

**PARTITIONED HISTORIES AND BEYOND: A SEMI-  
AUTOMATED APPROACH TO IDENTIFICATION AND  
MITIGATION OF LINGUISTIC BIASES IN 1947  
PARTITION NARRATIVES**

**Ph.D. Thesis**

By  
**JUSTY JOSEPH**



**DISCIPLINE OF ENGLISH  
INDIAN INSTITUTE OF TECHNOLOGY INDORE  
FEBRUARY 2024**

**PARTITIONED HISTORIES AND BEYOND: A SEMI-  
AUTOMATED APPROACH TO IDENTIFICATION AND  
MITIGATION OF LINGUISTIC BIASES IN 1947  
PARTITION NARRATIVES**

**A THESIS**

*Submitted in partial fulfillment of the  
requirements for the award of the degree  
of*  
**DOCTOR OF PHILOSOPHY**

*by*  
**JUSTY JOSEPH**



**DISCIPLINE OF ENGLISH  
INDIAN INSTITUTE OF TECHNOLOGY INDORE  
FEBRUARY 2024**



# INDIAN INSTITUTE OF TECHNOLOGY INDORE

I hereby certify that the work which is being presented in the thesis entitled **PARTITIONED HISTORIES AND BEYOND: A SEMI-AUTOMATED APPROACH TO IDENTIFICATION AND MITIGATION OF LINGUISTIC BIASES IN 1947 PARTITION NARRATIVES** in the partial fulfillment of the requirements for the award of the degree of **DOCTOR OF PHILOSOPHY** and submitted in the **SCHOOL OF HUMANITIES AND SOCIAL SCIENCES, Indian Institute of Technology Indore**, is an authentic record of my own work carried out during the time period from December, 2019 to February, 2024 under the supervision of Prof. Nirmala Menon, Professor, School of Humanities and Social Sciences, Indian Institute of Technology, Indore

The matter presented in this thesis has not been submitted by me for the award of any other degree of this or any other institute.

(JUSTY JOSEPH)

---

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

(Prof. NIRMALA MENON)

---

**JUSTY JOSEPH** has successfully given her Ph.D. Oral Examination held on **02 July 2024**.

(Prof. NIRMALA MENON)

---



## ACKNOWLEDGEMENTS

I am profoundly grateful to my Supervisor, Prof. Nirmala Menon, whose unwavering guidance has been instrumental throughout this research. Even in moments of self-doubt, particularly when navigating the complexities of computational tools which is outside the realm of my formal training, her confidence in my abilities has been a source of unwavering encouragement. I greatly appreciate the freedom she has given me to find my path and the support she offered when needed which prevented several wrong turns. I appreciate her attention to the smallest details and ensuring the sheer perfection of my work. Her dedication to my success and commitment to excellence have left an indelible mark on this journey, for which I am truly grateful.

I would like to thank my PSPC members, Prof. Abhishek Srivastava and Dr. Ananya Ghoshal for their invaluable suggestions. I thank Dr. Deborah Sutton, Dr. Pieter Francois, and Dr. Priyanka Tripathi for their input regarding this work. I extend my gratitude to all the faculty members and non-teaching staff of the School of HSS for their support over the years. My gratitude extends to the Learning Resource Centre at IITI, and online academic resource-sharing platforms for providing access to secondary research materials. I acknowledge the Bodleian and Weston Libraries at the University of Oxford, Libraries at Bharathiar University, and Lancaster University, and the British Library for the access to resources and data used for this study. I acknowledge the financial support from the Ministry of Education for my Teaching Assistantship and SPARC for facilitating my visit to the UK. I would like to acknowledge our data annotators for their valuable assistance.

I am deeply indebted to Dr. B. Padhmanabhan, my M.Phil supervisor for all the foundational lessons in research and academic writing and for embodying and teaching me academic integrity. Additionally, I extend my thanks to Mohanapriya T and Christy Joseph for their guidance in comprehending computational models and for patiently clarifying my doubts.

Completing this research would have been all the more difficult were it not for the friendship and support of my friends and fellow scholars. I thank my colleagues at the DH lab and JPN Centre and scholars from SHSS, IITI, and the Dept. of History at Lancaster University for their assistance and constructive feedback. I am profoundly grateful to Kavi for doubling as a sounding board and guinea pig to badly written first drafts. I thank Sonal for helping me navigate the unfamiliar terrain of IIT Indore, especially during my bouts of illness. I extend my gratitude to Shanmuga, Shaifali, Reema, and Ankit for their invaluable insights, experienced opinions, and clarifications. I am thankful to Lalith, and Shenba for bringing a piece of Coimbatore back to me. As someone eager to jump and easier to bruise, I am grateful to Bhanu for holding me every time I stumbled. To Hussain, I am thankful for seeing the light for me and being the light for me. I owe it all to Aparna and Muhsi for the persistent clearing of cobwebs and being counselors to life drama.

DH Lab has been more than just a building for the last four years. I am grateful for the space that has been a safe abode for all my thoughts. To Jacku, Kili, and Wildie, I was fortunate to have had your canine companionship, especially through my night strolls. And at moments of despair, Real Madrid and Kerela blasters; Ronaldo and Vini Jr- Thank you for the football.

I would not have had the courage to embark on this research journey in the first place without the support of my family. Their sacrifice and love have molded me into what I am now and they are all my reasons. I am grateful to my parents for putting up with a PhD process they didn't quite understand. I deeply thank my brother, for shouldering all the weight and being there for the family, allowing me to complete this. I am grateful to Michu Kutty for all the warmth and the feel of home that I carry along. To Achach, Ben, and Lijin Chetayi, thank you for being patient recipients and prompt responders of anxious calls and hazy messages.

Finally, to all my friends back home, thank you for staying through it all, even through my absences, I owe you all more than you can imagine.

*To Amma, Appa, Unni, and Michu.*

*To the lands ensnared in the occupation's grip, to the lives  
claimed at the borders of hate and bias, and to the unbowed  
resistance of the subjugated...*





## SYNOPSIS

### **PARTITIONED HISTORIES AND BEYOND: A SEMI-AUTOMATED APPROACH TO IDENTIFICATION AND MITIGATION OF LINGUISTIC BIASES IN 1947 PARTITION NARRATIVES**

The 1947 Partition of the subcontinent resulted in the largest migration in human history and violence across the borders of India and the newly formed state of Pakistan. Mushirul Hasan, in his introduction to *The Partition Omnibus* (2002) writes “Both as an event and memory, it [Partition] has to be interpreted and explained afresh in order to remove widely held misconceptions. This is both a challenge and a necessity and it is indeed a theme where the historian’s craft must be used deftly”. Partition literature originates from official histories in India and Pakistan, where narratives either celebrated independence, downplayed Partition’s disruptions, or attributed responsibility for the violence that occurred on the ‘other’. Much of the early historiography was concerned with the political process that accompanied the process of transferring power. Indian nationalists see Partition as the outcome of divisive colonial policies eroding cultural unity, while Pakistani writers often viewed the establishment of a separate homeland as a response to the aspiration to preserve community values in the face of what they perceived as a tyrannical Hindu majority (Virdee, 2013). In the early 1980s, a transformative historiographical movement, led by scholars like Ranajit Guha, known as the Subaltern Studies School, emerged to provide an alternative history from the prevailing populist nationalist narrative. By the 1990s, this approach influenced Partition Studies, shifting from a ‘great men of history’ focus to a ‘history from below,’ especially prompted by the regional studies’ move. The catalysts included chilling parallels between Partition violence and the 1984 anti-Sikh riots, and the Golden Jubilee of Independence, prompting introspective and reflective writing on previously inconspicuous subjects like violence, rape, and the abduction of women.

Recently, scholars like Ayesha Jalal, have examined the causative factors, repercussions, and lived experiences of Partition. Urvashi Butalia's focus on oral histories and their impact on marginalized communities, alongside contributions from historians such as Mushirul Hasan and Gyanendra Pandey, have contributed to an enriched understanding of the political and communal dimensions of partition violence. Gendered perspectives, as elucidated by Ritu Menon and Kamla Bhasin, have highlighted the experiences of women during this tumultuous period. The historiographical landscape has evolved, moving beyond traditional narratives to embrace interdisciplinary approaches that consider the complex interplay of politics, society, and testimonies. The plethora of scholarly studies on the discursive analysis of partition literature has explored concepts ranging from linguistic loss, narrative structures, linguistic register across spatial axes, and the language of remembering. This thesis expands the scope of scholarly inquiries looking into the semantics and symbolism in partition literature by harnessing newly accessible digital tools to study the 'language of bias'.

Bias analysis is an overarching term applied to methods that estimate the rationale, direction, and magnitude of systematic errors and their sources that influence measures of associations in a text (Lash et al, 2016; Frampton et al, 2022). Understanding biases is of high societal and policy relevance as it allows readers to critically and independently review and reason information. Bias analysis covers a broad range of methods, extending from sentiment analyses to in-depth probabilistic analyses requiring substantial manual and computational labor. It is generally conducted by reviewing the data collection methods, selection and retention of data, commission, and omission of information in the text, choice of words and labels, and placement of the content. A realistic model of bias analysis is likely to be complex, as bias is a relative term and is seldom undefined.

Humanities research has an extensive tradition of investigating biases, initially focusing on media biases in the 1950s and gradually expanding to encompass bias mitigation in research and historical documentation. Recent studies concentrate on diverse bias categories such as publication bias, communocentric bias, and gender bias, employing metrics like event selection, source choice, content placement, and word choice for identification and minimization. Methodologies like content analysis, frame analysis, and meta-analysis, are predominantly employed in recognizing and quantifying biases. Manual bias analysis is not scalable to the quantity and pertinence of the produced data and disparities like subjective interpretation, individual preferences and prejudices, the subtlety of the bias cues, and the difficulty of the sentence and word level identification often make manual mitigation of biases difficult.

Parallely in computational linguistics and computer sciences, bias investigations that originated with Lin et al.'s work in 2006. Various dimensions of media bias, such as perspective, ideology, truthfulness, and hyper-partisanship, have been explored using different computational methods, including low-level lexical information, linguistic cues, and deep learning techniques. However, there is a lack of consensus in defining media bias, with automated metrics often providing vague interpretations, and the literature review emphasizes the need to consult the comprehensive literature on bias analysis in humanities and social sciences for more effective automated tools. Identifying this research gap, this research advocates for the integration of humanities and computational methodologies to factor in the limitations in the analysis of bias patterns in complex narratives, thereby developing more efficient and effective tools.

## Research Objectives

With the above background, this thesis has four primary objectives.

1. To test and compare four media bias-detecting metrics to identify the gap and devise a target-oriented bias analysis tool to identify and mitigate biases in texts on the 1947 partition.
2. To understand the multifaceted roles of historical texts, testimonies, school textbooks, and social media narratives in the construction and mediation of sustainable mnemonic hegemonies and the dissemination of 'subjugated knowledges' pertaining to Partition.
3. To study the Linguistic patterns of narrative construction and bias manifestation in partition historiography focusing on the consequential social implications of such biases.
4. Through a methodological probe, the study also looks into the applications, challenges, and ethical concerns in employing AI-integrated models for digital humanities research, especially in interpreting individual texts as opposed to patterns in a corpus.

## Methodology

The tool developed for the linguistic bias analysis for texts at different levels of granularity ranging from words, sentences, and paragraphs to the entire text identifies, categorizes, and suggests the manifestation across ten linguistic bias categories. The dataset used to develop the model consists of 6,850 sentences from select partition texts and sentences from Wikipedia articles on partition under the NPOV dispute category. For the detection module, this study adapts the methods from the metrics of Chen et.al (Chen, 2020; Chen, 2020; Chen, 2018) and Pryzant et.al (Pryzant, 2020). RNN classifiers were employed for the classification tasks, where each cell is a Gated Recurrent Unit (GRU) of a recurrent rate size of 32. The pre-trained word embedding model GloVe with a word embedding dimension of 50, the optimizer Adam, and a learning rate of 0.001 are employed. Figure 1 demonstrates the architecture and workflow

of the devised model. In the analysis of the selected texts, the RNN class works effectively in capturing semantic information at multiple granularities. An in-depth reverse feature analysis is used to disseminate biases from the article level to other levels of granularity. The biased sentences are then correlated with Linguistic Inquiry and Word Count (LIWC) categories (Pennebaker et al., 2015), and the sequential patterns of bias and other contextual collective information are identified.

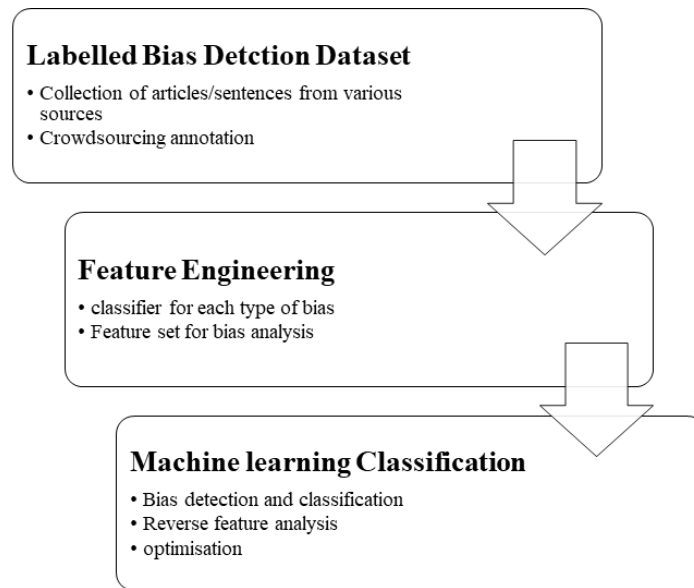


Figure 1: Architecture and workflow of the model

Close reading in the context of this research is defined as the practice of examining closely the language of a literary work or a section of it to view the minute or the unseen. “Close reading need not involve detailed interpretation of literary passages” it is about “attention to how meaning is produced or conveyed, to what sorts of literary and rhetorical strategies and techniques are deployed to achieve what the reader takes to be the effects of the work or passage” (Culler, 2011). The thesis follows the theoretical framework for close reading in Martin Paul Eve’s *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell’s Cloud Atlas* and the “back to the text” argument in Alan Liu’s, “The State of the Digital Humanities: A Report and a Critique,” and Tanya

E. Clement's, "Text Analysis, Data Mining, and Visualizations in Literary Scholarship," to close read Urvashi Butalia's *The Other Side of Silence*, Nisid Hajari's *Midnight's Furies*, NCERT and Pakistani Board school textbooks and Tweets on 14,15 August 2021 about the declaration of Partition Horrors remembrance Day using the devised bias analysis tool.

## **Thesis Structure and Key Findings**

### **Chapter 1: Introduction**

The first chapter introduces the concept of linguistic biases and its types. It also summarises the recent and ongoing studies on Partition history and bias studies. The chapter serves both as a detailed overview of the background, literature review, research gaps, and primary objectives of the thesis and a roadmap for how this thesis moves between primary texts of partition, secondary textual criticism, computational tools and methodologies as well as the concepts of linguistic bias.

### **Chapter 2: Computational Methods for Identifying and Categorizing Bias in Textual Data Across Various Granularity Levels**

This chapter traces the history of the methodological framework for bias analysis by outlining its major approaches in social sciences and computer sciences. We test and compare four exquisite media bias-detecting metrics namely,

1. Bias Determination using content sentiment analysis algorithm.
2. "Linguistic Models for Analyzing and Detecting Biased Language" (Recasens2013linguistic)
3. "Automatically Neutralizing Subjective Bias in Text" (Pryzant, 2020)
4. "Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity" (Chen, 2020)

in terms of accuracy, performance while using longer texts, analysis speed, labeling and inclusion of different types of biases, optimal

embedding approach, effective utilization of defined approaches in humanities and social sciences, sentence-to-word and sentence-to-article level accuracy and efficiency on events not covered in the training data. Based on the inferences from the comparative analysis, the chapter proposes a semi-automated target-oriented memory and narrative bias detection and classification tool for texts of the 1947 partition of India and Pakistan at different levels of granularity.

### **Chapter 3: Historical Memory, Language, and Biases: Re-reading *Midnight's Furies* and *The Other Side of Silence***

The first section of this chapter studies Nisid Hajari's *Midnight's Furies: The Deadly Legacy of India's Partition* (2015). The section briefs about the facilitation of historiography through archival study and the twofold manifestation of biases in Hajari's work, primarily in the narrative pattern and secondarily arising out of the archiving and archival study. The linguistic bias analysis using the tool described in the methodology section underscores the susceptibility to bias through misinterpretation, omissions, citation failures, and misleading explanations in the work. A notable deficiency in conveying the profound sentiments driving the call for Pakistan is observed, as the absence of Muhammad Iqbal's influence leaves the historical impetus solely attributed to Jinnah. The narrative is marred by inaccuracies, such as the misattribution of quotes and excessive reliance on secondary sources. Hajari's bias is primarily rooted in slanted objectivity and framing effect, particularly against Jinnah. His religious and political biases are revealed through a lack or an overloading of evidence. Moreover, his reliance on selective evidence and references to present-day incidents introduces biases like misinformation effect, attribution, confirmation, and selection biases. The section also intrigues bias manifestation through the chain of causality of events and metaphors employed by the author. The section highlights the pitfalls in Hajari's approach to history as a perennial conflict between two men who are central to the events. Discussing the subtle religious and gender biases in Hajari's historiography, the section

argues that the complex history of 'nationalist and patriarchal fetish on women's sexuality' is reinforced without criticism in Hajar's narrative. Hajri's work is clear evidence of representational deficiency in the social and cultural historiography of the 1947 Partition irrespective of the Oral history wave.

The second section studies Urvashi Butalia's *The Other Side of Silence: Voices from the Partition of India* by tracing the pathway of employing testimonies in Partition historiography and the associated challenges. The computational analysis reveals inherent biases, notably memory biases, often influenced by the author's acquiescence to interviewee perspectives, particularly entrenched patriarchal notions. The reluctance to acknowledge women as perpetrators of violence results in their repositioning into symbolic roles of sacrifice or victimhood. The author also submits to the pervasive perception of men as heroic soldiers or perpetrators of violence and women as victims. The narratives consistently attribute violence during Partition as an external act perpetrated by individuals outside the affected communities, contributing to the preservation of a communal 'other.' Besides the observable biases like egocentric, patriotic, gender, confirmation, selection, ad hominem, and framing effects, the text underscores childhood amnesia's impact on the reliability of memories recounted by adults reflecting on childhood experiences during Partition. The fallibility of memory emerges as the most conspicuous bias, leading to the rewriting of events with new errors, intricate fabrications, and susceptibility to the Zeigarnik effect. Discrepancies in narratives from individuals within the same family or locale regarding communal attitudes before Partition further highlight the challenges in historical interpretation. Urvashi Butalia navigates the complexities of writing a social history by employing a diverse array of sources while consciously avoiding rigid classifications such as official, factual, historical, or fictional. Recognizing the inherently narrative nature of social history, Butalia acknowledges her subjective position as a narrator,



emphasizing the interpretative aspect inherent in the crafting of any historical account. She deliberates on the decision to integrate her perspectives and understandings into the narrative, embracing a method that weaves together oral accounts and her interpretations. The thematic organization of the book under headings like 'Blood,' 'Facts,' 'Women,' and others allows the interviews to transcend individual stories, providing a comprehensive exploration of diverse themes. The narrative strategy, while potentially merging specificities, transparently communicates the subjective nature of historical interpretation, offering a distinct approach that openly acknowledges the author's agency and perspective in presenting history. The absence of interview questions and the interviewer in the book raises questions about the dynamics of the interviews, underscoring the omnipresent role of the author/narrator who holds exclusive agency in representing the interwoven stories. The author's narrative approach involves omitting questions, and interventions, and shaping interviews, losing nuances like inflections, hesitations, and body language in the transition to text. The narrative though centered around the concept of intellectual honesty, is composed of selective silence of certain aspects to avoid potential implications. Her attempt to challenge the notion of formal history as sacrosanct on the one hand and personal stories or oral histories as less worthy of formal documentation on the other is a thread that runs throughout the book. The chapter further explores how using the language of the Holocaust allows the authors to capture the politically volatile moments of the Partition and its aftermath adequately, as attempted with analyses of the Cambodian, Rwandan, Armenian, and Guatemalan genocides and the linguistic patterns of bias manifestation and narration of violence.

## **Chapter 4: ‘Master Narratives in (Re-En)Action’: 1947 Partition in School Textbooks and Social Media Narratives**

The first section of this chapter studies the Social Studies textbooks from Grade IV to Grade VIII and Pakistan Studies from Grade IX to Grade X compiled by the Punjab Textbook Board and NCERT textbooks from Grade VI to XII using the Semi-Automated tool. The chapter traces the history of school textbooks in India and Pakistan and discusses how Partition is portrayed in them. Bias analysis reveals a notable prevalence of positive sentences, indicating a linguistic positivity bias. This bias stems from writers' efforts to conform to established societal perceptions of the Partition, emphasizing objective circumstances and affective states. The victim-centered narration in both textbooks, along with the use of affect-inducing language to depict hardships, suggests a negative attribution bias. Notably, the higher use of passive voice construction in NCERT textbooks may impact emotional association and content recall for younger readers. This bias-inducing pattern is attributed to the concept of 'historical distance,' where writers strategically manipulate emotional responses to subtly instill their ideological values in readers. The section further argues that textbooks are perceived as 1. Memory Practice 2. ‘Boundary Object’ 3. National Ideology 4. Reinforcements of Homogeneity in understanding partition. It also highlights the Language of Causality and Othering, Selective Amnesia of Conflicts and Violence, and geopolitical implications of imagined cartographies in Partition Textbooks.

The second section titled ‘Recollection and Transmedial Reiteration: Twitter and the New Epistemology of 'Subjugated knowledges' about Partition’ studies the nature of discourses and the manifestation of biases in Tweets about the 1947 Partition on 14 and 15 August 2021 around Partition Horrors Remembrance Day. The analysis suggests a correlation between the degree of hostile comments and inherent bias in the original tweets, proposing the potential use of comment hostility as a proxy for

measuring bias in tweets. The manifestation of political bias in tweets is often seen through entities and hashtags, with nouns and noun phrases playing a crucial role in introducing bias. The introduction of bias is facilitated by sentiment-charged language, predominantly employing negative emotion words and emphasizing group identity. When expressing dissent towards the declaration of Partition Horrors Remembrance Day, both in primary tweets and replies, Twitter users employed polite and apologetic language, incorporating phrases such as 'In my humble suggestion' and 'I request to.' The section further argues that Twitter can be perceived as 1. Mediated sites of contention 2. Mnemonic medium 3. Site of false memory formation 4. Creator and manipulator of collective consciousness in recollecting and reiterating Partition. The study further highlights the transmedial structure of Twitter aiding the Recollection of the Past Within a Social Network's Collective Memory.

