K. (2023). #MeToo in British schools: Gendered differences in teenagers' awareness of sexual violence. *European Journal of Cultural Studies*. 00(0), pp. 1-20.
- Kohn, L. S. (2018). #Metoo, wrongs against women, and restorative justice. *Kansas Journal of Law & Public Policy*. 28(3), pp. 561-586.
- Langen, H. (2024). The impact of the# MeToo movement on language at court A text-based causal inference approach. *PLoS one*. 19(5), pp. 1-25.
- Me Too Movement. (2024a). *History and inception*. [Online]. [Accessed 25 July 2024]. Available from: <https://metoomvmt.org/get-to-know-us/history-inception/>
- Me Too Movement. (2024b). *Tarana Burke: Founder*. [Online]. [Accessed 5 August 2024]. Available from: <https://metoomvmt.org/get-to-know-us/tarana-burke-founder/>
- Milano, A. (2017). *If you've been sexually harassed or assaulted write 'me too' as a reply to this tweet*. [X]. 15 October. [Accessed 3 August 2024]. Available from: https://x.com/Alyssa_Milano/status/919659438700670976
- Ministry of Justice. (2021). *The end-to-end rape review report on findings and actions*. London: H.M. Government.
- Nichols, L. (2019). Smashing the Patriarchy: the Political Possibilities in# MeToo and Feminist Political Theory. *Newcastle Business School Student Journal*. 2(1), pp. 14-20.
- Prasad, V. (2018). If anyone is listening, #MeToo: Breaking the culture of silence around sexual abuse through regulating non-disclosure agreements and secret settlements. *Boston College Law Review*. 59(7), pp. 2507-2550.

Seales, R. (2018). What has #MeToo actually changed? *BBC News*. [Online]. 12 May. [Accessed 4 August 2024]. Available from: <https://www.bbc.co.uk/news/world-44045291>

The National Archives. (2024a). *Find case law: Judgments and decisions from 2003 onwards*. [Online]. [Accessed 15 July 2024]. Available from: <https://caselaw.nationalarchives.gov.uk/>

The National Archives. (2024b). *What we have*. [Online]. [Accessed 30 July 2024]. Available from: <https://www.nationalarchives.gov.uk/>

Turner, A. (2024). Sexism and misogyny as traits of police culture: Problems, red flags and solutions. *International Journal of Police Science & Management*. 26(2), pp. 279-291.

World Health Organisation. (2021). Violence against women prevalence estimates, 2018: global, regional and national prevalence estimates for intimate partner violence against women. WHO: Switzerland.

Appendix

Appendix 1 - Pre-#MeToo TF-IDF

Appendix 2 - Post-#MeToo TF-IDF

Appendix 3 - Elbow plot

Appendix 4 - Emotion sentiment plot pre

Appendix 5 - Emotion sentiment plot post

Appendix 6 - Full code transcript