## **Chapter 5: AI Biases: A Digital Humanities Response**

This chapter traces the distinguishable phases in the evolution of automated practices in Digital Humanities. It briefs how DH uses AI in natural language processing, archiving, image processing, and text analytics. AI has revolutionized data curation and analysis in the humanities, offering researchers unprecedented insights beyond traditional approaches. AI challenges existing methods and disciplinary boundaries in DH, facilitating collaboration, and experimentation. However, the shift towards methodological rigor and data determinism in AI-enabled humanities research raises concerns about transparency, reproducibility, and the need for critical engagement with AI systems, positioning DH researchers as creative technologists, digital archivists, and critical humanists navigating the intersection of technology and humanistic inquiry. The chapter argues that Bias can infiltrate AI systems during data generation and management, dataset creation, and algorithm composition and evaluation. Factors contributing to dataset bias include non-existent or

biased data points, historical and current inequities, and subjective or discriminatory data labels, leading to datasets that may misrepresent or reinforce societal injustices. Algorithmic bias arises from narrow objectives, usage of biased datasets, lack of consideration for target populations, and inadequate auditing, resulting in inaccurate and discriminatory outcomes, compounded by a lack of accountability and transparency. At the user stage, biased AI outputs may occur when used in different contexts, altered discriminatorily, or interpreted with bias. The chapter also analyses biases in AI-integrated offensive speech detection by examining the results generated by TILT an online Hate Speech Detection Solution for Twitter as a ‘user case’. The analysis infers that though the Prior studies noted that terms related to 'Muslim' could trigger an algorithm's overreaction in predicting offensive speech, this study reveals that 54% of Islamophobic tweets went unnoticed, as the tweeps employed alternative terms and special characters to evade detection. The TILT algorithm exhibited a higher rate of false positives with identity terms, emphasizing the challenges in accurately identifying offensive content, especially in the context of dialectical variations and regional language usage in short text strings like tweets. The chapter also highlights the challenges of a DH researcher working on AI and concludes by proposing a set of ethical guidelines for humanities researchers working on or with AI models.

## **Chapter 6: Conclusion**

The concluding chapter summarises and reflects upon the analysis carried out in the four major chapters of the thesis. It briefly summarises the main research findings of the thesis and discusses its limitations and future scopes.

## References

- Chen W, Khatib K, Wachsmuth H et al (2020), Analyzing political bias and unfairness in news articles at different levels of granularity. arXiv preprint.
- Culler, Jonathan. (2011), “The closeness of Close reading.” Adfl, 8–13, (DOI:10.1632/adfl.41.3.8).
- Eve, Martin Paul. (2019), Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell’s Cloud Atlas. Stanford University Press.
- Frampton, G., Whaley, P., Bennett, M., (2022) Principles and framework for assessing the risk of bias for studies included in comparative quantitative environmental systematic reviews. Environ Evid 11, 12 (DOI: 10.1186/s13750-022-00264-0).
- Hajari N (2015) Midnight’s Furies: The deadly legacy of India’s partition. Houghton Mifflin Harcourt.
- Kumar J (2022) History, National Identity and Politics of History Textbooks in India. J Pol Sci Pub Aff. 10:417.
- Lash TL, Fox MP, Cooney D, Lu Y, Forshee RA (2016) Quantitative Bias Analysis in Regulatory Settings. Am J Public Health. Jul;106(7):1227-30. doi: 10.2105/AJPH.2016.303199.
- Liu, Alan (2012), The State of the Digital Humanities: A Report and a Critique, Arts and Humanities in Higher Education, 11, 8–41, (DOI: 10.1177/1474022211427364).
- Noble, Safiya Umoja. (2018), Algorithms of Oppression. New York, NY: New York University Press.
- Pryzant R, Martinez RD, Dass N, et al (2020) Neutralizing subjective bias in text. In: Proceedings of the aaai conference on artificial intelligence, 480–489.
- Recasens M, Mizil C, Jurafsky D (2013) Linguistic models for analyzing and detecting biased language. In: Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, 1650–1659.

Smit, R., Heinrich, A., & Broersma, M. (2018). Activating the past in the Ferguson protests Memory work, digital activism and the politics of platforms. *New Media & Society*, 20(9), 3119– 3139.

Tanya E. Clement, “Text Analysis, Data Mining, and Visualizations in Literary Scholarship,” *MLA Commons, Literary Studies in the Digital Age: An Evolving Anthology* (New York: Modern Language Association, 2013): <https://dlsanthology.mla.hcommons.org/text-analysis-data-mining-and-visualizations-in-literary-scholarship>.

# LIST OF PUBLICATIONS

## Published Papers

Menon, N., Shanmugapriya, T., **Joseph, J.** and Sutton, D. (eds.)  
(2023). *Indian Electronic Literature Anthology: Volume I*. India:  
IITI KSHIP. DOI: <https://doi.org/10.57004/book1>. License: CC BY  
4.0.

**Joseph, J.**, & Menon, N. (2023). "Textual Data Categories and  
Descriptors in Select Automated Bias Analysis Metrics: Methods,  
Applications, and Challenges", *Categories in DH*, Digital  
Humanities Quarterly, Vol. 17, no. 3.

**Joseph, J.**, & Menon, N. (2023). "Applications and Developments of  
NLP Resources for Text Processing in Indian Languages: Shared  
Multilingual Corpora Building and Pre-trained Models". *Digital  
Research in Arts and Humanities*.  
<https://doi.org/10.4324/9781003393696>.

**Joseph, J.**, Palanisamy K., Justin, J., & Menon, N. (2023). Digital  
Humanities Workshops in India: Effective Organizing Pedagogies  
and Sustainable Contributions to Academia. *Digital Humanities  
Workshops*, 48–55. <https://doi.org/10.4324/9781003301097-6>.

**Joseph, J.**, & Menon, N. (2023). "Digital Humanities Teaching Pedagogy  
in Virtual classrooms of India", MLA 22 special Issue, Duke  
University Press.

**Joseph, J.**, & Menon, N. (2022). Electronic literature in India: Where is  
it? does it even exist? Literary cultures and digital humanities in  
India (pp. 296-311) doi:10.4324/9781003354246-22

**Joseph, J.**, & Menon, N. (2022) "Enumerating Identities in the Certitude  
of African Indigeneity in Chimamanda Ngozi Adichie's *Americanah*",  
*IUP Journal of English Studies*. 17(1).

**Joseph, J.**, & Menon, N. (2020) "Traversing the obsolete past, irrefutable  
present and anonymous future through memories in Michael  
Ondaatje's *Warlight*.", *Rupkatha Journal for Humanities*, Vol. 12,  
No. 6, DOI: 10.21659/rupkatha. v12n6.18.

## **Papers Accepted**

**Joseph, J., & Menon, N.** “A Semi-Automated Approach to Identification and Mitigation of Biases in Textual Data”, Humanities and Social Sciences Communications. Nature Communications. (revisions submitted)

**Joseph, J., & Menon, N.** “AI Biases: A Digital Humanities Response”. In collection: Artificial intelligence: Digital humanities, business, & society, Humanities and Social Sciences Communications, Nature Publications. (paper proposal accepted)

## **Papers Under review**

**Joseph, J., & Menon, N.** “Named Entity Recognition Frameworks and Evaluation Strategies for Malayalam: Methods, Parameters and Challenges”, International Journal of Digital Humanities, Springer nature. (submitted on 12 September 2022)

**Joseph, J., & Menon, N.** “Who Owns the Data When Computer Learns your Mother Tongue? Ethics and Data Governance in Creating Machine Learning Models for Low Resourced Languages”, ACM Transactions on Asian and Low-Resource Language Information Processing. (submitted on 23 October 2023)



# TABLE OF CONTENTS

	<b>Pages</b>
List of Figures .....	xxxi
List of Tables .....	xxxii
List of Acronyms .....	xxxiii
Chapter 1: Introduction .....	1- 30
1.1.Intersection of History and Narrative: An Overview .....	1
1.2.Historical narrative and linguistic biases.....	3
1.3.Bias Analysis .....	6
1.4.Computational Tools for Text Analytics .....	7
1.5.The 1947 Partition and its Scholarly Landscape .....	10
1.6.Linguistic Biases .....	15
1.6.1. Confirmation Bias .....	15
1.6.2. Misinformation Effect.....	16
1.6.3. Framing Effect Bias .....	17
1.6.4. False memory effect.....	17
1.6.5. Political Bias .....	18
1.6.6. Religious Bias .....	19
1.6.7. Geographical bias .....	19
1.6.8. Hindsight bias .....	20
1.6.9. Representation Heuristic Bias .....	20
1.6.10. Selection Bias .....	20
1.6.11. The Zeigarnik Effect.....	20
1.6.12. Ego-Centric Bias.....	21
1.7.Research Objectives .....	21
1.8.Overview of the Thesis .....	21
1.8.1. Chapter 1: Introduction .....	22
1.8.2. Chapter 2: Computational Methods for Identifying and Categorizing Bias in Textual Data Across Various Granularity Levels .....	22
1.8.3. Chapter 3: Historical Memory, Language, and Biases: Re-reading Midnight’s Furies and The Other Side of Silence .....	23
1.8.4. Chapter 4: ‘Master Narratives in (Re-En)Action’: 1947 Partition in School Textbooks and Social Media Narratives .....	24
1.8.5. Chapter 5: AI Biases: A Digital Humanities Response .....	25

1.8.6. Chapter 6: Conclusion .....	26
References .....	28
Chapter 2: Computational Methods for Identifying and Categorising Bias in Textual Data Across Various Granularity Levels .....	31-76
2.1. Approaches to Textual Bias Analysis <sup>i</sup> in Humanities and Social Sciences .....	32
2.2. Bias Analysis Using Computational Metrics .....	34
2.3. A Comparative Study of Select Automated Media Bias Analysis Approaches .....	37
2.4. Bias Determination Using Content Sentiment Analysis Algorithm.....	38
2.5. “Linguistic Models for Analyzing and Detecting Biased Language” (Recasens,2013) .....	42
2.6. “Automatically Neutralizing Subjective Bias in Text” (Pryzant,2020) ....	46
2.7. “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen, 2020) .....	50
2.8. Semi-Automated Target-oriented Identification of Select Linguistic Biases at Different Levels of Granularity .....	57
2.9. ‘Close Reading with Computers’ .....	64
References .....	70
Chapter 3 Historical Memory, Language, and Biases: Re-reading <i>Midnight’s Furies</i> and <i>The Other Side of Silence</i> .....	77-136
Section 1.....	80-106
3.1.1. Archival Study and Historical Writing .....	80
3.1.2. Partition Historiography Through Archival Research .....	81
3.1.3. “Midnight’s Furies: The Deadly Legacy of India’s Partition” .....	83
3.1.4. Linguistic Bias Analysis of <i>Midnight’s Furies</i> .....	85
3.1.5. Archival Biases in Hajari’s Historiography.....	88
3.1.6. Bias Manifestation through ‘chain of causality’ of events developed by Hajari .....	91
3.1.7. ‘Jinnah the Fallen Angel’ Metaphor and Linguistic Bias Manifestation .	94
3.1.8. The Nehru Bias .....	97
3.1.9. Historiography as a Two-men issue and manifestation of Narrative Biases .....	100
3.1.10. Religious Bias .....	102
3.1.11. Portrayal of Women by Hajari and Linguistic Biases .....	103
3.1.12. Linguistic indicators of biases .....	105

Section 2 .....	106-133
3.2.1. Testimony, Memory, and Representation in Oral Histories .....	106
3.2.2. Oral History Narratives in Partition Studies .....	109
3.2.3. The Other Side of Silence – Urvashi Butalia .....	117
3.2.4. Linguistic Biases in <i>The Other Side of Silence</i> .....	119
3.2.5. Bias Manifestation Through Author’s Interventions .....	123
3.2.6. Selection of Interviews and Resultant Biases .....	124
3.2.7. Dual Perspectives: Butalia’s Use of Personal Narratives in Historical Inquiry and Doubts on its Efficiency .....	124
3.2.8. The one-sidedness of Butalia’s Narrative .....	126
3.2.9. Linguistic Indications of Biases .....	127
3.3. Language of Describing Violence in <i>The Midnight’s Furies</i> and <i>The Other Side of Silence</i> .....	127
3.4. The language of the Holocaust in narrating partition in <i>The Midnight’s Furies</i> and <i>The Other Side of Silence</i> .....	129
References .....	134
Chapter 4 ‘Master Narratives in (Re-En)Action’: 1947 Partition in School Textbooks and Social Media Narratives .....	137-200
Section 1 .....	140-178
4.1.1. Partitioned Histories: Portrayal of 1947 Partition in Select School Textbooks from India and Pakistan .....	140
4.1.2. Teaching History in School Classrooms .....	143
4.1.3. History of School Textbooks in India and Pakistan .....	144
4.1.4. History of NCERT Textbooks in India .....	145
4.1.5. Controversies Waged Against NCERT History Textbooks in India .....	147
4.1.6. Changes in NCERT History Textbooks Across Ages .....	148
4.1.7. History of Textbooks in Pakistan .....	151
4.1.8. Portrayal of Partition in NCERT (Indian) Textbooks .....	154
4.1.9. Portrayal of Partition in Punjab Textbook Board (Pakistani) Textbooks .....	157
4.1.10. Linguistic Bias Analysis of Chapters on Partition in Indian and Pakistani School Textbooks .....	160
4.1.11. The Language of Causality in Partition Textbooks and Manifestation of Biases .....	163
4.1.12. The Language of ‘Othering’ in Partition Textbooks and Manifestation of Biases .....	165
4.1.13. Geographical Bias in Textbooks .....	166

4.1.14. Selective Amnesia of Conflicts and Violence and Linguistic Biases	169
4.1.15. Understanding the Conceptualisation of History Textbooks in Remembering Partition .....	172
Section 2 .....	178-195
4.2.1. Recollection and Transmedial Reiteration: Twitter and the New Epistemology of 'Subjugated knowledges' about Partition .....	178
4.2.2. Partition Horrors Remembrance Day .....	181
4.2.3. Portrayal of Partition in Tweets .....	182
4.2.4. Linguistic Bias Analysis of Tweets .....	187
4.2.5. Partition Through the Lens of Twitter .....	191
4.2.6. Recalling the Past Within a Social Network's Collective Memory: History in the Digital Age .....	195
References .....	196
Chapter 5: AI Biases: A Digital Humanities Response .....	201-228
5.1. Digital Humanities, Artificial Intelligence, and Big Data.....	205
5.2. How does the Digital Humanities use AI? .....	206
5.3. Changing roles of DH research and researcher .....	210
5.4. Challenges of a Digital Humanist working with AI .....	212
5.5. Ethical and Social Implications of AI .....	214
5.6. AI Bias .....	215
5.7. Offensive speech detection in Tweets.....	218
5.8. Ethical Framework for Digital Humanists Engaging with AI .....	221
References.....	225
Chapter 6: Conclusion .....	229-238
6.1. Summary and Key Findings of the Chapters.....	231
6.2. Limitations .....	235
6.3. Scope for further research .....	236
References.....	238
Appendix A .....	239-242
Bibliography .....	243-254

## LIST OF FIGURES

Figure 2.4.1.: Architecture of the Sentiment Analysis Model.....	39
Figure 2.4.2.: Content Sentiment Analysis of Newspaper Reports .....	41
Figure2.4.3.: Content Sentiment Analysis of The Other Side of Silence.....	41
Figure 2.5.1.: Analysis of Newspaper Report.....	45
Figure 2.5.2.: Analysis of The Other Side of Silence .....	45
Figure 2.6.1.: Analysis of Newspaper Reports .....	49
Figure 2.6.2.: Analysis of The Other Side of Silence .....	50
Figure 2.7.1.: Bias Strength in Newspaper Reports.....	54
Figure 2.7.2.: Bias Strength in The Other Side of Silence.....	54
Figure 2.8.1.: Workflow of the Model .....	61
Figure2.8.2.: Analysis of Newspaper Reports .....	63
Figure 2.8.3.: Bias Analysis of The Other Side of Silence .....	63
Figure 3.1.4.1.: Sentence-level bias analysis of Midnight’s Furies .....	87
Figure 3.1.4.2.: word cloud of biased words in Midnight’s Furies.....	88
Figure 3.2.4.1.: Range of Linguistic Biases in The Other Side of Silence .....	121
Figure 3.2.4.2.: Wordcloud of Bias Inducing Words in The Other Side of Silence .....	122
Figure 4.1.10.1.: Sentence-level bias analysis of Chapters on Partition in Pakistani (Punjab Board )School Textbooks.....	161
Figure 4.1.10.2.: Sentence-level bias analysis of Chapters on Partition in Indian (NCERT) School Textbooks .....	161
Figure 4.1.10.3.: Word Cloud of chapters on partition in Punjab Board School Textbooks.....	163
Figure 4.1.10.3.: Word Cloud of chapters on partition in Punjab Board School Textbooks.....	163
Figure 4.2.4.1.: Bias Analysis of Tweets .....	188
Figure 4.2.4.2.: Sentence level Linguistic Bias Analysis of Tweets .....	189
Figure 4.2.4.3.: Tweet level Linguistic Bias Analysis .....	189
Figure 4.2.4.4: Word Cloud of Identified Bias Inducing Words .....	191
Figure 5.5.1.: Word Cloud of AI Now Reports .....	215
Figure 5.6.1.: Bias in AI Systems .....	217

## LIST OF TABLES

Table 2.8.1: Feature Set of the Tool .....	62
Table 4.2.3.1: Topic Categories of Analysed Tweets .....	185-187

# LIST OF ACRONYMS

AI	Artificial Intelligence
BASIL	Balanced Active Semi-supervised Learning
BERT	Bidirectional Encoder Representations from Transformer
DH	Digital Humanities
GloVe	Global Vectors for Word Representation
GRU	Gated Recurrent Unit
LIWC	Linguistic Inquiry and Word Count
LSTM	Long Short-Term Memory
NLP	Natural Language Processing
NPOV	Neutral Point of View
OCR	Optical Character Recognition
POS	Part-of-Speech
RFA	Reverse Feature Analysis
RMSProp	Root Mean Square Propagation
RNN	Recurrent Neural Network
WNC	Wiki Neutrality Corpus

# Chapter 1

## Introduction

*“Not enough to shed tears, to suffer anguish, not enough to nurse love in secret...Today, walk in the public square fettered in chains”*

*-- Faiz Ahmed Faiz,  
“Aaj Bazaar Me Pa Ba Jola Chalo” (“Let us walk in the market in shackles”)*

### 1.1. Intersection of History and Narrative: An Overview

Throughout human history, spanning epochs and civilizations, the intricate tapestry of our shared past has found a sanctuary within the realm of narratives. Despite the proliferation of literacy and the ubiquitous presence of media, the true essence of our collective narrative often remains veiled, emerging subtly through the unassuming discourse of everyday interactions. As aptly articulated by Grele (1975), it is within these informal exchanges that the vast majority of individuals forge their foundational understanding of historical epochs. Such conversational engagements serve as portals to the past, offering glimpses into the lives, struggles, and triumphs of those who have gone before us. As Roth explains, the concept of narrative explanation involves connecting events in a sequence to understand their outcomes (Roth, 1988). Through this narrative explanation, historiographers aim to uncover the objective truth of past events. This leads to a major challenge in the utilization of narratives—namely, the fundamental question regarding the existence of an ontological "true" history and the accompanying challenge of reconciling the multiplicity of individual narrative accounts of historical events. The crux of this dilemma lies in discerning the epistemological status of narratives: given their genesis within the realm of individual imagination, how might they aspire to encapsulate truth? Adding to the complexity is the unavoidable task of determining truthfulness among the numerous conflicting narratives and what distinguishes those narratives deemed "true" from those relegated to fiction or speculation.



Until recently, historians upheld the notion of a "Universal History," positing the existence of a singular, overarching narrative about humanity awaiting revelation (Mink, 1987). This concept of Universal history presupposed a uniformity in human nature across temporal and cultural contexts. However, by the nineteenth century, increasing recognition of cultural diversity and the rise of nationalism led to a decrease in the prominence of this idea. While this decline facilitated the acknowledgment of multiple historical narratives and diverse interpretations of shared events, the underlying notion of universal history persists. It manifests in the belief that the past unfolded in a specific manner, despite the inherent challenge of fully capturing objective reality. Thus, scholars persist in referring to "established facts" and "authentic depictions" of past events. The notion of verifiable accounts poses a challenge in reconciling the multitude of historical interpretations presented in diverse narratives. Historians opt to circumvent this challenge by outright dismissing the utility of narratives in historical discourse. Conversely, scholars such as William Cronon (1992) embrace a narrative-centric approach, despite its often-vexing nature. In his examination of historians' narratives Cronon observes that it is feasible "to narrate the same evidence in radically divergent manners." This observation implies a conceptualization of history as an ongoing contestation among competing narrative frameworks. However, he reassures both the reader and himself that his objective throughout "has been to acknowledge the immense power of narrative while still defending the past as a tangible entity to which our storytelling must somehow adhere, lest it deviate entirely from the realm of history" (Cronon, 1992). The rising question here is -What drives Cronon and other historians to employ narrative despite their reservations? There are at least three significant attributes that the narrative approach brings to historical inquiry:

- 1) it enables historians to glean insights into how individuals have construed their lives and the events surrounding them;

- 2) it serves as a framework for organizing and imbuing significance into specific historical contexts; and
- 3) it presents diverse perspectives on the human experience.

History is often perceived as a sequence of chronological facts, ranging from isolated occurrences to long-spanning processes. Each historical event unfolds within a complex network of interconnected events and ongoing processes. Here, Narratives serve as a means of structuring historical narratives. Historians not only draw upon others' narratives but also construct their own as a means of presenting history. As Cronon elucidates, a narrative plot can provide order and meaning to the overwhelming and disorderly chronology of reality (Cronon, 1992). Lastly, narratives constitute a substantial portion of the material comprising the new social history and history from below".

## **1.2. Historical narrative and linguistic biases**

In recent times, scholars such as Hayden White and F. R. Ankersmit have delved into the literary and subjective origins of historical interpretation. Their work suggests that while narratives can derive specific facts about the past from available evidence, the manner in which they imbue these facts with meaning, by delineating connections among them, is fundamentally shaped by the individual creative faculties of the writers, unconstrained by any specific cognitive parameters. White and Ankersmit contend that historical descriptions, interpretations, and explanations are inevitably influenced by the writer's personal inclinations and conceptualizations of past events. Consequently, they would view the notion of bias in historical accounts as obvious. To them, it is evident that historians' depictions of the past are colored by their individual interests and perspectives, rendering the idea of an objective standard for interpretation, against which accounts could be assessed for bias, untenable.

There exist four prevalent ways through which historical writing can exhibit bias. Initially, writers may misconstrue evidence,

thereby lacking justification in asserting the veracity of their drawn inferences regarding past occurrences. For instance, they may selectively acknowledge evidence suggesting how a certain event is transpired, while disregarding contradictory evidence indicating its impossibility. Subsequently, when writers construct accounts of historical subjects, their narratives may be credibly justified yet fail to encompass significant facets of the subject matter, resulting in an unbalanced or unjust portrayal. This bias, termed unfairness, might entail accentuating individuals' virtues while overlooking their vices, thus imparting an inequitable impression of their character. Thirdly, bias may manifest in overarching depictions of the past that insinuate factual assertions contradicting available evidence. Lastly, bias can emerge in the provision of causal explanations for historical events, wherein only select causes are delineated, leading to a misrepresentation of the event's causal genesis and conferring a deceptive understanding upon the reader.

Writers, when delineating individuals, societies, or events, often direct their attention towards specific facets that captivate their interest, such as an individual's demeanor, politically active factions within a society, or pivotal developments amid an event. However, to furnish an equitable depiction, it becomes imperative to encompass all salient attributes of the chosen aspect, thereby averting any semblance of distortion. Nevertheless, presenting an exhaustive portrayal runs the risk of incorporating trivial details, potentially encumbering the narrative with superfluous information. Consequently, writers frequently strive to deliver succinct yet comprehensive depictions as the act of description inherently entails selectivity, prompting scrutiny regarding its impartiality and objectivity.

The difficulty in defining historical subjects extends to concepts like character and political structure. Different perspectives, influenced by philosophical, cultural, or ideological backgrounds, shape a historiographer's interpretations. For instance, disparate perspectives on character encompass a spectrum of attributes, ranging from virtues

and vices to ethical precepts and adherence to customary norms. Similarly, notions of political structure differ, with some including personal, judicial, and military groups, while others focus solely on elected officials and legislative bodies. The selection of which groups to consider in a political structure portrayal further illustrates the subjective nature of historical narratives. Divergent viewpoints, influenced by philosophical, cultural, or ideological frameworks, mold historians' interpretations.

A description is considered deceptive if it offers a broad portrayal of the subject that conflicts with existing evidence or excludes significant causes or characteristics. Bias can arise when writers neglect crucial aspects of the subject due to preconceived notions or ideological inclinations. However, biases may also result from the restricted or skewed information accessible to the writers, resulting in incomplete and biased descriptions. The notion of bias encompasses more than deliberate distortion; it also encompasses unintentional misconceptions stemming from vague subject definitions. Here, readers may misconstrue the extent of historical investigations, leading to biased assessments of the impartiality and thoroughness of historical narratives. Interpretations within historical discourse entail diverse yet equally defensible representations of a given historical subject, characterized by variations in their overarching descriptions and colligatory terms. Writers select from these interpretations based on their capacity to impart significance to a wide array of facts and events, rendering them comprehensible. Nevertheless, interpretations may exhibit bias if they suggest erroneous facts about the subject or overlook critical causes or elements. Meanwhile, causal explanations in historical analysis strive to encompass all events that substantially shaped the unfolding of a specific historical occurrence. Biased explanations arise when historiographers favorably emphasize certain causes while disregarding others of comparable significance.

Despite the inherent difficulties, recognizing and rectifying biases within historical narratives holds significant importance.

Historians have taken steps to address biases by shedding light on marginalized perspectives, such as women's history or the colonial experience from the viewpoint of the colonized. Biases are also evident in national histories, where stereotypes shape perceptions of self and others. Cultural biases, deeply embedded within historical accounts, often go unnoticed but have far-reaching implications. Acknowledging and openly discussing these biases are crucial steps toward fostering a more nuanced and comprehensive understanding of history.

Bias in historical narratives is significant for several reasons. Firstly, biased historical accounts claim to offer an impartial view but present a skewed perspective that aligns with the writer's interests, resulting in unjust representations. Secondly, biased histories can perpetuate misunderstanding by attributing social phenomena to individual actions rather than systemic factors. Biased narratives may also overlook the perspectives and motivations of various groups involved, further distorting the understanding of historical events. Despite arguments that biased history can rectify historical injustices by highlighting marginalized experiences, there is a risk of oversimplification and demonization when focusing solely on the perspectives of oppressed groups, without considering the broader context and motivations of all parties involved. Ultimately, bias in narratives should be minimized to fulfill the writer's social responsibility of providing fair and accurate descriptions, interpretations, and explanations of past events. However, achieving this goal is challenging due to inherent biases, subjective interpretations, and cultural influences.

### **1.3. Bias Analysis**

Bias analysis encompasses various approaches aimed at assessing the reasoning, orientation, and extent of systematic errors, along with their origins, which impact measurements of associations in a text. Humanities research has an extensive tradition of investigating biases, initially focusing on media biases in the 1950s and gradually

expanding to encompass bias mitigation in research and historical documentation. Recent studies concentrate on diverse bias categories such as publication bias, communocentric bias, and gender bias, employing metrics like event selection, source choice, content placement, and word choice for identification and minimization. Methodologies like content analysis, frame analysis, and meta-analysis, are predominantly employed in recognizing and quantifying biases. Manual bias analysis is not scalable to the quantity and pertinence of the produced data and disparities like subjective interpretation, individual preferences and prejudices, the subtlety of the bias cues, and the difficulty of the sentence and word level identification often make manual mitigation of biases difficult.

Parallelly in computational linguistics and computer sciences, bias investigations originated with Lin et al.'s work in 2006. Various dimensions of media bias, such as perspective, ideology, truthfulness, and hyper-partisanship, have been explored using different computational methods, including low-level lexical information, linguistic cues, and deep learning techniques. However, there is a lack of consensus in defining media bias, with automated metrics often providing vague interpretations, and the literature review emphasizes the need to consult the comprehensive literature on bias analysis in humanities and social sciences for more effective automated tools. (A detailed discussion of methods for narrative bias analysis in both computer sciences and humanities can be found in the second chapter which discusses the methodology for this study.) Identifying this research gap, this research advocates for the integration of humanities and computational methodologies to factor in the limitations in the analysis of bias patterns in complex narratives, thereby developing more efficient and effective tools.

#### **1.4. Computational Tools for Text Analytics**

The history of computer-assisted studies in literature, emerging in the 1960s and 1970s, initially focused on leveraging electronic

media to identify patterns and strings within texts. his methodological approach empowered scholars to conduct a more exhaustive analysis of textual elements, surpassing traditional methodologies that relied on limited sample sizes. The introduction of word lists and concordances, initially in printed formats and later transitioned into digital formats, facilitated a more in-depth exploration of textual content. This transformation challenged the prevalent belief in literary criticism that often disregarded significant sections of text, as highlighted by J.F Burrows in his work "Computation to Criticism". Literary computing sought to remedy this limitation by offering insights into stylistic patterns and thematic elements, thus positioning itself within a broader discourse on methodology. Through the utilization of electronic tools, scholars acquired access to comprehensive datasets and could engage in comparative analyses spanning extensive literary collections. However, the efficacy of computer-based analyses relied on establishing connections between textual phenomena and discernible surface attributes, typically achieved through pattern-matching methodologies. Despite the potential offered by computer-assisted studies, persistent challenges remained. Certain scholars voiced concerns regarding the emphasis on quantitative methodologies, contending that intricate statistical calculations might not adequately capture nuanced textual nuances. Moreover, the perception of computer-based literary analysis as primarily centered on "word counting" impeded its recognition within mainstream scholarly circles. Nevertheless, numerous studies underscored the importance of integrating empirical data with critical interpretation, emphasizing the interdependent nature of human insight and computational analysis.

Over the years, advancements in both hardware and software have enabled more intricate analyses, empowering literary scholars to explore textual structures and intertextual connections in greater depth. Despite these strides, computer-assisted literary criticism has remained somewhat on the fringes of broader academic discourse. Traditional literary critics have often greeted it with skepticism, and the field has

grappled with the task of demonstrating the tangible benefits of digital tools in enriching interpretive methods. Computer-assisted studies have continued to evolve, embracing interdisciplinary approaches and actively engaging with theoretical discussions surrounding textuality and interpretation. Through the exploration of innovative methodologies, such as randomized text deformations, scholars have endeavored to uncover textual elements that transcend conventional human perception.

In the context provided, this thesis seeks to understand the linguistic biases prevalent in historical narratives further by analyzing the narratives surrounding the significant event of the 1947 Partition of the Indian Subcontinent. Its principal objective lies in employing advanced computational tools for text analysis to understand the manifestation of these biases, shedding light on how they have shaped and influenced the portrayal of Partition. By utilizing nuanced computational techniques and computational methodologies for close reading texts, this study attempts to uncover linguistic subtleties and implicit linguistic biases within historical accounts of the Partition. Through the analysis of textual data, including primary sources, textbooks, and social media discourse, the thesis aims to identify instances of linguistic bias, distortion, and selective interpretation that may have permeated historical narratives over time.

Furthermore, this research takes a critical stance towards the role of computational tools in mitigating biases within historical texts. While computational analysis offers unprecedented opportunities for uncovering and addressing biases, it also poses challenges and limitations. Therefore, this thesis will critically examine the efficacy and limitations of computational approaches in mitigating biases, considering factors such as algorithmic biases, data selection, and interpretative frameworks.



### 1.5. The 1947 Partition and its Scholarly Landscape

The partition of the Indian subcontinent in 1947 led to unprecedented human migration and widespread violence along the borders between India and the newly established nation of Pakistan, marking one of the most significant events in South Asian history in terms of population displacement and conflict. Partition is often described as “not a static event of the past, but an evolving moment in history.” (Singh et.al, 2016). Estimates suggest that over two million people lost their lives, and there were approximately 15 million refugees as a consequence of the Partition of the subcontinent. The extensive massacres that set off an unforeseen mass migration, which extended across much of North India, had Punjab at its epicenter. Only in Punjab, “(it) may be estimated that about five and half million people traveled each way across the new India-Pakistan border” (Nag 2001). In the introduction to *The Partition Omnibus* (2002) Partition historian Mushirul Hassan writes “Both as an event and memory, it [Partition] has to be interpreted and explained afresh in order to remove widely held misconceptions. This is both a challenge and a necessity and it is indeed a theme where the historian’s craft must be used deftly”.

Partition literature originates from authorized historical accounts in both India and Pakistan, where narratives either celebrated independence, downplayed Partition's disruptions, or attributed responsibility for the violence that occurred on the ‘other’. Early Partition historiography focused on the Declaration of Independence and the transfer of power from the colonizer. Indian nationalists see Partition as the outcome of divisive colonial policies eroding cultural unity. At the same time, Pakistani writers often viewed the establishment of a separate homeland as a response to the aspiration to preserve community values in the face of what they perceived as a tyrannical Hindu majority (Virdee, 2013). Early Historical studies have concentrated on the causes of Partition rather than its aftermath. Extensive exploration of its consequences has mostly been confined to narratives primarily aiming to assign responsibility for the associated

massacres. In this significant human event, the voices of individuals were notably muted. In “Partition and Post-Partition Acts of Fiction: Narrating Painful Histories” Sukeshi Kamra writes: “Partition’s erasure in Indian historiography has in no small way contributed to the sense many survivors have of a history that does not include or represent them.”

Partition literature is most appropriately situated within the developments in historiography and literary criticism. The disciplinary realm of history has only recently recognized the imperative for a social history of Partition, simultaneously, expanding literary criticism to accommodate discussions of traumatic literature. In contrast to these two disciplines, novelists, poets, and short story writers have extensively engaged with the profound human suffering entwined with the Partition. Numerous novels, short stories, and plays have chosen this theme, with a considerable portion serving as a cathartic outlet for writers to articulate their personal experiences. Writers such as Saadat Hasan Manto, Faiz Ahmed Faiz, Khushwant Singh, Allama Muhammad Iqbal Bapsi Sidhwa, Nasir Kazmi, Bhisham Sahni, Qurratulain Hyder Intizar Hussain, Amitav Gosh, Amrita Pritam, and Rajinder Singh Bedi have played a significant role in capturing the intricate human experiences, complexities, and personal narratives entwined with the Partition, thereby contributing significantly to the understanding of this historical period. As Alok Bhalla (1999) categorically notes:

The best of them (novelists)...do not repeat what the historians already know – that there was the violence of such fiendishness that each reminder of it still comes as a shock to our decencies and still violates our sense of a common humanity. Instead, they seek to make connections with the social and cultural life of a community in its entirety within a historically specifically period. (3120)

In the early 1980s, a transformative historiographical movement, led by scholars like Ranajit Guha, known as the Subaltern Studies School, offered an alternative perspective to history deviating from the prevailing mainstream nationalist narrative. By the 1990s, this approach influenced Partition Studies, shifting from a “great men of history” focus to a “history from below,” especially prompted by the regional studies' move.

The subaltern historiography envisioned the expansion of historical narratives by juxtaposing elite histories with marginalized voices (Roy, 2010). However, the omission of women's experiences during the Partition era was a notable lacuna that persisted within this framework as highlighted by Rituparna Roy (Roy, 19). This historiographical gap was subsequently filled by scholarly investigations into women's encounters with Partition, particularly gaining recognition in the literary contributions of feminist historiographers like Anis Kidwai, Kamla Bhasin, Urvashi Butalia, and Ritu Menon. The catalysts of this shift included chilling parallels between Partition violence and the 1984 anti-Sikh riots, and the Golden Jubilee of Independence, prompting introspective and reflective writing on previously inconspicuous subjects like Gender-based violence and violence against minorities. Recently, scholars like Ayesha Jalal and Ananya Jahanara Kabir, have examined the causative factors, repercussions, and lived experiences of Partition. Their works portray the integration of methodologies integrating oral history, memory theorization, and intersectionality which has significantly broadened the array of available evidentiary material. Urvashi Butalia's focus on oral histories and their impact on marginalized communities, alongside contributions from historians such as Mushirul Hasan and Gyanendra Pandey, have contributed to the expanded understanding of partition violence and its communal and political foregrounding. Gendered perspectives, as elucidated by Ritu Menon and Kamla Bhasin, have highlighted the experiences of women during this tumultuous period. These changes have marked a ‘Memory turn’ in

Partition studies extensively drawing upon the discourses and memory practices shaped by the historical experiences of the Holocaust and World Wars in the twentieth century. Scholars like Anita Inder Singh, Gyanesh Kudaisya, and Ian Talbot have studied the social, economic, and psychological implications of displacement and migration during Partition, focusing on themes of identity, belonging, and the construction of national narratives. Similarly, Yasmin Khan, Joya Chatterji, and Mushirul Hasan examine the concept of recollection, forgetting, and memorialization, tracing how memories of Partition are constructed, contested, and commemorated across generations. Narrative Histories like "Freedom at Midnight" by Larry Collins and Dominique Lapierre (1975), "The Great Partition: The Making of India and Pakistan" by Yasmin Khan (2007), "The Partition of India" by Ian Talbot (2009), "The Long Partition and the Making of Modern South Asia: Refugees, Boundaries, Histories" by Vazira Fazila-Yacoobali Zamindar (2007), "Remembering Partition: Violence, Nationalism and History in India" by Gyanendra Pandey (2001) and "The Partition of India" by H. V. Hodson (1969) provides vivid and detailed accounts of the events leading up to and following the partition. In contrary Oral Histories and Personal Accounts like "The Other Side of Silence: Voices from the Partition of India" by Urvashi Butalia (1998), "Train to Pakistan" by Khushwant Singh (1956), "The Shadows of Partition: Stories from a Divided India" edited by Jasbir Jain (2003), "Freedom at Midnight" by Larry Collins and Dominique Lapierre (1975), and "The Penguin Book of Partition Stories" edited by Saadat Hasan Manto (2011) provides lived experiences and individual perspectives of those directly impacted by this historic event. Works such as "Partition and Its Aftermath: Violence, Migration, and the Prospects for Reconstruction" edited by Mushirul Hasan (1997), "India after Gandhi: The History of the World's Largest Democracy" by Ramachandra Guha (2007), "Freedom, Trauma, Continuities: Northern India and Independence" by Chandrika Kaul (2008) and "The Partition of British India: A Study in Political Evolution" by A. H. Dani (1985) explored various aspects of the partition, including its causes, consequences, and

legacies, while also critically examining the role of key actors and factors contributing to the event. There were also legal analyses such as "The Punjab Boundary Force and the Problem of Order, 1947–48: A Revision" by Ian Copland (1990), "The Punjab Boundary Force, the Radcliffe Award, and Punjab's Administrative Integration" by Ian Copland (1991), "The Administration of Punjab since 1947" by Raj Kumar Gupta (1975) and "Judgment at Islamabad: The Role of the British Government in Partition, 1947" by Roger D. Long (1993) that analyzed the legislative and governmental frameworks that shaped the event and its aftermath.

“Ignited by the fear of impending disappearance of survivors of Partition 1947” (Roy, 2019) there has been an urgency to archive Partition in the past few decades which resulted in a growing interest in establishing memory archives of Partition in the last two decades. Consequently, there were digital and physical archival projects around the Partition such as ‘The 1947 Partition Memory Archive’, ‘The Partition Museum’ in Amritsar, ‘The Citizenship Archive of Pakistan’, and the ‘Indian Memory Project’. A recent observation in Partition scholarship centers around the concept of the "long Partition." Which Vazira Zamindar defines as ‘the unfinished post-colonial burden of this political Partition that unsettles ... [the] national closure given to Partition’s displacements.’”

The historiographical landscape of partition has evolved, moving beyond traditional narratives to embrace interdisciplinary approaches that consider the complex interplay of politics, society, and testimonies. More than seven decades post their attainment of independence, both India and Pakistan continue to witness the emergence of scholarly studies on Partition history. This sustained academic endeavor not only yields novel discoveries but also brings to light heretofore undocumented voices and narratives within the discourse.

The plethora of scholarly studies on the discursive analysis of partition literature has explored concepts ranging from linguistic loss, narrative structures, linguistic register across spatial axes, and the language of remembering. This thesis expands the scope of scholarly inquiries looking into the semantics and symbolism in partition literature by harnessing newly accessible digital tools to study the ‘language of bias’.

## **1.6. Linguistic Biases**

In the context of this thesis, linguistic bias can be defined as a consistent imbalance in the selection of words, reflecting the social perceptions attributed to the specified demographic or entity (Beukeboom and Burgers, 2017). Linguistic biases arise from and expedite the dissemination of essentialist perceptions regarding social categorizations. One prominent area of study regarding linguistic bias focuses on the labels employed to denote individuals or groups within society. Previous research in this field explores how language, particularly in the context of describing people and their actions, can subtly reveal stereotypical beliefs about the individuals and social categories being described. Epistemological bias pertains to linguistic expressions that influence the perceived credibility of a proposition. It manifests through the presentation of opinion as factual and factual information as opinionated. It also arises from expressing the author's personal stance on the subject matter in irrelevant textual contexts through linguistic choices that imply judgments on the truth or validity of the discussed content.

This thesis attempts to understand in detail the manifestation of the following 12 linguistic biases and their sub-categories in the select narratives.

### **1.6.1. Confirmation Bias**

Humans possess a tendency to disregard information that undermines their prior choices and judgments. This phenomenon, known as confirmation bias, is characterized by the pursuit or interpretation of

evidence in a manner that conforms to preconceived notions, anticipations, or hypotheses. Contemporary scholars regard confirmation bias as ubiquitous and one of the most well-established notions of inferential error in human reasoning. This concept is not new; early thinkers like Francis Bacon recognized similar patterns in human understanding. Bacon astutely observed that once an opinion is adopted, the human mind tends to draw all other evidence to support and reinforce it, disregarding or discounting contradictory evidence. This selective process serves to uphold the authority of preconceived conclusions, even in the face of contrary evidence. Indeed, the presence of confirmation bias becomes evident in everyday discourse, particularly in discussions of politics where individuals often exhibit a steadfast adherence to their own beliefs, despite evidence to the contrary.

#### **1.6.2. Misinformation Effect**

The phenomenon known as the misinformation effect pertains to the impairment of narrative memory consequent to exposure to misleading information. The impact of misinformation is profound, as the discordance in credibility between reliable content and misleading data can create a schism between accepted knowledge and public perception. Texts aiming to counter misinformation, through explicit refutation, may diminish belief in false information. However, when supplementary details are provided alongside misinformation, this additional information can enhance susceptibility to a greater degree than details contradicting the false information. For instance, in a study by Huff and Umanath, participants were tasked with reading fictional narratives and responding to questions containing both additive and contradictory misinformation statements. It was observed that queries pertaining to narratives accompanied by additive misinformation garnered higher endorsement rates compared to those with contradictory misinformation, despite participants being explicitly instructed to identify inaccuracies. Likewise, in a study by Pennycook et.al utilizing a false information framework, subjects persisted in

endorsing false information as true, even after these articles had been identified as disputed by fact-verifiers. Hence, offering explanations alongside false statements can bolster their credibility, while labeling misinformation as debatable proves less successful in reducing trustworthiness. Furthermore, the propagation of misinformation is exacerbated by people's eagerness to share information, particularly in cases where the veracity of the information remains unchallenged.

### **1.6.3. Framing Effect Bias**

Framing entails the deliberate selection of linguistic elements employed in presenting a subject matter, encompassing the choice of vocabulary, descriptors, and the associations drawn with individuals or occurrences (Entman, 1993). This encompasses the utilization of terms endorsing specific perspectives or accentuating aspects of the topic in a distinctly favorable or unfavorable manner. The concept originates from the framing effect bias, a cognitive predisposition where presenting a decision as either a loss or a gain influences the reader's decision (Tversky and Kahneman, 1974; Plous, 1993). A thorough comprehension of the framing surrounding a topic is essential for capturing its subtleties, as highlighted by the studies of Rein and Shoen (1996). This understanding enables individuals to delve deeper into the underlying perspectives and implications embedded within the discourse, thereby facilitating a more comprehensive grasp of the framing effect and thus reducing its impact. Nonetheless, detecting the subtle nuances in framing language poses a challenge, exacerbated by a dearth of established methodologies for its identification (Chong and Druckman, 2007).

### **1.6.4. False memory effect**

The concept of false memory encompasses instances wherein individuals recall and recount events in a manner divergent from their actual occurrence or, in extreme cases, conjure memories of events that never transpired. False memory bias manifests within a narrative through the distortion of perception, interference, and inferences



regarding the narrated event, stemming from the conflation or misattribution of familiarity pertaining to the recounted incident. This phenomenon underscores the intricate nature of memory formation and retrieval, wherein subjective interpretations and external influences can significantly alter one's recollection of past events. Furthermore, false memory bias underscores the susceptibility of human memory to various cognitive distortions, highlighting the complexities inherent in the construction of personal narratives.

#### **1.6.5. Political Bias**

Political bias entails the inclination or perceived inclination to skew or manipulate information in favor of a particular political position or candidate. Within political discourse, bias arises when individuals demonstrate an incapacity or disinclination to comprehend viewpoints that diverge from their own political leanings. There are varied types of political biases :

- a. **Concision bias** pertains to the practice of presenting perspectives succinctly, prioritizing brevity and efficiency over thoroughness, particularly in the communication of complex or unconventional views. This bias seeks to enhance communicative effectiveness by selectively emphasizing essential information while minimizing redundancy. In political discourse, concision bias may manifest as the omission of seemingly extraneous details, which can introduce bias depending on the subjective determination of what constitutes non-essential information. Consequently, political opinions are often reduced to simplistic party affiliations or belief systems, with potentially challenging or nuanced information omitted from consideration in the presentation
- b. **Coverage bias** occurs when a narrative allocates varying degrees of attention to different topics or issues. This discrepancy in attention can create the illusion of certain issues being more widespread or significant than others, thereby potentially influencing public perceptions of their importance or necessity.

- c. **False consensus bias** arises when individuals perceive their own opinions, beliefs, and values as more prevalent or widely accepted than they actually are. This bias becomes particularly pronounced within group settings, where the collective opinion of the group is erroneously assumed to reflect the broader population's views, often without significant challenge from other groups.
- d. **Speculative bias** refers to narratives that emphasize potential future outcomes, often employing speculative language such as "may," "what if," and "could," rather than focusing on concrete evidence or confirmed occurrences. When articles lack explicit labeling as opinion or analysis pieces, this can exacerbate speculative bias.
- e. **Gatekeeping bias** manifests through the selective inclusion, exclusion, or alteration of stories based on individualized ideological perspectives. This bias operates by individuals wielding influence over which stories are disseminated to the public, often filtering content to align with their own ideological leanings.

#### **1.6.6. Religious Bias**

Religious bias refers to the biased portrayal of religious beliefs, practices, or individuals within a discourse. This bias can manifest in various means, such as depicting certain religions or religious groups as inherently positive or negative, reinforcing stereotypes about specific religious practices, or marginalizing particular religious perspectives.

#### **1.6.7. Geographical bias**

Geographical bias refers to a partiality or inclination towards particular regions or geographic areas over others. This bias can also stem from the illusory sense of attachment or belonging to a particular geographic region having closer links to the concept of 'psychological dislocation'. This is where people may feel disconnected or dissatisfied with their current environment, leading

them to fantasize about belonging to a different place that they perceive as more desirable or aligned with their idealized self-image. This feeling of connection to a specific place can lead individuals to develop biases towards others who are perceived as outsiders or not belonging to the same geographical community.

#### **1.6.8. Hindsight bias**

The hindsight bias, or the "knew-it-all-along" effect, characterizes our inclination to retrospectively perceive an unpredictable event as easily foreseeable. Scholars have recently posited the existence of three distinct manifestations of hindsight bias: memory distortion, inevitability, and foreseeability

#### **1.6.9. Representation Heuristic Bias**

The representativeness heuristic bias occurs in a narrative when the narrator makes judgments regarding the likelihood of an event embodying the characteristics of a known prototypical event. This heuristic operates by evaluating an event in relation to a stereotype.

#### **1.6.10. Selection Bias**

Selection bias arises from the non-random selection of individuals, groups, or information, resulting in a failure to achieve proper randomization and thereby compromising the representativeness of the narrative in relation to the intended population.

#### **1.6.11. The Zeigarnik effect**

The Zeigarnik effect, attributed to the work of psychologist Bluma Zeigarnik, refers to the phenomenon wherein individuals demonstrate enhanced recall and narration of activities that have been interrupted. This effect posits that incomplete or interrupted tasks are better retained in memory compared to those that have been successfully completed.

### **1.6.12. Ego-Centric Bias**

Egocentric bias is characterized by individuals' propensity to perceive the world predominantly from their own subjective viewpoint, thereby overestimating the significance of their personal thoughts and experiences. This bias may prompt individuals to formulate judgments and decisions primarily influenced by their subjective experiences, rather than relying on objective evidence or considering alternative perspectives from others.

## **1.7. Research Objectives**

With the above background, this thesis has four primary objectives.

- To test and compare four media bias-detecting metrics to identify the gap and devise a target-oriented bias analysis tool to identify and mitigate biases in texts on the 1947 partition.
- To understand the multifaceted roles of historical texts, testimonies, school textbooks, and social media narratives in the construction and mediation of sustainable mnemonic hegemonies and the dissemination of knowledge pertaining to Partition.
- To study the linguistic patterns of narrative construction and bias manifestation in partition historiography pointing towards possible consequential social implications of such biases.
- Through a methodological probe, the study also looks into the applications, challenges, and ethical concerns in employing AI-integrated models for digital humanities research, especially in interpreting individual texts as opposed to patterns in a corpus

## **1.8. Overview of the Thesis**

The thesis comprises of six main chapters: an Introduction, four major chapters, and a Conclusion. The first major chapter, entitled "Computational Methods for Identifying and Categorizing Bias in Textual Data Across Various Granularity Levels," serves as the

methodological backbone of the study, detailing the approach taken to analyze linguistic bias within textual data. Chapter 3, titled "Historical Memory, Language, and Biases: Re-reading Midnight's Furies and The Other Side of Silence," examines Nisid Hajari's "Midnight's Furies" and Urvashi Butalia's "The Other Side of Silence" in two distinct sections. In Chapter 4, "'Master Narratives in (Re-En)Action': 1947 Partition in School Textbooks and Social Media Narratives," the focus shifts to the analysis of school textbooks and Tweets pertaining to the Partition of 1947, presented in two separate sections. The final major chapter, "AI Biases: A Digital Humanities Response," delves into the exploration of artificial intelligence (AI) both as a research method and a subject of study within the Humanities domain, offering insights into the complexities of addressing biases inherent in computational approaches.

### **1.8.1. Chapter 1: Introduction**

This chapter introduces the concept of linguistic biases and their types. It also summarises the recent and ongoing studies on Partition history and bias studies. The chapter serves both as a detailed overview of the background, literature review, research gaps, and primary objectives of the thesis and a roadmap for how this thesis moves between primary texts of partition, secondary textual criticism, computational tools and methodologies as well as the concepts of linguistic bias.

### **1.8.2. Chapter 2: Computational Methods for Identifying and Categorizing Bias in Textual Data Across Various Granularity Levels**

The second chapter traces the history of the methodological framework for bias analysis by outlining its major approaches in social sciences and computer sciences. We test and compare four exquisite media bias-detecting metrics namely,

1. Bias Determination using content sentiment analysis algorithm.

2. “Linguistic Models for Analyzing and Detecting Biased Language” (Recasens2013linguistic)
3. “Automatically Neutralizing Subjective Bias in Text” (Pryzant, 2020)
4. “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen, 2020)

in terms of accuracy, performance while using longer texts, analysis speed, labeling and inclusion of different types of biases, optimal embedding approach, effective utilization of defined approaches in humanities and social sciences, sentence-to-word and sentence-to-article level accuracy and efficiency on events not covered in the training data. Based on the inferences from the comparative analysis, the chapter briefs the architecture and workflow of the semi-automated target-oriented memory and narrative bias detection and classification tool for texts of the 1947 partition of the Indian subcontinent at different levels of granularity developed for this study.

### **1.8.3. Chapter 3: Historical Memory, Language, and Biases: Re-reading *Midnight’s Furies* and The Other Side of Silence**

The first section of this chapter studies Nisid Hajari’s *Midnight’s Furies: The Deadly Legacy of India’s Partition* (2015). The section briefs about the facilitation of historiography through archival study and the twofold manifestation of biases in Hajari’s work, primarily in the narrative pattern and secondarily arising out of the archiving and archival study. The linguistic bias analysis using the tool developed for the study underscores the susceptibility to bias through misinterpretation, omissions, citation failures, and misleading explanations in the work. The section also intrigues bias manifestation through the chain of causality of events and metaphors employed by the author. It accentuates the pitfalls in Hajari’s approach to history as a perennial conflict between two men who are central to the events.

Additionally, it delves into the subtle religious and gender biases in Hajari's historiography.

The second section studies Urvashi Butalia's *The Other Side of Silence: Voices from the Partition of India* (2000) by tracing the pathway of employing testimonies in Partition historiography and the associated challenges. Through computational analysis, the section seeks to identify and elucidate the manifestation and patterns of linguistic biases in Butalia's text. Additionally, the chapter aims to understand the repositioning of women in partition narratives delving into the attribution of violence and preservation of a communal 'other.' By highlighting observable biases such as egocentric, patriotic, gender, confirmation, and framing effects, the section underscores the fallibility of memory and challenges in its historical interpretation. Urvashi Butalia's narrative strategy and thematic organization within the book are also examined, shedding light on the complexities of writing social history and the author's subjective position as a narrator.

The chapter further examines the utilization of Holocaust rhetoric as a means to aptly encapsulate the politically tumultuous episodes of the Partition era and its subsequent aftermath and the linguistic patterns of bias manifestation and narration of violence.

#### **1.8.4. Chapter 4: 'Master Narratives in (Re-En)Action': 1947 Partition in School Textbooks and Social Media Narratives**

The first section of this chapter studies the Punjab Textbook Board Social Studies and Pakistan Studies materials from Grade Class IV to Class VIII and Class IX and X respectively and NCERT textbooks from Grade VI to XII using the Semi-Automated tool. The chapter traces the history of school textbooks in India and Pakistan and discusses how Partition is portrayed in them. Through an examination of linguistic patterns of biases and their manifestation, this section aims to uncover and elucidate the presence of positivity and negativity biases within the texts. Specifically, it seeks to understand how positive and negative language choices shape readers' perceptions of historical

events and individuals involved in the Partition. Additionally, it delves into the impact of passive voice construction and 'historical distance' in bias manifestation. The section attempts to provide insight into the multifaceted roles of textbooks as memory practices, boundary objects, and vehicles for reinforcing national ideology and homogeneity in understanding the Partition. Moreover, it aims to highlight the language of causality, othering, selective amnesia, and geopolitical implications embedded within Partition textbooks.

The second section of the chapter titled 'Recollection and Transmedial Reiteration: Twitter and the New Epistemology of knowledge about Partition' studies the nature of discourses and the manifestation of biases in Tweets about the 1947 Partition on 14 and 15 August 2021 around Partition Horrors Remembrance Day. The section through understanding the relationship between language use on Twitter and the manifestation of bias establishes the correlation between hostile comments and the inherent bias in original tweets. The section explores how political bias is reflected in tweets through entities, hashtags, and the use of sentiment-charged language, particularly emphasizing negative emotion words and group identity. Furthermore, the section conceptualizes Twitter as mediated sites of contention, mnemonic mediums, sites of false memory formation, and creators and manipulators of collective consciousness in recollecting and reiterating Partition. The study further highlights the transmedial structure of Twitter aiding the Recollection of the Past Within a Social Network's Collective Memory.

#### **1.8.5. Chapter 5: AI Biases: A Digital Humanities Response**

This chapter traces the distinguishable phases in the evolution of automated practices in Digital Humanities. It briefs how DH uses Artificial Intelligence (AI) in natural language processing, archiving, image processing, and text analytics. The chapter encompasses a critical examination of the transformative impact of AI on data curation



and analysis within the humanities, while also addressing the challenges and concerns that arise from its implementation. The chapter seeks to highlight how AI revolutionizes traditional approaches in the humanities, fostering collaboration and experimentation while challenging existing disciplinary boundaries. Furthermore, it underscores the importance of methodological rigor and critical engagement with AI systems to address concerns regarding transparency, reproducibility, and bias. Specifically, the chapter explores the infiltration of bias into AI systems at various stages, including data generation, dataset creation, algorithm composition, and evaluation. Through analysis of biases in AI-integrated offensive speech detection, the chapter provides insights into the complexities and limitations of AI systems in accurately identifying and addressing societal injustices. The chapter also highlights the challenges of a DH researcher working on AI and concludes by proposing a set of ethical guidelines for humanities researchers working on or with AI models.

#### **1.8.6. Chapter 6: Conclusion**

The concluding chapter summarises and reflects upon the analysis carried out in the four major chapters of the thesis. It begins by summarizing the methodology employed for linguistic bias analysis, emphasizing the integration of humanities and computational methodologies to analyze bias patterns within intricate narratives. The chapter highlights the development of a specialized tool for bias detection and classification, tailored to analyze texts related to the 1947 partition at various levels of granularity. Through close reading with computers, the thesis aims to bridge the gap between traditional literary-critical pursuits and digital humanities, enabling scholars to analyze texts at a deeper level and reveal fine-grained interpretations. It reflects on the critical analysis conducted in preceding chapters, where specific texts such as "Midnight's Furies" by Nisid Hajari and "The Other Side of Silence" by Urvashi Butalia were examined for biases in narrative patterns. Through this analysis, the chapter highlights the linguistic biases present in historical narratives and their

impact on societal perceptions and attitudes toward partition events. Furthermore, it discusses linguistic biases identified in school textbooks and social media narratives, emphasizing the role of language and discourse in shaping collective memory and identity. Building upon these insights, the conclusion chapter reflects on the broader implications of the research, including the potential for further interdisciplinary investigations into linguistic biases and the ethical considerations surrounding AI-driven research in the humanities. It underscores the importance of critical engagement with bias in historical narratives, offering insights into the complexities of language, memory, and representation in the context of partition literature. It further briefs the limitations and scope for further research.

## References

- Ahluwalia, M. (2017). *Divided by partition, united by resilience: 21 inspirational stories from 1947*. Rupa Publications.
- Ankersmit, F. R. (1998). Hayden white's appeal to the historians. *History and Theory*, 37(2), 182–193.
- Baumer, E. P., Polletta, F., Pierski, N., & Gay, G. K. (2017). A simple intervention to reduce framing effects in perceptions of global climate change. *Environmental Communication*, 11(3), 289–310.
- Beukeboom, C. J., & Burgers, C. F. (2017). Linguistic bias. In H. Giles, & J. Harwood (Eds.), *Oxford Encyclopedia of Communication* (pp. 1-19). Oxford University Press.
- Bhalla, A. (1999). Memory, History and Fictional Representations of the Partition. *Economic and Political Weekly*, 34(44), 3119–3128.
- Burrows, J. (1987). *Computation into criticism : a study of Jane Austen's novels and an experiment in method*. Retrieved from <http://ci.nii.ac.jp/ncid/BA00976007>
- Butalia, U. (1998). *The other side of silence: Voices from the partition of India*. Duke University Press.
- Chandra, B., Mukherjee, M., Mukherjee, A., Panikkar, K. N., & Mahajan, B. (1995). *Freedom, partition, and communalism*. National Book Trust.
- Chong, D., & Druckman, J. N. (2007). Framing theory. *Annual Review of Political Science*, 10, 103–126.
- Collins, L., & Lapierre, D. (1975). *Freedom at midnight*. Simon & Schuster.
- Cronon, W. (1992). A Place for Stories: Nature, History, and Narrative. *Journal of American History*, 78(4), 1347–1376. <https://doi.org/10.2307/2079346>
- Gamson, W. A., & Modigliani, A. (1989). Media discourse and public opinion on nuclear power: A constructionist approach. *American Journal of Sociology*, 95(1), 1–37.
- Ghosh, A. (1988). *The shadow lines*. Penguin Books.
- Grele, R. J. (1975). *Envelopes of Sound: Six Practitioners Discuss the Theory, Method, and Practice of Oral History*. Editor. Chicago: Precedent.

- Hasan, M. (Ed.). (1997). *Partition and its aftermath: Violence, migration, and the prospects for reconstruction*. Oxford University Press.
- Huff, M. J., & Umanath, S. (2018). Evaluating suggestibility to additive and contradictory misinformation following explicit error detection in younger and older adults. *Journal of Experimental Psychology: Applied*, 24(2), 180–195.
- Kamra, S. (2007). Partition and Post-Partition Acts of Fiction: Narrating Painful Histories. In A. G. Roy & N. Bhatia (Eds.), *Partitioned Lives: Narratives of Home, Displacement, and Resettlement*. Pearson.
- Khan, Y. (2007). *The great partition: The making of India and Pakistan*. Yale University Press.
- Mink, L. O. (1987). *Historical understanding*. Edited by B. Fay, E. O. Golob & R. T. Vann. Ithaca: Cornell University Press.
- Nag, S. (2001). Nationhood and Displacement in Indian Partition. *Economic and Political Weekly*, 36(51), 4753–4760.
- Pennycook, G., Cannon, T. D., & Rand, D. G. (2018). Prior exposure increases perceived accuracy of fake news. *SSRN Electronic Journal*.
- Plous, S. (1993). *The Psychology of Judgment and Decision Making*. McGraw-Hill Book Company.
- Rein, M., & Schon, D. (1996). Frame-critical policy analysis and frame-reflective policy practice. *Knowledge and Policy*, 9(1), 85–104.
- Roth, P. A. (1988). Narrative Explanations: The Case of History. *History and Theory*, 27(1), 1–13.  
<https://doi.org/10.2307/2504958>
- Roy, A. G. (2020). *Memories and Postmemories of the Partition of India*. Routledge.
- Roy, H., & Menon, V. (Eds.). (2006). *Partition of India: Legend and reality*. Oxford University Press.
- Roy, R. (2010). *South Asian Partition Fiction in English*. IIAS/Amsterdam University Press.
- Sengupta, D., & Bose, S. (Eds.). (2018). *The partition of Bengal: Fragile borders and new identities*. Cambridge University Press.
- Singh, A. I., & Hasan, M. (n.d.). *The Partition Omnibus: Contains Four Books*. Oxford University Press.

- Singh, A., Iyer, N., & Gairola, R. K. (2016). Introduction. In *Revisiting India's Partition* (pp. 15–35). Lexington Books.
- Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131.
- Virdee, P. (2013). Remembering partition: women, oral histories and the Partition of 1947. *Oral History*, 41(2), 49–62.
- White, H. V. (1982). The Politics of Historical Interpretation: Discipline and De-sublimation. *Critical Inquiry*, 9.
- White, H. V. (1991). Historical Emplotment and the Problem of Truth. In S. Friedlander (Ed.), *Probing the Limits of Representation: Nazism and the Final Solution* (pp. 37-53). Cambridge, Mass.
- White, H. V. (1992). Writing in the Middle-Voice. *Stanford Literary Review*, 9.
- Zamindar, V. F. (2007). *The long partition and the making of modern South Asia: Refugees, boundaries, histories*. Columbia University

## Chapter 2

### Computational Methods for Identifying and Categorising Bias in Textual Data Across Various Granularity Levels

*“Countless times since that day, a more experienced mechanic has pointed out to me something that was right in front of my face, but which I lacked the knowledge to see. It is an uncanny experience; the raw sensual data reaching my eye before and after are the same, but without the pertinent framework of meaning, the features in question are invisible. Once they have been pointed out, it seems impossible that I should not have seen them before”.*

*-Matthew B. Crawford, ‘Shop Class as Soulcraft’*

Linguistic bias analysis encompasses approaches that evaluate the rationale, sources, and repercussions of systematic errors and their sources that influence measures of associations in a text (Lash et al, 2016; Frampton et al, 2022). Understanding these biases is of high societal and policy relevance as it allows readers to critically and independently review and reason information. Linguistic bias analysis includes a broad range of methods, extending from sentiment analyses to in-depth probabilistic analyses involving substantial manual and computational labor. It is generally conducted by reviewing the data collection methods, selection and retention of data, inclusion, and exclusion of information, choice of words and labels, and placement of the content. A realistic model of bias analysis is presumptively complex, as bias is a relative term and is seldom defined. The influence of bias on shaping societal perception and formulating social consciousness has been explored extensively in humanities and social sciences, but the automated and semi-automated detection of bias in textual data has recently piqued interest of scholars in computer science and related fields of study.

This chapter traces the history of the methodological framework for linguistic bias analysis in textual data by outlining its major

approaches in social sciences and computer sciences. The researchers test and compare four media bias detecting metrics namely, “Determination of biasedness using content sentiment analysis algorithm”, Marta Recasen et.al's “Linguistic Models for Analyzing and Detecting Biased Language” (Recasens,2013), “Automatically Neutralizing Subjective Bias in Text” by Pryzant et.al. (Pryzant,2020}, and Chen et.al's model on “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen,2020analyzing). The chapter also briefs the architecture and functioning of the target-oriented bias analysis tool developed for the study to identify and analyze memory and narrative biases in texts on the 1947 partition of India and Pakistan.

This chapter is divided into four major sections where the first section critically studies the bias mitigation approaches in social sciences while section two outlines the approaches in computer sciences. Section three is a comparative study of four select automated bias analysis models and section four proposes a methodology for target-oriented bias analysis in texts longer than newspaper reports.

## **2.1. Approaches to Textual Bias Analysis<sup>i</sup> in Humanities and Social Sciences**

The study of biases in humanities and social sciences dates back to the 1950s, with a rich history primarily focused on media biases. Very recently there have been studies focusing on bias mitigation in research as well as in historical documentation. These studies focus on publication bias (Franco,2014) communocentric bias (Gordon,1990), and gender bias (Yip,2020) in texts. Topic selection, source choice, inclusion and exclusion of data, data labeling, lexical selection, content arrangement, allocation of space, spin, generalizability, and reliability are the key metrics defined for bias identification and minimization. Scholars from the social sciences mainly conduct content analysis (McCarthy,2008), frame analysis (Entman,1993), and meta-analysis (Hunter, 1986) to identify and quantify biases. Content analysis is to

identify and quantify bias by studying the patterns of communication and identifying and characterizing instances within the text. Researchers employ both quantitative and qualitative methods for content analysis. A qualitative content analysis attempts to identify all the subtle instances of biases that need human interpretation of the text while quantitative content analyses take into account the metadata for analysis. Whereas, in frame<sup>ii</sup> analysis, the researcher investigates how readers perceive textual information by focusing on the features of the information itself and the patterns of narration. scholars depend on the results of numerous previous researches to derive their conclusions in meta-analysis.

Pioneering researchers in bias analysis across the social sciences include Latanya Sweeney, whose groundbreaking investigations have centered on data privacy and discrimination within the realms of healthcare and criminal justice. Claudia von Vacano, focuses her research on uncovering biases prevalent in social inequality and political discourse. Gary King's significant contributions to quantitative methods have played a pivotal role in identifying biases within the domains of political science and public policy. Meanwhile, Susan Fiske's extensive work in social cognition has delved deeply into the analysis of biased language found in various forms of communication and media. Cynthia M. Frantz's research expertise lies in exploring biases embedded within language processing and cultural stereotypes. Similarly, Margaret L. Kern specializes in text analysis and sentiment expression, particularly within the context of social media and discourse surrounding mental health. Collectively, these scholars represent a diverse array of expertise aimed at comprehensively understanding and addressing biases inherent within textual data across various social science disciplines.

In the field of humanities, scholars have increasingly turned their attention to the analysis of biases present in textual data, particularly within the realm of digital humanities and critical studies. Posner (2016) investigates biases in digital humanities projects, highlighting issues of representation and inclusivity in data selection, digitization, and interpretation. Similarly, the edited volume by Liu and Svensson (2019)



explores biases from a decolonial perspective, challenging dominant narratives and advocating for methodologies that center marginalized voices. Krane's (2019) critical feminist technoscience study of sex testing in sports examines biases in textual data from scientific literature, policy documents, and media coverage, shedding light on power dynamics and discriminatory practices. Noble (2018) discusses biases in algorithmic culture, emphasizing how algorithms can perpetuate and amplify biases present in textual data, particularly within internet search engines. Seaver's curated reading list (2017) provides a comprehensive overview of scholarship on bias analysis and critical algorithm studies within the humanities, while Benjamin's (2019) exploration of race and technology interrogates biases in technological systems, including textual data, through a critical race theory lens. D'Ignazio and Klein's (2020) work on data feminism advocates for a feminist approach to data analysis and visualization, stressing the importance of reflexivity and intersectionality in addressing biases in textual data and computational systems. Collectively, these research underscores the interdisciplinary nature of bias analysis, integrating perspectives from critical theory, feminist studies, postcolonial studies, and digital humanities to critically examine biases in textual data.

In humanities and social sciences methodologies, manual bias analysis is not scalable to the quantity and pertinence of the produced data. Disparities like subjective interpretation, individual preferences and prejudices, the subtlety of the bias cues, and the difficulty of the sentence and word level identification often make manual mitigation of biases difficult. These fallibilities suggest the need for a hybrid model combining both qualitative and quantitative methodologies.

## **2.2. Bias Analysis Using Computational Metrics**

There are only a few systematic approaches for bias analysis in computational linguistics and computer sciences. Bias, especially specific to media has been a subject of computer-assisted study since Lin et al.'s research in 2006. (Lin,2006). Media bias has been

investigated in varied formats categorising it as perspective (Lin,2006), ideology (Iyyer,2014), truthfulness (Rashkin,2017), and hyper partisanship (kiesel2019). Early bias analysis metrics relied on low-level lexical information for bias prediction for example the research conducted by Greene and Resnik (Greene, 2009) employed "kill verbs" and "domain-specific verbs" to identify pro-Israeli or Palestinian perspectives in articles. Recasens et. al. (Recasens,2013) studied linguistic indicators to assess whether sentences from Wikipedia articles adhere to its neutral point of view policy. Besides this research started by the Natural Language Processing (NLP) community, researchers in journalism have also generated media bias mitigation approaches. While Gentzkow and Shapiro (Gentzkow,2010) used the preferences of phrases, Groseclose and Milyo (Groseclose, 2005) in their research utilized think-tank citations for bias mitigation. The adoption of deep learning methodologies for bias detection represents a recent trend. For instance, in the study conducted by Iyyer et al. (Iyyer, 2014), Recurrent Neural Networks were employed to determine the polarity of individual words, thereby facilitating the estimation of bias at the sentence level. Additionally, Gangula et al. (Gangula, 2018) devised a methodology leveraging headline attention mechanisms to categorize the bias present in articles. This approach understands biases by focusing only on key elements of textual content. Furthermore, Li and Goldwasser (Li, 2019) contributed to this field by introducing Convolutional Neural Network Graphs<sup>iii</sup>, which effectively encode social information to identify and analyze biases. By integrating social context into the analytical process, their work sheds light on the intricate relationship between language, society, and bias. A significant deviation in the detection of media bias lies in the varying levels of granularity, patterns, and features of the text under analysis. This spectrum ranges from examining individual tokens, as demonstrated by Fan (2019), to analyzing entire sentences, as explored by Aleksandrova (2019). Furthermore, this analysis extends to broader textual units such as articles, as evidenced by the work of Chen (2020), and encompasses the evaluation of specific information sources, as studied by Baly (2018).

Moreover, bias detection can also be directed towards understanding the biases inherent in the users themselves, as highlighted by Owen (1982). Though machine learning models on different levels facilitate the estimation of the manifestation and magnitude of biases at different levels of granularity, Lin et al. (Lin,2006) and Chen et al. (Chen,2020) are the only two studies that differentiate between word-level, sentence-level, and discourse-level bias detection.

Bolukbasi et al. (2016) conducted pioneering research on biases present in word embeddings, revealing how these representations could perpetuate societal biases. Zhao et al. (2018) extended this investigation to coreference resolution systems, uncovering gender biases and proposing debiasing methods to enhance system performance. Kiritchenko et al. (2018) delved into sentiment analysis systems, highlighting disparities based on the gender and race of text authors, thereby emphasizing the need for inclusive and fair computational models. Ethayarajh (2019) explored the use of euphemisms to attenuate biases in word embeddings, offering insights into alternative strategies for bias mitigation. Building upon these findings, Bolukbasi et al. (2016) proposed quantitative measures to identify and quantify biases in word embeddings, along with techniques to mitigate their effects. Gonen and Goldberg (2019) applied word embeddings to analyze gender bias in literary texts, showcasing the broader applicability of computational techniques in studying societal biases, particularly exemplified by their analysis of gender bias in the Harry Potter series. Additionally, Malmasi and Dras (2020) investigated gender biases in authorship attribution within computer science research papers, shedding light on disparities in collaboration patterns. Collectively, these research endeavors underscore the significance of identifying, quantifying, and mitigating biases in textual data to ensure fairness, equity, and inclusivity in computational systems and applications, particularly within the realm of NLP.

Many automated metrics adopt a broad perspective when addressing media bias, often defining it as “differences of news

coverage” (Park,2011) “diverse opinions” (Munson,2010), “different perspectives” (Hamborg,2020), or “topic diversity” (Munson,2009), leading to findings that lack decisiveness or depth. They “analyze media bias from the perspective of news consumers while neglecting both the established approaches and the comprehensive models that have already been developed in social sciences” (Hamborg,2020). For an efficient automated bias identification tool, the vast literature on bias analysis and bias categorization available in humanities and social sciences should be consulted. Former studies haven't also addressed the adequacy of these models in studying texts longer and more diverse than media reports.

### **2.3. A Comparative Study of Select Automated Media Bias Analysis Approaches**

This section tests and compares four recent bias analysis models that have proven accuracy in the test phase namely,

1. Bias Determination using content sentiment analysis algorithm.
2. “Linguistic
3. Models for Analyzing and Detecting Biased Language” (Recasens2013linguistic)
4. “Automatically Neutralizing Subjective Bias in Text” (Pryzant, 2020)
5. “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen, 2020)

in terms of accuracy, performance while using longer and complex texts, analysis speed, labeling and inclusion of different types of biases, optimal embedding approach, effective utilization of defined approaches in humanities and social sciences, sentence-to-word and sentence-to-article level accuracy and efficiency on events not included in the training dataset. The focus on these models is because they offer over 95.2 percent accuracy on the test dataset and 99.27 percent accuracy on the training dataset. The availability of the dataset and related resources, efficiency and recency of the preferred technology, utilization of recent

classification models, and positive feedback from the media houses are the other factors behind the selection of these models.

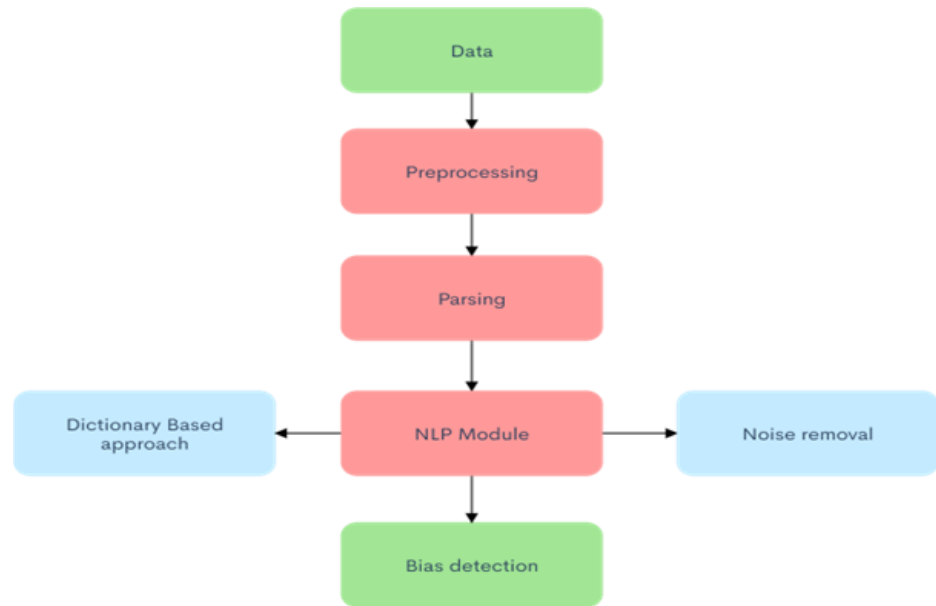
The dataset used for testing in this analysis is news reports published in The Economic Times, Hindustan Times, The Wire, Scroll, The Hindu, The Indian Express, The Print, and Zee News on the declaration of August 14th as Partition Remembrance Day. As the longer text the oral history narrative *The Other Side of Silence* (1998) by Urvashi Butalia is analyzed. Butalia documents stories across borders, of victims and perpetrators, men and women, families and villages, Dalits, and children as testimonials of Partition. We used the Tesseract Optical Character Recognition (OCR) engine to convert the book into machine-readable text data.

#### **2.4. Bias Determination Using Content Sentiment Analysis Algorithm.**

This model is designed to classify and quantify the polarity of sentences, which refers to whether a sentence expresses a positive, negative, or neutral sentiment. It leverages the Sentiment Analysis algorithm, a computational technique used to automatically determine the sentiment conveyed by a piece of text. Sentiment analysis has been widely employed by multiple researchers as a powerful tool to detect bias in textual data. By analyzing the emotional tone and subjective expressions within text, researchers can uncover implicit biases and subjective viewpoints present in the language, facilitating a deeper understanding of societal attitudes and perceptions encoded within textual datasets.

##### *Dataset and Experiment*

The training dataset for the model (Bharathi, 2019) analyzed in this chapter includes 2,550 newspaper reports extracted from 20 news websites in the United Kingdom, USA, and India. The architecture and workflow of the cited model is given in Figure 2.4.1.



*Figure 2.4.1: Architecture of the Sentiment Analysis Model*

The test data is pre-processed to remove noises and is parsed into individual sentences in the initial stage of analysis. Subsequently, individual sentences are forwarded to the Natural Language Processing (NLP) module, where the Part-of-Speech (POS) tagger assigns grammatical categories to each word, while a lexicon is employed to identify opinion words and their respective polarities.<sup>iv</sup>. Utilizing the polarity scale ranging from +1 to -1, the sentiment score is computed based on the emotional orientation or tone conveyed by the data. This score quantifies the overall sentiment expressed within the analyzed text, providing insight into whether the content is predominantly positive, negative, or neutral in nature. Further, depending on whether the sentence score is positive or negative, the sentence is categorized into biased, unbiased, and neutral. Employing a sentiment scoring algorithm, sentences are assessed based on a scale from 0.0 to 1.0, where scores falling within this range are categorized as positive and indicative of unbiased content. Conversely, sentences with scores ranging from -1.0 to 0.0 are classified as negative, suggesting a biased tone or perspective. In instances where the sentiment score equals 0.0, the sentence is designated as neutral, indicating a lack of discernible bias or emotional inclination.

## Results and Discussion

The test data set with newspaper reports on the declaration of Partition Horrors Remembrance Day was analyzed using the model and the results are summarised in figure 1.2. Out of the 50 sentences analyzed, Economic Times News reports had 27 biased, 16 unbiased, and 5 neutral sentences respectively. Hindustan Times had 20, 11, and 8 biased, unbiased, and neutral sentences respectively. The Wire had the lowest percentage of biased sentences with a range of 14 unbiased, 8 biased, and 5 neutral sentences. Out of the 94 sentences extracted from The Scroll 37 were unbiased, 41 were biased and 15 were neutral. The Hindu had 32 percent (25 sentences) of the analyzed sentences categorized as biased. The Indian Express being the biggest dataset showed 67 biased sentences making 41 percent of the dataset biased. The Print and Zee News had 10 and 4 biased sentences respectively. All the articles except the ones extracted from The Hindu and The Wire displayed 40 percent or above biased sentences according to the Sentiment analysis algorithm. The Economic Times articles displayed the highest percentage of biased sentences, and The Hindu displayed the highest number of neutral sentences. 15 sentences weren't processed by the model mostly because only stop words were available for analysis. In the experimental study, the longer complex text, *The Other Side of Silence*, with 5653 pre-processed sentences, showed 2011 unbiased sentences, 1712 biased sentences, and 1768 neutral sentences. The model failed to analyze 162 sentences, because of the presence of regional language words from Hindi, Punjabi, and Sindhi. Figures 2.4.2 and 2.4.33 show the chapter-wise summary of the analysis. The model showed a higher accuracy with newspaper reports when compared to the longer text as *The Other Side of Silence* had significantly large data that contained subjective responses.

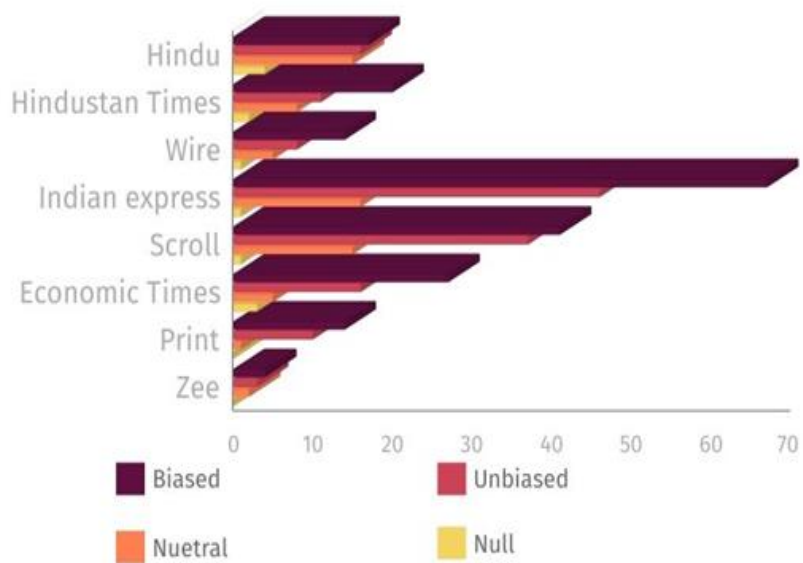


Figure 2.4.2.: Content Sentiment Analysis of Newspaper Reports

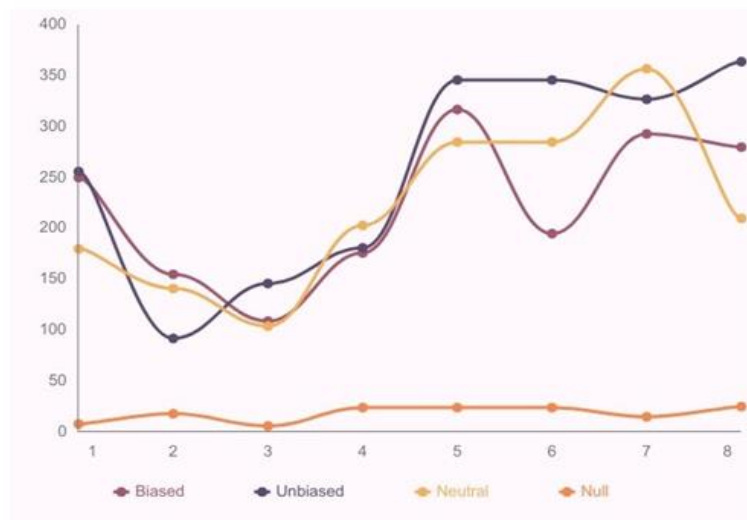


Figure2.4.3.: Content Sentiment Analysis of The Other Side of Silence

The model had difficulty in finding subjective sentiments and properly analyzing them for their intended tone. This resulted in a higher number of false positives and negatives. The text used multilingual words, and the model failed to analyze these, resulting in a higher number of null sentences. With both long and short texts, in conjugations



of words, the mid-polarity was left out. The biggest setbacks that resulted in a huge number of false positives were because of the model's inadequacy in interpreting sarcasm, idioms, negations, and comparative sentences. The model also doesn't propose an effective way to decide the sentence sentiment score when the words have multi-polarities<sup>v</sup>.

## **2.5. “Linguistic Models for Analyzing and Detecting Biased Language” (Recasens,2013)**

This model is capable of discerning two distinct types of bias: framing effect bias and epistemological bias. Framing effect bias refers to the phenomenon where language is employed in a manner that favors or promotes a particular perspective. This bias often manifests through the use of praising language or words that are specific to a particular viewpoint. For example, the inclusion of positive adjectives or descriptors that portray one side of an argument in a favorable light while overlooking or downplaying alternative viewpoints can contribute to framing effect bias. On the other hand, epistemological bias occurs when presupposed propositions are accepted as true without undergoing critical examination or encountering conflicting evidence. This type of bias can lead to the uncritical acceptance of certain beliefs or ideas, potentially influencing the interpretation and presentation of information. Epistemological bias may arise when assumptions or assertions are presented as established facts without adequate justification or consideration of opposing viewpoints.

Recasens et al. identifies common linguistic cues associated with framing effect bias and epistemological bias. These cues encompass a variety of language features, such as factive verbs<sup>vi</sup>, subjective intensifiers, implicatives, and hedges<sup>vii</sup>. To identify these cues, they employ a logistic regression model<sup>viii</sup> trained on a feature vector<sup>ix</sup>. The training dataset is extracted from the Neutral Point of View (NPOV) corpus of Wikipedia articles. The Neutral Point of View (NPOV) corpus of Wikipedia articles refers to a collection of articles from Wikipedia that have been curated or annotated to adhere to the principle of neutral

point of view. The NPOV principle is a core tenet of Wikipedia, requiring that articles present information objectively, without bias, and without favoring any particular perspective. The corpus typically consists of articles that have been reviewed and edited to ensure that they accurately represent a balanced view of the subject matter, presenting multiple viewpoints, if applicable, while avoiding advocacy or opinionated language. Researchers may use the NPOV corpus for various purposes, such as studying language use, analyzing bias, or developing computational models for natural language processing tasks. Within this model, words that induce bias are categorized as positive, while all other words are classified as a negative category.

#### *Dataset and Experiment Setting*

Recasens et al. assembled the NPOV corpus, comprising “7,464 Wikipedia articles” categorized under the "NPOV disputes category"<sup>x</sup>. These articles were specifically selected due to their non-compliance with Wikipedia's neutral point of view (NPOV) policy, indicating the presence of biased content or contentious viewpoints. They used Stanford Core NLP tools to tokenize and split the text in the corpus into sentences. The logistic regression model was trained using the NPOV dataset, where each word in the NPOV sentences from the training set was assigned to either the positive class, representing bias-inducing words, or the negative class, encompassing all other words. This approach involved constructing a feature vector for every word, with the aim of discerning patterns and associations between specific linguistic features and bias categories. The types of features such as word, lemma, Parts of Speech (POS), position in sentence, hedge, factive verb, assertive verb, implicative verb, report verb, strong subjective, weak subjective, entailment, polarity, positive word, negative word, grammatical relation, bias lexicon, and collaborative features are used in the logistic regression model individually and also in relation to the context.

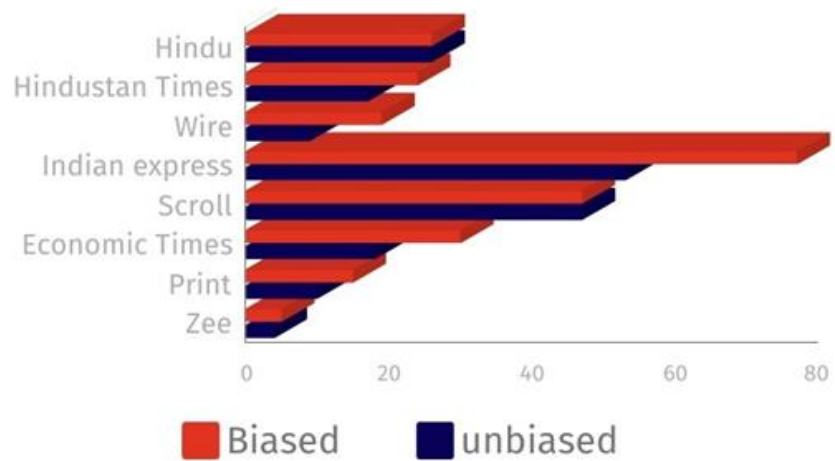
The bias detection task was carried out in three phases:

1. finding the biased phrase
2. identifying the bias-inducing word
3. rewording

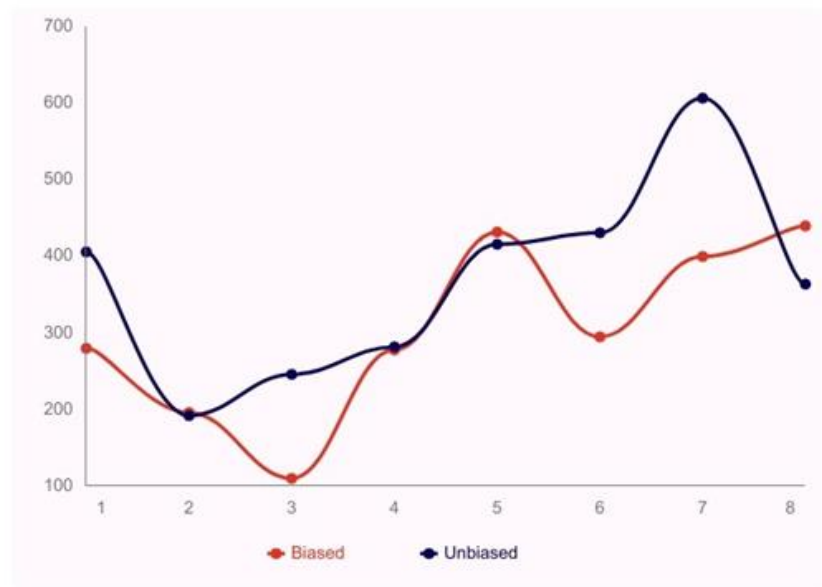
### *Results and Discussion*

Though the NPOV dataset mentioned in the work proposes to include NPOV policy-based edits, it contains a lot of insignificant edits carrying out spelling corrections, hyperlink additions, and rectification of grammatical errors. The corpus contained sentences accompanied by links and titles, necessitating extensive cleaning procedures to extract meaningful and coherent sentences. Subsequently, Stanford's CoreNLP software was employed to perform various linguistic analyses, including tokenization, sentence splitting, lemmatization, tagging, and dependency parsing. These analyses facilitated the extraction of linguistic features, which were then used to build the model. Specifically, the features were derived from the linguistic tags assigned to each word during the preprocessing stage. The model leveraged a linear classifier to associate these features with individual words, enabling the identification of patterns indicative of biased language usage within the textual data. It was observed that modeling as a sequence labeling problem would have provided more contextual information.

Figures 2.5.1 and 2.5.2 consolidate the results of the first phase of analysis attempting to identify the biased phrases for select Newspaper reports and “The Other Side of Silence” respectively. The linear model displayed poor performance in the longer text when compared to that of shorter newspaper reports. The sentences with multilingual words were omitted from this analysis and thus only 5450 sentences were studied using the model.



*Figure 2.5.1: Analysis of Newspaper Reports*



*Figure 2.5.2: Analysis of The Other Side of Silence*

Though the model successfully proposes a heuristic to build a labeled corpus with biased sentences, identifying the bias-driven edits based on the comments the editors left at commit<sup>xi</sup> is hectic considering the nature of the comments. The model fails at the identification of bias-inducing words as well as the replacement of them. This limitation primarily arises from the inadequacy of the corpus used and the

emphasis on detecting bias-inducing words in isolation. This approach presupposes that bias in a sentence can stem solely from individual words, overlooking the potential contribution of linguistic cues and resources. In contrast, the work of Bruce et al. (bruce,2002) introduces a more comprehensive framework that leverages rich signals and resources, including subjectively biased words or phrases. A significant portion, specifically 62 percent, of the replacements made during the removal module were deemed inaccurate. This inaccuracy stemmed from the fact that many of the corrections focused on altering terms that the model considered factual rather than biased. For instance, the term "journalist" was replaced with "businessman", indicating a misunderstanding or misinterpretation of the context or intended meaning of the original text. Such factual corrections failed to address the underlying bias within the content and instead resulted in inaccuracies or distortions of the information conveyed. Only unspecified places and events, relative time references, expressions of doubt, such as supposed, apparent, unsupported attributions as Some people say, and contentious labels like perverted, myth, controversial, and bigot were removed and replaced.

## **2.6. “Automatically Neutralizing Subjective Bias in Text” (Pryzant,2020)**

This model focuses on automatically reducing inappropriate subjectivity (“subjective bias”) bias in text. Subjective bias occurs when neutral textual language becomes deviant consciously or unconsciously because of the feelings, opinions, and/or dispositions of the writer. In practice, this model identifies subjective bias by employing Recasens et.al's model (recasens,2013), using Wikipedia’s neutral point of view (NPOV) policy. The model attempts to mitigate bias and neutralize texts by proposing edits aimed at removing biased language. In contrast, earlier research focused on debiasing textual representations through the removal of various dimensions of distortions present in word embeddings. This approach, as demonstrated by studies conducted by Bolukbasi et al. (2016), Gonen et al. (2019), and Ethayarajh et al. (2019),

involves identifying and addressing biases encoded within the semantic representations of words. By eliminating these biases from word embeddings, researchers aimed to enhance the fairness and neutrality of natural language processing models. However, the model under consideration adopts a different strategy, targeting biased language directly within the text itself. Pryzant et.al., introduce the Wiki Neutrality Corpus (WNC), a parallel corpus of biased language extracted from text-neutralizing editorial suggestions from Wikipedia editors. The corpus has “180,000 biased and neutralized sentence pairs along with contextual sentence information and metadata”. The approach put forth by this model introduces a methodology that involves human intervention in the process of bias analysis, as well as in the broader domain of generative language modeling. This is achieved through the utilization of sequence-to-sequence algorithms that are specifically designed to enable interactions between human annotators and the model, thereby facilitating a collaborative approach to bias detection and mitigation. By integrating human expertise with machine learning techniques, this approach aims to enhance the accuracy and effectiveness of bias analysis while also advancing the capabilities of generative language modeling in capturing and representing diverse linguistic nuances and contexts.

#### *Dataset and Experiment Setting*

The dataset used for training this model is the “Wiki Neutrality Corpus with aligned sentences pre- and post-neutralization by English Wikipedia editors”. This is an extension of the NPOV corpus developed by Recasens et al. This model crawled 423823 Wikipedia revisions suggested by the editors during the period between 2004 and 2019. The model has two modules

1. Detection module
2. Editing module

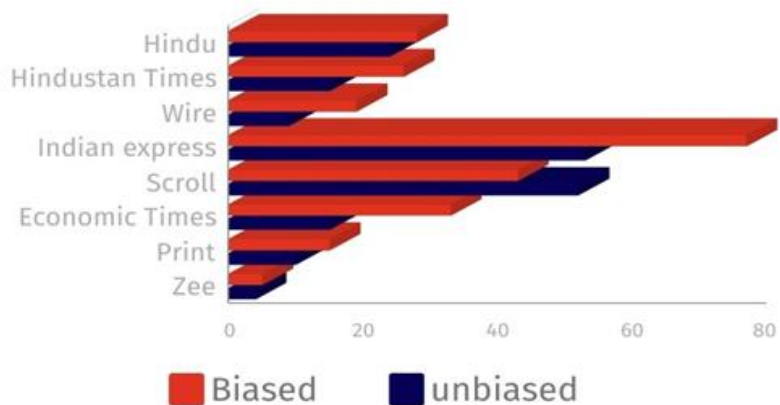
Which utilizes Bidirectional Encoder Representations from Transformer (BERT) based detection and Long Short-Term Memory (LSTM) based

editing. BERT is a pre-trained natural language processing model developed by Google AI. It utilizes the Transformer architecture, which enables it to capture contextual information from both left and right contexts of a word, unlike previous models that only consider one direction. BERT is trained on large corpora of text data using unsupervised learning techniques, allowing it to learn deep contextual representations of words and sentences. These pre-trained representations can then be fine-tuned on specific downstream tasks, such as text classification, question answering, or named entity recognition, achieving state-of-the-art performance across various natural language processing tasks. BERT has had a significant impact on the field of natural language processing and has been widely adopted in both research and industry applications. LSTM based editing refers to a method or technique that utilizes Long Short-Term Memory networks, a type of Recurrent Neural Network (RNN), for editing tasks in natural language processing (NLP). LSTM networks are designed to capture long-range dependencies and remember information over extended sequences, making them particularly effective for tasks involving sequential data like text. In editing tasks, LSTM-based approaches can be used to generate or revise text, correct grammatical errors, improve fluency, or enhance readability. By leveraging the memory and contextual understanding capabilities of LSTM networks, these methods can produce high-quality edits that preserve the coherence and meaning of the original text while addressing specific editing objectives. The identification component within this model employs a neural sequence tagger to assess the likelihood of each word in the input being imbued with subjective bias. Subsequently, within the editing phase, the system selects the most likely subjective source sentence and undergoes training to generate edits that substitute biased words with neutral alternatives while preserving the intended meaning of the sentence. This process ensures that the resulting output maintains fidelity to the original message while mitigating any potential bias present in the initial text. Though the detection module was effective, this analysis failed to

remodel the editing module as it wasn't open-sourced and there were technical limitations in training the model.

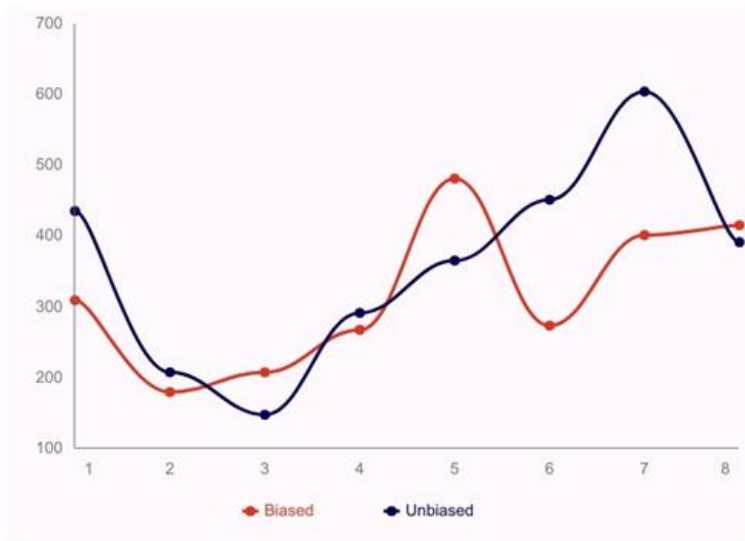
### *Results and Discussion*

There weren't any whopping differences in the results of bias analysis of newspaper reports when comparing the effectiveness of Recasens et.al's model and Pryzant et.al's, model. The MODULAR algorithm <sup>xiii</sup> improves human control and interpretability. In employing the model for bias analysis most of the errors were due to the subtlety and complexity in processing and reasoning language by the model rather than in the process of generation of a fluent text. Figure 2.6.1 shows the consolidated results of the bias analysis of the select newspaper reports.



*Figure 2.6.1: Analysis of Newspaper Reports*





*Figure 2.6.2: Analysis of The Other Side of Silence*

Compared to the Recasens et al model, this analysis provides a higher accuracy in terms of longer texts as shown in Figure 2.6.2. The wider and better-annotated database, human control over the analysis, and the effective usage of a BERT-based classifier are reasons for this augmented performance.

While this model hasn't shown perfect accuracy, it possesses the capability to offer valuable recommendations for mitigating subjective bias in textual data, particularly in contexts involving substantial political content. Its utility lies in providing suggestions aimed at enhancing the objectivity of such texts, thereby contributing to a more balanced discourse. It also provides immense possibilities for a wider data set creation and advancement.

## **2.7. “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen, 2020)**

This model is an extension of Chen et.al's work titled “Detecting Media Bias in News Articles using Gaussian Bias Distributions” (Chen, 2020) which in turn draws inspiration from techniques for discerning specific and overarching sentiments in sentiment analysis (Wachsmuth,

2015). The study titled "Detecting Media Bias in News Articles using Gaussian Bias Distributions" delves into the augmentation of the detection module's efficacy through the incorporation of supplementary data and metadata pertaining to biased sentences within a given text. "Gaussian Bias Distributions" typically refer to bias distributions that follow a Gaussian or normal distribution pattern. In statistical terms, a Gaussian distribution is characterized by a bell-shaped curve, where the majority of data points cluster around the mean, with fewer data points located further away from the mean. In the context of bias distributions, Gaussian bias distributions may imply that biases in a system or dataset are distributed according to this pattern. This means that most biases are centered around a certain value, with fewer biases exhibiting extreme values. Chen et.al's research explores how factors beyond the textual content itself, such as secondary information and accompanying metadata, contribute to enhancing the model's ability to identify instances of bias within news articles. By considering a broader spectrum of contextual cues, the model seeks to refine its detection capabilities. In particular, Chen et.al, "utilize the probability distributions of the frequency, positions, and sequential order of lexical and informational sentence-level bias in a Gaussian Mixture Model" (Chen, 2020). In the model utilized for this study, Chen et al. employ a neural framework aimed at identifying bias tendencies across various levels of text granularity, ranging from individual words to the broader discourse within articles, within the bias assessment module. This neural model is specifically designed to analyze and detect biases inherent in texts, enabling a multi-dimensional understanding of bias manifestation throughout the entirety of the textual content.

#### *Dataset and Experiment Setting*

The database is an expansion of Chen et al.'s previous corpus (Chen, 2018), extracting news articles categorized by topics and labels specific to political bias. This enhanced corpus within the model comprises "6964 news articles", incorporating the most recent updates sourced from websites dedicated to analyzing media bias, such as

"allsides.com" and "adfontesmedia.com". Every sentence within the corpus undergoes labeling for its topic, political bias, and perceived unfairness, aligning with the fairness criteria established by "adfontesmedia.com".

Using the outlined corpus, a "Recurrent Neural Network (RNN) architecture" was developed to acquire classification insights for the bias identification module, employing "pre-trained word embeddings from GloVe" (Pennington, 2014). A Recurrent Neural Network (RNN) is a type of artificial neural network designed to handle sequential data by capturing temporal dependencies. Unlike traditional feedforward neural networks, which process each input independently, RNNs have connections that form directed cycles, allowing them to maintain a state or memory of previous inputs as they process new ones. The key feature of RNNs is their ability to retain information from past inputs, making them well-suited for tasks involving sequences, such as time series prediction, language modeling, speech recognition, and natural language processing. An RNN class will work effectively in capturing semantic information in a text at multiple granularities according to previous studies. GloVe, which stands for Global Vectors for Word Representation, is a word embedding technique used to represent words as dense vector representations in a continuous vector space. Developed by Stanford researchers, GloVe aims to capture semantic relationships between words by leveraging global co-occurrence statistics from large text corpora. The focus of GloVe is to learn word embeddings by modeling the probability of word co-occurrences in a corpus. It constructs a co-occurrence matrix that counts how often each word appears in the context of other words within a fixed window size. By factorizing this matrix, GloVe learns vector representations for words that encode semantic similarities and relationships based on their co-occurrence patterns. One of the advantages of GloVe is its ability to produce word embeddings that capture both syntactic and semantic information. An in-depth reverse feature analysis (RFA) <sup>xiii</sup> was also used to disseminate biases from the article level to other levels of

granularity. while most biased sentences were correlated with Linguistic Inquiry and Word Count (LIWC) categories <sup>xiv</sup>at the word level (Pennebaker et al., 2015), the parts of an article that are most unfair are revealed in the sentence and paragraph levels. At the discourse level, the sequential patterns of media bias were identified.

### *Results and Discussion*

This model is highly significant as it looks biased in an article and discourse level beginning from word and sentence. This helps in a better understanding of the reasons behind the bias manifestations. This analysis proves that the model can process high-level language information including semantic and lexical information which is significant to bias prediction. The study confirms that narrative biases are mostly manifested in larger granularity levels such as paragraph or discourse level than in word or sentence level, significantly the final paragraphs of the analyzed newspaper articles had the most number of biased sentences compared to the beginning and middle paragraphs. The classifier GloVe is capable of detecting topic-independent bias clues in the text. In the individual classes of bias, RNN could predict the absence of bias more precisely than its presence in a sentence.

The sentence level bias analysis gives almost similar results to Pryzant's model. Below figures 2.7.1 and 2.7.2 summarize the article-level bias strength. For a better understanding, the longer text is visualized in chapters. The graphs show the Patterns and types of bias (political, Unfairness, and non-objectivity). The values undergo normalization to achieve a mean of zero and a standard deviation of one. Positive values signify a heightened degree of bias, whereas negative values suggest either a lower level of bias or impartiality within the text. This normalization process ensures consistency in the representation of bias levels, allowing for more accurate analysis and interpretation of the text's inherent biases.

Though there wasn't much difference in the results when analyzed in the longer text the biggest shortcoming was that the model

has taken into account only three types of biases. The text “The Other Side of Silence” displayed higher unfairness and non-objectivity biases as expected owing to the peculiarities of an oral history text.

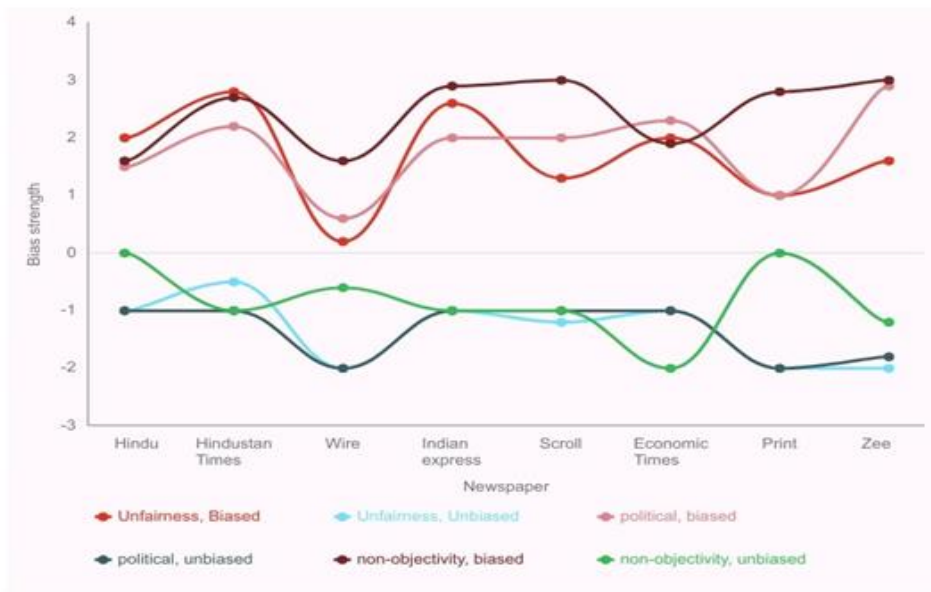


Figure 2.7.1: Bias Strength in Newspaper Reports

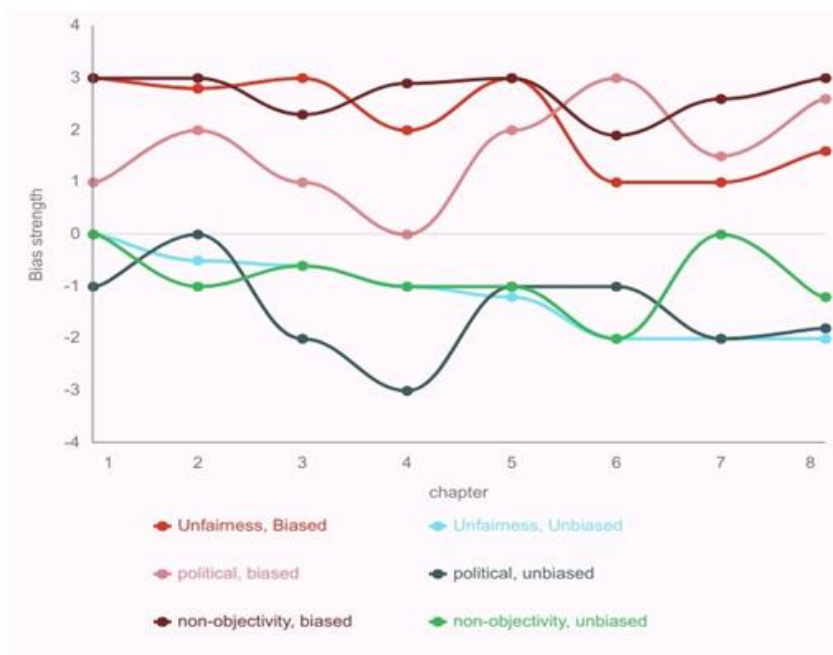


Figure 2.7.2: Bias Strength in The Other Side of Silence

The above graphs show that newspaper articles as well as chapters in the longer text ‘The Other Side of Silence’ start in a neutral tone (value closer to mean) and the biasedness of the text becomes prominent towards the final paragraphs of the text. Among the three bias types that are studied using this model unfairness bias displays the highest strength in the analyzed texts. While the choice of words dissimilar from the rest of the text is an indicator of unfairness bias, positive and negative expressions in describing facts are the common indicators of political bias. Despite of its few limitations, this model shows a larger prospect to be developed as a bias mitigation tool with an extended database and bias categorization strategy.

This study is limited to four models to evaluate the proposed bias detection and mitigation modules. Though bias detection is of a high research significance the unavailability of an adequate annotated database is the biggest challenge in achieving the necessary accuracy. All the available datasets are inadequate and most of them are not sufficiently annotated to accommodate varied forms of biases. Though as part of the study, we have also experimented with the Balanced Active Semi-supervised Learning (BASIL) dataset (Fan, 2019) which has 300 sets of articles from FOX, New York Times, and HPO, it wasn’t used further because it wasn’t large enough to accommodate the bias categories and language complexities especially in “The Other Side of Silence”. Pryzant et.al’s extension of the NPOV corpus is a promising consolidation, which can be further elaborated. There are huge databases developed for bias studies (for example Hamborg et, al. that contain “1,6m articles gathered from almost 4,000 publishers from over 100 countries”) which are not open-sourced yet. There is also a crucial question of the generalizability of the assorted data. The lack of generalizability often results in a high accuracy difference between the analysis of test data and training data. The larger size of the dataset often didn't guarantee diversity of data and representation in the case of the used dataset, which affects reusability. Thus, developing a generalizable,

adequate, properly annotated dataset should be the primary concern in the field of automated and semi-automated bias studies.

### *Word Embedding*

This study conducted experiments utilizing two-word embedding tools: Google's word2vec, recognized for its adeptness in implementing both the continuous bag-of-words and skip-gram architectures, enabling the computation of vector representations for words.; and Stanford's GloVe, An algorithm for unsupervised learning crafted to generate vector representations for words. Despite rigorous experimentation, the results did not reveal any embedding and model combination that significantly outperformed the others. This suggests that no single embedding-tool and model pairing exhibited superiority over the rest, underscoring the importance of further exploration and refinement in this domain. It is inferred that all the word embedding tools used for the analysis provide almost similar accuracy and their accuracy in combination with the machine learning models used for this analysis is also similar.

### *Accuracy*

Amongst the experimented models, “Analyzing Political Bias and Unfairness in News Articles at Different Levels of Granularity” (Chen, 2020) provides the highest accuracy in the detection module, and “Automatically Neutralizing Subjective Bias in Text” (Pryzant, 2020) provides the highest accuracy in the removal module. A comparative assessment of accuracy isn't highly effective because all the models in the analysis attempt to mitigate different types of textual biases. All the models gave less accurate results compared to the training and validation data used. The replacement module proved ineffective as a majority of the edits impacted the coherence and fluidity of the original text, occasionally introducing new biases.

Current methods for detecting bias are often adapted from simpler text classification tasks, lacking the complexity needed to effectively analyze articles comprehensively. These approaches

primarily focus on capturing basic lexical features, rendering them inadequate when applied at the article level, especially for topics not present in the training data. This limitation arises from the subtlety and infrequency of bias indicators, particularly those unrelated to specific events. Consequently, solely modeling low-level information fails to sufficiently identify bias at the article level. Chen et.al proposes an effective workbench to look into biases in word, sentence, paragraph, and discourse levels.

#### *Utilization of bias detection metrics in Humanities*

All the studied models failed to utilize the available bias detection metrics in the humanities. As Felix Hamborg writes:

One important reason for the comparably poor performance of the technically superior computer science methods for the automatic identification of instances of media bias is that such approaches currently tend to not make full use of the knowledge and expertise on this topic from the social sciences (Hamborg, 2019automated).

Another common error is approaching the removal module without classifying the type of bias. The detection module is fallible because of the exclusion of most of the bias types emphasized in humanities scholarship during the annotation of the dataset. Although this study does not reveal any superiority among embedding techniques, it underscores the efficacy of certain machine learning models for detection while exposing limitations in the removal modules employed. Despite this, it provides a springboard for future research.

#### **1.7.Semi-Automated Target-oriented Identification of Select Linguistic Biases at Different Levels of Granularity**

Based on the inferences from the comparative analysis, we propose a semi-automated target-oriented linguistic bias detection and classification tool for texts at different levels of granularity ranging from words, sentences, and paragraphs to the entire text. The tool identifies,



categorizes, and suggests the manifestation across twelve bias categories namely, confirmation bias, Misinformation effect, framing effect bias, false memory effect, political bias and its six sub- categories, political bias, geographical bias, hindsight bias, representativeness heuristic, selection bias, The Zrigarnik effect and Ego-centric bias as described in the introductory chapter.

### *Dataset*

Owing to the technical and infrastructural limitations we developed a diverse and extendable target-oriented dataset for this study. The Wikipedia articles on partition under the NPOV dispute category were extracted assuming that there would be a high density of ‘bias-driven edits’ (Recasens, 2013) considering the platform’s serious engagement with Partition and related content. But there were edits on spelling errors, language simplification, precision, and vandalism in the NPOV dispute category. Thus, the study took into account the comments of the editors to select the sentences from Wikipedia for the corpus. Further open-sourced articles from the 1947 Archive<sup>xv</sup>, Newspaper articles, and editorials from *The Hindu*, *The Indian Express*, *Hindustan Times*, *The Times of India*, and *The Telegraph* were included. The corpus also contains selected chapters from Maulana Abdul Kalam Azad's “*India Wins Freedom*” (Azad, 2021), Ram Manohar Lohia's “*Guilty Men of India's Partition*” (Lohia,2020), Barney White Spinner's “*Partition: The Story of Indian Independence and the Creation of Pakistan*” (Spinner, 2020), Urvashi Butalia's “*The Other Side of Silence*” (Butalia, 2017), Yasmin Khan's “*The Great Partition: The Making of India and Pakistan*” (Khan, 2017), Aanchal Malhotra's “*Remnants of Seperation: A ghara and a gaz: From Lahore to Amritsar to Delhi*” (Malhotra,2016), Nisid Hajari's “*Midnight's Furies: The deadly legacy of India's partition*” (Hajari, 2015), Ayesha Jalal's “*Pity of Partition*” (Jalal, 2013), Kavita Daiya's “*Violent Belongings: Partition, gender, and national culture in postcolonial India*” (Daiya, 2011) and Sarila's “*The Shadow of the Great Game: The Untold Stories of India's Partition*” (Sarila, 2017). The corpus had 6,850 sentences

from the listed sources. Along with the NPOV edits the annotations for these sentences were also gathered by crowd-sourcing them as micro tasks to volunteers. The annotators were between the age group of twenty to fifty-five, from academia<sup>xvi</sup>. All the annotators were located in India and had English as their second language. 63 percent of them were women and were from diverse political and religious backgrounds. Each annotator was given fifty randomly shuffled sentences and they were briefed about the motif of this study and given general instructions about their task. The task was explained to them in detail with examples and the annotators were asked to highlight the sentence and/or phrase and/or word that induced the specified type of bias (if there). They were also asked to classify the sentence with the tags biased, unbiased, and neutral with a tag of the bias category. The annotations were confirmed only when half of the annotators responded similarly.

#### *Architecture and Workflow*

For the detection module, this study adapts the methods from the metrics of Chen et.al (Chen,2020; Chen,2020; Chen,2018) and Pryzant et.al (Pryzant, 2020). RNN classifiers were employed for the classification tasks, where each cell is a Gated Recurrent Unit (GRU) of a recurrent rate size of 32. A Gated Recurrent Unit (GRU) is a type of recurrent neural network (RNN) architecture that addresses some of the limitations of traditional RNNs, such as the vanishing gradient problem and difficulty in capturing long-term dependencies in sequential data. The GRU in this model introduces gating mechanisms that control the flow of information within the network, allowing it to selectively update and forget information over time. The key components of a GRU include:

1. Update Gate: This gate controls how much of the past information should be passed along to the current time step. It combines information from the previous time step with the current input to generate an update gate value between 0 and 1, indicating how much of the past information to retain.

2. Reset Gate: This gate determines how much of the past information should be forgotten or reset. It combines information from the previous time step with the current input to generate a reset gate value between 0 and 1, indicating how much of the past information to forget.

3. Hidden State: The hidden state of the GRU represents the current memory or state of the network. It is updated based on the input, previous hidden state, and the update gate, allowing the GRU to capture temporal dependencies and learn representations of sequential data.

By incorporating these gating mechanisms, the model attempts to effectively capture long-range dependencies in sequential data while mitigating the vanishing gradient problem. In this study, we utilize the GloVe pre-trained word embedding model, featuring a dimensionality of 50 for word embeddings. Additionally, we employ the Adam optimizer with a learning rate set at 0.001. The Adam optimizer is chosen as it is an adaptive learning rate optimization algorithm that combines techniques from both momentum-based optimization and RMSProp (Root Mean Square Propagation) optimization. The features of Adam such as:

1. Adaptive Learning Rates: Adam adjusts the learning rates for each parameter in the neural network based on estimates of the first moment (mean) and second moment (uncentered variance) of the gradients. This adaptive learning rate helps in converging faster and more reliably compared to fixed learning rates.

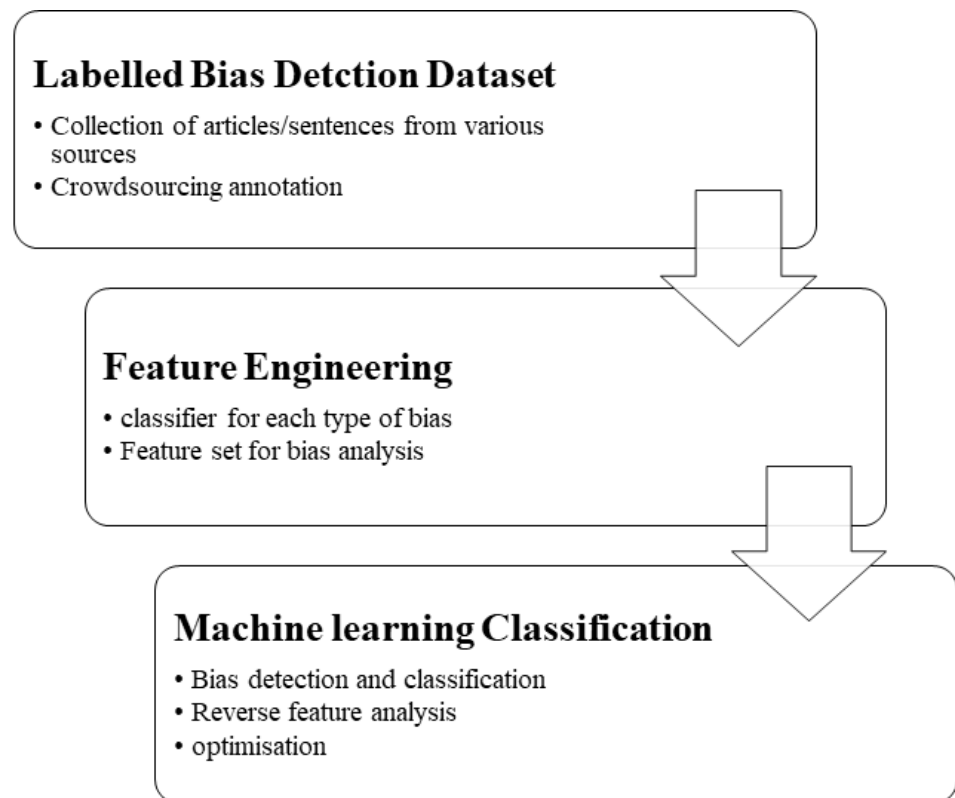
2. Bias Correction: Adam incorporates bias correction to compensate for the initial bias of the estimates of the first and second moments. This correction is particularly important during the early stages of training when the estimates may be inaccurate.

3. Momentum: Adam utilizes momentum similar to other optimization algorithms, which helps accelerate the optimization process by accumulating gradients from past time steps.

4. Regularization: Adam includes L2 weight decay regularization to prevent overfitting by penalizing large parameter values.

Was significant for the optimization of the developed model.

Figure 2.8.1 consolidates the architecture and workflow of the developed model in labeled bias detection dataset generation, feature engineering, and machine learning classification phases.



*Figure 2.8.1: Workflow of the Model*

The dataset was divided into two groups to be used in the development and training set. In reverse feature analysis, the model uses the developed classifiers to ascertain the probability of bias within test articles. The process of Reverse Feature Analysis for this model involved the following steps:

1. Training the model using all features in the dataset.
2. Evaluating the model's performance on a validation set and cross-validation.

3. Removing one feature from the dataset and retraining the model.
4. Evaluating the model's performance again.

If the model's performance significantly deteriorated after removing a feature, it was inferred that the feature was important for prediction. If the performance remained relatively unchanged, the feature was considered less important. Steps 3 and 4 were repeated for each feature in the dataset. Based on the estimation of bias strength at the word level and sentence level experiments were designed. The study also captures useful collective information and features from the context. The table 2.8.1 shows the set of features used in this approach and their description

Feature	Description
POSTag	Providing insight into the syntactic function of the word and its dependencies of words in the text
Linguistic features of the word	Whether it is implicative, assertive, and/ or hedge
Polarity	the polarity of the analyzed word according to Riloff and Wiebe
LIWC features	Drawing from psychological and psychometric analyses encompassing structural, cognitive, and emotional aspects of language.

*Table 2.8.1: Feature Set of the Tool*

### *Results and Discussion*

Figures 2.8.2 and 2.8.3 consolidate the results of the model for newspaper reports and *The Other Side of Silence* in the test phase. At the sentence level, the model analyses context-based features and bias strength. Whereas, at the discourse and article level the study is conducted completely manually as the approach based on the ‘inverted pyramid’ (Po, 2003) structure of an article, doesn't suit the structural complexity of texts analyzed in this study.

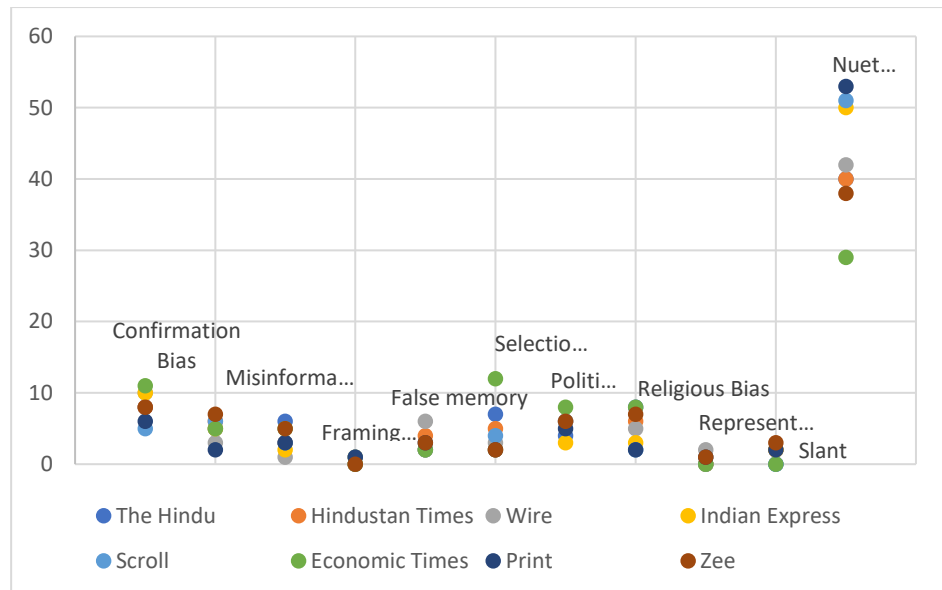


Figure2.8.2: Analysis of Newspaper Reports

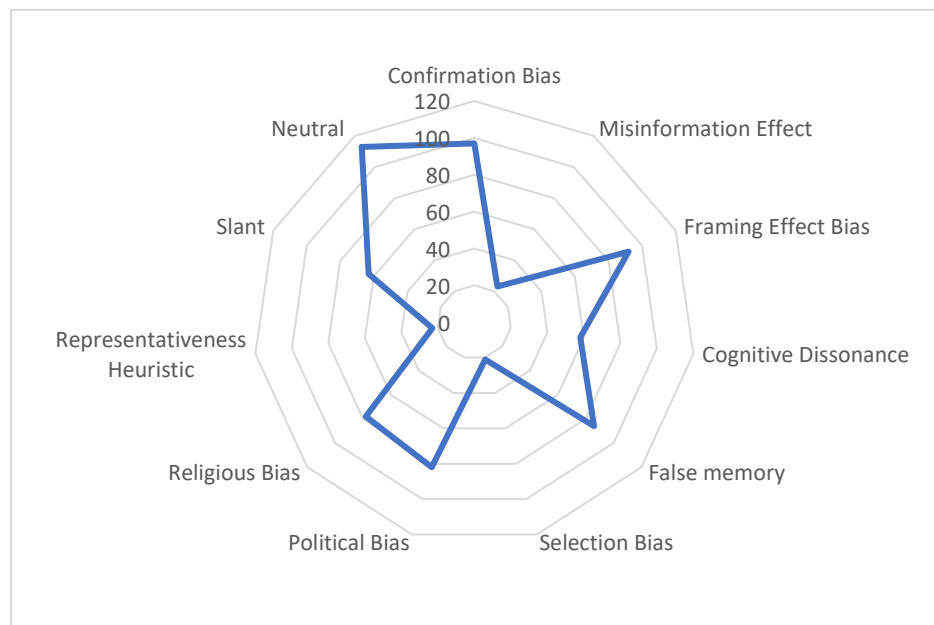


Figure 2.8.3.: Bias Analysis of The Other Side of Silence

All classifiers outperform the majority baseline, achieving an accuracy of 75.60% for political bias, 83.42% for religious bias, 85.6% on selection bias, 88% on cognitive dissonance and representative heuristic bias, 75.42% on slant, 86% on confirmation bias and misinformation effect, 83% on framing effect bias and 93.45% on neutral. This illustrates the proficiency of the classifiers in identifying linguistic characteristics that transcend specific topics. Upon closer examination of the distribution of classes, it becomes evident that while

RNN performs well in predicting the absence of bias, it faces challenges in accurately detecting its presence across various bias classes. This observation underscores the need for further refinement and exploration in leveraging RNN for bias detection tasks, particularly in capturing expressions of bias within different contexts.

The dataset used for this study, though it is extendable is target-oriented in its present form which affects its adaptability. The dataset cannot be termed as fully diverse as user-generated sites like Wikipedia have their structural factors including moderation practices that make them less accessible for marginalized communities. This dataset was curated between the period of 2020-2021. There was a lot of development and shift of frames in the data during and post this period. The model has to be retrained continuously and the training data should also be updated dynamically, which is not feasible. Thus, there should be sustainable curating practices to find appropriate data and also effective strategies to capture reframing and techniques to evaluate and rule out encoding bias.

The developed tool is subsequently employed for linguistic bias analysis in chapters 3 and 4, while chapter 5 provides an in-depth examination of the tool's workflow and the methodological framework utilized throughout this thesis to apply the tool effectively. This comprehensive analysis serves to elucidate the practical implementation and theoretical underpinnings of the tool, enhancing the understanding of its utility in linguistic bias assessment within the context of the study.

## **2.8. 'Close Reading with Computers'**

In employing the above-discussed tool for the computational analysis of linguistic biases this thesis employs the methodological framework of 'close reading with computers'. Computational literary studies as a field have emerged through 'distant reading' combining big data and the rhetorical force of numbers. Considering the prevalent scholarly perception of digital and quantitative methodologies in literary analysis in digital humanities, microlevel digital analyses of texts have

been obtruded from the mainstream. But “The processes of iteration, repetition, and quantitative analysis that are made possible by computational methods have an analogy not just in the telescope but also in another optical instrument: the microscope” (Eve, 2019). The developed tool studies the linguistic biases in partition texts using the framework of close reading with computers.

This study defines close reading as the practice of scrutinizing the language used in a literary work or a specific excerpt thereof to view the minute or the unseen. “Close reading need not involve detailed interpretation of literary passages (though there is plenty of that around in close reading, especially when the texts in question are difficult to understand),” it is about “attention to how meaning is produced or conveyed, to what sorts of literary and rhetorical strategies and techniques are deployed to achieve what the reader takes to be the effects of the work or passage” (Culler, 2011). In most cases, close reading focuses on linguistic details to have fresh optics on literary interpretation. Very few scholars have employed close reading in digital humanities. Most of those studies in fact are bridges between close reading and distant reading. Martin Paul Eve in his work “Close Reading with Computers” (2019) focuses on David Mitchell’s “Cloud Atlas” (2004) and employs a set of computational methods to examine different formal aspects of the novel. scholars, Ichiro Kumagai, Miyuki Yamada, and Yuichi Murai have focused on the visualization of linguistic features of Charles Dickens’s *A Christmas Carol* (1843). Another example of digital close reading would be Stefanie Posavec’s “visualization of Jack Kerouac’s *On the Road*” (1957) in 2008, and the “Novel Views project” that examined “*Les Misérables*”. This study follows Martin Paul Eve’s “*Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell’s Cloud Atlas*” and the “back to the text” argument in Alan Liu’s, “The State of the Digital Humanities: A Report and a Critique,” and Tanya E. Clement’s, “Text Analysis, Data Mining, and Visualizations in Literary Scholarship,” to close read of



select 1947 Partition texts using computational tools. The study is also a methodological probe to answer the following questions

- “What can the computer see, in its repetitive and unwavering attention, that was less (or even in-) visible to a human reader?” (Eve, 2008)
- How computational methods can aid in gathering evidence and interpreting individual texts as opposed to patterns in a corpus?
- Can propounding close reading projects balance out the necro reading ratio in digital humanities and insist closer focus on individual texts?

## Notes

<sup>i</sup>Used by researchers as an overarching term to describe all types of biases including but not limited to cognitive, textual, media, and communocentric biases

<sup>ii</sup> “A frame is a selection and/or emphasis on a specific part or event in the text”

<sup>iii</sup> Convolutional Neural Network (CNN) graphs depict the architecture of CNNs, illustrating the arrangement of layers and their connections. They provide a visual representation of how data flows through the network during forward propagation, aiding in understanding the network's structure and facilitating optimization and debugging.

<sup>iv</sup> “Polarity is a float which lies in the range of  $[-1,1]$  where 1 means positive statement and -1 means a negative statement”

<sup>v</sup> Multi-polarities in sentiment analysis involve categorizing text into multiple sentiment categories beyond just positive or negative

<sup>vii</sup> hedges are linguistic devices used to mitigate the strength or certainty of a statement, expressing a degree of uncertainty, doubt, or politeness.

<sup>viii</sup> A logistic regression model is a statistical method used for binary classification tasks, where the dependent variable is categorical with two possible outcomes (e.g., yes/no, true/false). It estimates the probability that a given input belongs to one of the two categories by fitting a logistic function to the observed data, which transforms the output of a linear regression into a probability value between 0 and 1.

<sup>ix</sup> A feature vector is a numerical representation of an object or entity in a machine learning or data analysis context. It consists of a set of features or attributes that describe the characteristics of the object, typically represented as a vector (i.e., an ordered list of numbers).

<sup>x</sup> “[http://en.wikipedia.org/wiki/Category:All\\_NPOV\\_disputes](http://en.wikipedia.org/wiki/Category:All_NPOV_disputes)”

<sup>xi</sup> operation which sends the latest changes of the source code to the repository

<sup>xii</sup> The MODULAR algorithm is a clustering algorithm used for community detection in complex networks. It is based on the principle of modularity, which measures the strength of division of a network into communities.

<sup>xiii</sup> Reverse Feature Analysis (RFA) is a technique used in machine learning and feature selection to identify the most relevant features in a dataset that contribute to the prediction of a target variable. Unlike traditional feature selection methods that aim to select a subset of features based on their individual predictive power, RFA works by iteratively removing features from the dataset and measuring the impact on the model's performance.

<sup>xiv</sup> Linguistic Inquiry and Word Count (LIWC) categories refer to a set of linguistic dimensions or features used to analyze and quantify the language used in text. Developed by James W. Pennebaker and colleagues, LIWC is a text analysis software program that categorizes words into various psychological and linguistic dimensions based on their semantic properties.

Some common LIWC categories include:

- a. Emotion: This category includes words related to different emotions such as joy, sadness, anger, and fear. It helps analyze the emotional tone of the text.
- b. Cognitive processes: This category covers words related to cognitive processes such as thinking, understanding, and perceiving. It provides insights into the cognitive aspects of the text.
- c. Social processes: This category includes words related to social interactions, relationships, and communication. It helps analyze the social dynamics conveyed in the text.
- d. Personal concerns: This category covers words related to personal experiences, needs, and desires. It provides insights into the individual's concerns and motivations expressed in the text.

- e. Language style: This category includes words related to linguistic style, such as pronouns, articles, and prepositions. It helps analyze the linguistic characteristics of the text, such as formality and complexity.
- f. Psychological processes: This category covers words related to psychological states, such as certainty, tentativeness, and discrepancy. It provides insights into the psychological processes underlying the text.

<sup>xv</sup> “<https://www.1947Partitionarchive.org>”

<sup>xvi</sup> postgraduate and Doctoral students and Professors from the English Departments of Bharathiar University, Coimbatore; Christ University, Bangalore; Devagiri College, Calicut; IIT Kharagpur; IIT Madras; IIT Gandhinagar; Karunya University, Coimbatore; Kerela University; Mar Ivanious College, Trivandrum; Mercy College, Palakkad; NIT Trichy; Pazhashi Raja College, Pulpally; Providence College, Calicut; Providence College, Ooty; PSG College of Arts and Science, Coimbatore; St. Mary’s College, Sulthan Bathery; University of Calicut, Kozhikode and VIT, Vellore.

## References

- 1947 archive. (2020). Retrieved from <https://www.1947Partitionarchive.org/>
- Alan Liu. (2012). The state of the digital humanities: A report and a critique. *Arts and Humanities in Higher Education*, 11(1–2), 8–41. <https://doi.org/10.1177/1474022211427364>
- Aleksandrova, D., Lareau, F., & Ménard, P. A. (2019). Multilingual sentence-level bias detection in Wikipedia. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing* (pp. 42-51).
- Alessio, D., & Allen, M. (Year). Media bias in presidential elections: A meta-analysis. *Journal of Communication*, 50(4), 133–156.
- Azad, A. K. (2021). *India wins freedom*. Prabhat Prakashan.
- Baly, R., Karadzhov, G., Alexandrov, D., et al. (2018). Predicting factuality of reporting and bias of news media sources. arXiv preprint arXiv:1810.01765
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.
- Bharathi, S. V., & Geetha Angelina. (2019). Determination of news biasedness using content sentiment analysis algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, 882-889.
- Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems* 29.
- Bolukbasi, T., Chang, K., Zou, J. Y., et al. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29.
- Bruce, R., Bell, M., & Martin, M. (2002). *Learning subjective language*. Computer Science Technical Report, University of Pittsburgh.
- Butalia, U. (2017). *The other side of silence: Voices from the Partition of India*. Penguin UK.
- Carder, B., & Rooney, J. (2018). Thinking, fast and slow. *The Journal for Quality and Participation*, 41(2), 31–32.

- Chen, W., Khatib, K., Sten, B., et al. (2020). Detecting media bias in news articles using Gaussian bias distributions. arXiv preprint.
- Chen, W., Khatib, K., Wachsmuth, H., et al. (2020). Analyzing political bias and unfairness in news articles at different levels of granularity. arXiv preprint.
- Chen, W., Wachsmuth, H., Khatib, K., et al. (2018). Learning to flip the bias of news headlines. In Proceedings of 11th Conference of Natural Language Generation (pp. 79-88).
- Culler, J. (2011). The closeness of Close reading. *Adfl*, 41(3), 8–13. <https://doi.org/10.1632/adfl.41.3.8>.
- Daiya, K. (2011). *Belongings: Partition, gender, and national culture in post-colonial India*. Temple University Press.
- Das, A., Dantcheva, A., & Bremond, F. (2018). Mitigating bias in gender, age, and ethnicity classification: A multi-task convolution neural network approach. In European Conference of Computer Vision Workshops.
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.
- Entman, R. (1993). Framing: Towards clarification of a fractured paradigm. In McQuail, D. (Ed.), *McQuail's Reader in Mass Communication Theory* (pp. 390-397).
- Ethayarajh, K. (2019). Attenuating Bias in Word Representations: An Evaluation of Euphemisms. In Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing.
- Ethayarajh, K., Duvenaud, D., & Hirst, G. (Year). Understanding undesirable word embedding associations. arXiv preprint.
- Eve, M. P. (2019). *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*. Stanford University Press.
- Fan, L., White, M., Sharma, E., et al. (2019). In plain sight: Media bias through the lens of factual reporting. arXiv preprint.
- Frampton, G., Whaley, P., Bennett, M., et al. (2022). Principles and framework for assessing the risk of bias for studies included in comparative quantitative environmental systematic reviews. *Environmental Evidence*, 11, 12. <https://doi.org/10.1186/s13750-022-00264-0>

- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505.
- Gangula, R., & Mamidi (2018). Resource creation towards automated sentiment analysis in Telugu (a low resource language) and integrating multiple domain sources to enhance sentiment prediction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*.
- Gentzkow, M., & Shapiro, J. M. (Year). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35–71.
- Gonen, H., & Goldberg, Y. (2019). Analyzing Gender Bias in Harry Potter with Word Embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*.
- Gonen, H., & Goldberg, Y. (2019). Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. *arXiv preprint*.
- Gordon, E. W., Miller, F., & Rollock, D. (1990). Coping with communication centric bias in knowledge production in the social sciences. *Educational Researcher*, 19(3), 14–19.
- Greene, S., & Resnik, P. (2009). More than words: Syntactic packaging and implicit sentiment. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 503-511).
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237.
- Hajari, N. (2015). *Midnight's furies: The deadly legacy of India's Partition*. Houghton Mifflin Harcourt.
- Hamborg, F. (2020). Media bias, the social sciences, and NLP: Automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (pp. 79-87).
- Hamborg, F., Donnay, K., & Bela, G. (2019). Automated identification of media bias in news articles: An interdisciplinary literature

- review. *International Journal on Digital Libraries*, 20(4), 391–415.
- Hamborg, F., Meuschke, N., & Gipp, B. (2020). Bias-aware news analysis using matrix-based news aggregation. *International Journal on Digital Libraries*, 21(2), 129–147.
- Hunter, J., Schmidt, F., & Jackson (1982). Meta-analysis: Cumulating research findings across studies. *Educational Researcher*, 15(8), 20–21.
- Iyyer, M., Enns, P., Graber, J. B., et al. (2014). Political ideology detection using recursive neural networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (pp. 1113-1122).
- Jalal, A. (2013). *The pity of Partition*. Princeton University Press.
- Khan, Y. (2017). *The great Partition: The making of India and Pakistan*. Yale University Press.
- Kiesel, J., Mestre, M., Shukla, R., et al. (2019). Semeval-2019 task 4: Hyperpartisan news detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 829-839).
- Kiritchenko, S., Zhu, X., Mohammad, S. M., & Cherry, C. (2018). Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*.
- Krane, V. (2019). Bias in the Flesh: A Critical Feminist Technoscience Study of Sex Testing in Sport. *Science, Technology, & Human Values*, 44(6), 1019-1044.
- Lash, T. L., Fox, M. P., Cooney, D., Lu, Y., & Forshee, R. A. (2016). Quantitative Bias Analysis in Regulatory Settings. *American Journal of Public Health*, 106(7), 1227-1230.
- Li, C., & Goldwasser, D. (2019). Encoding social information with graph convolutional networks for political perspective detection in news media. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 2594-2604).
- Lin, W., Wilson, T., Wiebe, J., et al. (2006). Which side are you on? Identifying perspectives at the document and sentence levels. In *Proceedings of the Tenth Conference on Computational Natural Language Learning* (pp. 109-116).



- Liu, A., & Svensson, P. (Eds.). (2019). *Decolonizing Digital Humanities: A Reading and Dialogue*. University of Minnesota Press.
- Lohia, R. (2020). *Guilty Men of India's Partition*. B.R Publishing.
- Malhotra, A. (Year). Remnants of a separation: A ghara and a gaz: From Lahore to Amritsar to Delhi. *South Asian Popular Culture*, 14(1-2), 89–100.
- Malmasi, S., & Dras, M. (2020). An Analysis of the Relationship Between Gender and Authorship in Computer Science Research. arXiv preprint arXiv:2008.06666.
- McCarthy, J., Titarenko, L., McPhail, C., et al. (2008). Assessing stability in the patterns of selection bias in newspaper coverage of protest during the transition from communism in Belarus. *Mobilization: An International Quarterly*, 13(2), 127–146.
- Miyuki Yamada, Yuichi Murai, & Ichiro Kumagai. (2013). Story Visualization of Novels with Multi-theme Keyword Density Analysis. *Journal of Visualization*, 16(3), 247–257. <https://doi.org/10.1007/s12650-013-0163-4>.
- Munson, S., & Resnick, P. (Year). Presenting diverse political opinions: How and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1457-1466).
- Munson, S., Zhou, D., & Resnick, P. (2009). Sidelines: An algorithm for increasing diversity in news and opinion aggregators. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 130-137).
- Noble, S. U. (2018). Reckoning with Bias in Algorithmic Culture. In *Data & Society: Points*.
- Owen, R. (Year). Reader bias. *Jama*, 247(18), 2533–2534.
- Park, S., Ko, M., Kim, J., et al. (2011). The politics of comments: Predicting political orientation of news stories with commenters' sentiment patterns. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (pp. 113-122).
- Pennebaker, J. W., Boyd, R. L., Jordan, K., et al. (2015). The development and psychometric properties of LIWC.
- Pennington, S., Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 conference on*

- Empirical Methods in Natural Language Processing (pp. 1532-1543).
- Posner, M. (2016). Whose Stories Are Heard? Exploring Bias in Digital Humanities Projects. In *Debates in the Digital Humanities 2016*.
- Pottker, H. (Year). News and its communicative quality: The inverted pyramid when and why did it appear? *Journalism Studies*, 4(4), 501–511.
- Preethi, P. G., Uma, V., et al. (2015). Temporal sentiment analysis and causal rules extraction from tweets for event prediction. *Procedia Computer Science*, 48, 84–89.
- Pryzant, R., Martinez, R. D., Dass, N., et al. (2020). Neutralizing subjective bias in text. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 480-489).
- Rashkin, H., Choi, E., Jang, J., et al. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2931-2937).
- Recasens, M., Mizil, C., & Jurafsky, D. (2013). Linguistic models for analyzing and detecting biased language. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 1650-1659).
- Riloff, E., & Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing* (pp. 105-112).
- Sarila, N. S. (2017). *The shadow of the great game: The untold story of India's Partition*. Constable.
- Seaver, N. (2017). *Critical Algorithm Studies: A Reading List*. Retrieved from <https://www.nickseaver.net/algorithm-studies.html>
- Spinner, B. W. (2017). *Partition: The Story of Indian Independence and the Creation of Pakistan*. Simon Schuster UK.
- Stefanie Posavec. (2009). *Writing Without Words*. Retrieved from [www.stefanieposavec.com/writing-without-words](http://www.stefanieposavec.com/writing-without-words).
- Tanya E. Clement. (2013). Text Analysis, Data Mining, and Visualizations in Literary Scholarship. In *MLA Commons, Literary Studies in the Digital Age: An Evolving Anthology*.

Retrieved from <https://dlsanthology.mla.hcommons.org/text-analysis-data-mining-and-visualizations-in-literary-scholarship>.

- Wachsmuth, H., Kiesel, J., & Stein, B. (2015). Sentiment flow-a general model of web review argumentation. In Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (pp. 601–611).
- Yip, P. S. F., Xiao, Y., Wong, C., et al. (2020). Is there gender bias in research grant success in social sciences? Hong Kong as a case study. *Humanities and Social Sciences Communications*, 7(1), 1–10.
- Zhao, J., Wang, T., & Yatskar, M. (2018). Gender bias in coreference resolution: Evaluation and debiasing methods. *arXiv preprint*.
- Zhao, J., Wang, T., Yatskar, M., Ordonez, V., & Chang, K. W. (2018). Gender Bias in Coreference Resolution: Evaluation and Debiasing Methods. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

### Chapter 3

#### Historical Memory, Language, and Biases: Re-reading

##### *Midnight's Furies and The Other Side of Silence*

*“The suffering and grief of Partition are not memorialized at the border....A million may have died but they have no monuments. Stories are all that people have, stories that rarely breach the frontiers of family and religious community: people talking to their own blood”*

*-Urvashi Butalia*

Historical methodology involves the meticulous analysis of firsthand materials, assessing their reliability, authenticity, and significance in relation to the events they document. These materials, which include records, visual representations, and cultural artifacts, offer invaluable insights into the past when they were created contemporaneously with the events under investigation. Drawing parallels to the scientific method, historiographers have been likened to judges, tasked with correlating testimonies with material evidence to ascertain the truth of past events (Ginzburg, 1999). In the process of historical inquiry, the authenticity and reliability of sources undergo rigorous evaluation, as it's understood that these materials are not inherently objective. Instead, they bear the imprint of the politics, practices, and circumstances surrounding their creation. Often, such sources selectively document certain aspects of events, reflecting the biases or priorities of their creators. (Clemens and Hughes, 2002, p. 201). Historical inquiry involves a series of pivotal stages (Dill and Aminzade, 2007, p. 269):

1. identifying, locating, and handpicking pertinent sources;
2. arranging and classifying these sources to facilitate deeper scrutiny;
3. subjecting the amassed materials to meticulous evaluation, with special attention to the institutional dynamics that influenced their formation.
4. conducting analyses that may diverge in various directions.

This chapter primarily focuses on the understanding of two distinct historiographical methodologies: archival study and oral history narratives. Archival research serves as the cornerstone of historiographical methodology, providing a primary means of gathering original sources and uncovering previously unused materials and unexplored areas of inquiry. Through meticulous examination of archival materials, historians have the opportunity to unearth previously unutilized resources and explore uncharted territories of study. By scrutinizing elements such as procedural intricacies, contextual significance, chronological sequence, and temporal context, archival investigations offer invaluable insights that can substantially advance theory development, enrich data collection processes, and facilitate rigorous hypothesis testing (Frisch et al., 2012, pp. 11). Unlike social scientists, historians generally do not create their own datasets anew. Instead, they rely on preexisting founts of information, such as archival records, artifacts, and secondary literature, to construct narratives about the past. However, oral history presents a distinct approach. This methodological technique involves obtaining firsthand historical accounts, particularly focusing on aspects absent from official documentation (Blee, 2012, p. 623). It allows for the amplification of voices from those who participated in, opposed, or observed past events. Oral history enables researchers to reconstruct historical periods through the perspectives of individuals involved, while also exploring respondents' memories as they actively construct narratives about the past in the present. When integrated with complementary research methodologies, oral history emerges as a valuable instrument for exploring, uncovering, and interpreting intricate societal developments (Blee and Taylor, 2002, p. 93). Advocated for its capacity to capture the narratives of marginalized individuals excluded from traditional historical archives due to prevailing political, geographic, socioeconomic, gender, or ethnic biases, oral history stands as a vital avenue for amplifying diverse voices and perspectives within historical discourse. This exclusion may be deliberate from external forces or ingrained culturally. Utilizing oral history methods allows researchers to

explore the subjectivity and agency of these voices, which are often portrayed primarily through the perspectives of governmental institutions or influential social organizations in archival records.

The scholarly engagement with the Partition of British India has undergone a significant evolution over time, reflecting changing perspectives and methodologies within historiography. Initially, the Partition was predominantly studied through a political lens, rooted in the extensive body of colonial historiography based on archives that focused on the decisions, policies, and negotiations among political elites leading up to 1947. This approach, often referred to as the "master narrative," emphasized the political origins of the Partition and prioritized analysis of state structures, power dynamics, and nationalist movements. However, by the 1990s, there emerged a notable shift in historiographical approaches towards the Partition. This shift was characterized by a move away from the dominance of the political narrative and towards the utilization of oral histories as a means of understanding the lived experiences of individuals affected by the Partition. While questions about the causes of the Partition remained unresolved for many historians, they became less prominent as scholars turned their attention towards exploring the effects of the Partition on people's lives, identities, and communities. Feminist scholars played a particularly significant role in this paradigm shift by highlighting the gendered dimensions of Partition violence and displacement. Through the lens of oral history narratives, they brought attention to the ways in which women experienced and responded to violence during the Partition. These narratives challenged traditional interpretations of Partition violence as solely rooted in political or communal conflicts, instead revealing the intersecting factors of gender, class, religion, and caste that shaped women's experiences of violence and survival. As a result of this shift towards incorporating oral histories and archival studies, the breadth of understanding of the Partition expanded significantly. Rather than focusing solely on political events and elite actors, historians began to explore the everyday experiences and

livelihoods of ordinary individuals affected by the Partition. Oral history interviews provided a platform for voices that had been marginalized or silenced in mainstream historical narratives, allowing for a more nuanced and inclusive understanding of the human dimensions of Partition.

This chapter in two sections analyses Nisid Hajari's archival study of Partition "*Midnight Furies: The Deadly Legacy of India's Partition*" and Urvashi Butalia's oral history narrative "*The Other Side of Silence: Voices from the Partition of India*" using the computational tool to

- understand how historical memory is narrated through archival study and Oral histories
- study the manifestation and patterns of linguistic biases
- understand how the language of the Holocaust is used to narrate Partition and its aftermath
- study the patterns of describing violence in these works

## **SECTION -I**

### **3.1.1. Archival Study and Historical Writing**

The 19th-century conception of history was based on the German conception of critique of documentary evidence. "The essence of a historian's work and the credibility of historical facts was based on a careful and extensive reading of archival sources" (Blouin and Rosenberg 2011, 14–16). Past was perceived and corresponded to the present through an immersion into these sources either with a lens of correspondence and/or criticism. Sources are viewed cautiously, with a growing awareness that they are narrative constructs that reconstruct realities rather than direct portrayals of reality (Iggers 2005, 144). The term "archive," is defined as an institution storing records and documents from individuals, organizations, and governments that:

- a) stemmed from the substantial professional endeavors of a legal entity or individual
- b) are not relevant to current affairs
- c) are of immense permanent value categorized as "national heritage" or "records of permanent value".

The current archival practices of cataloging, registering, and preserving documents in accordance with the specific bureaucratic process responsible for their generation has its emergence in the sixteenth and seventeenth centuries. This specificity has become a key factor determining the authenticity of administrative documents in current archival practices.

The narrative pattern of archival historians recounts historical events by unraveling the intricate network of causes and effects in the sequence they unfold. "Historians enter the archives in order to represent the history of a field that resists consensus on its origin stories, and thrives on countering master narratives" (Ritter, 2012). But in most cases, the objective of curating an archive is not to achieve the 'narrative whole'. This raises the question of subjectivity, representation, truth, and bias.

### **3.1.2. Partition Historiography Through Archival Research**

Over the years, scholars have delved into the multifaceted dynamics of the Partition, drawing on archival studies to illuminate its causes, consequences, and enduring legacies. Larry Collins and Dominique Lapierre's "Freedom at Midnight" (1975) offers a gripping narrative of the events leading up to the Partition and its aftermath. While not solely based on archival studies, the authors draw from a wide range of primary sources, including interviews, letters, and government documents, to provide a detailed account of the political maneuvers, communal tensions, and human tragedies surrounding the Partition. Yasmin Khan's meticulously researched book "The Great Partition: The Making of India and Pakistan" (2007) delves into the political, social,



and human dimensions of the Partition. Drawing on archival materials from India, Pakistan, and the United Kingdom, Khan offers an analysis of the process of Partition, examining its underlying causes, and the role of colonial policies. "Partition: The Long Shadow" (2015), edited by Urvashi Butalia, brings together a diverse range of scholarly essays that examine various aspects of the Partition, incorporating archival studies as a central methodology. Gyanendra Pandey's "Remembering Partition: Violence, Nationalism, and History in India" (2001) critically analyzes the construction of memory and historiography surrounding the Partition. Drawing on archival sources, Pandey explores how narratives of violence, nationalism, and identity have shaped collective memory and historical interpretations of the Partition in India, underscoring the complex interplay between history, politics, and memory in postcolonial societies. Ian Talbot and Gurharpal Singh's "Divided Cities: Partition and Its Aftermath in Lahore and Amritsar, 1947-1957" (2006) offers a comparative study of the post-Partition trajectories of Lahore and Amritsar, two cities deeply affected by the Partition. Drawing on archival research from both India and Pakistan, the authors analyze the socio-political transformations, demographic changes, and communal dynamics that shaped the urban landscapes of these cities in the aftermath of Partition. Vazira Fazila-Yacoobali Zamindar's "The Long Partition and the Making of Modern South Asia: Refugees, Boundaries, Histories" (2007) offers a fresh perspective on the Partition by centering on the experiences of refugees and the creation of new national boundaries. Based on archival research conducted in India, Pakistan, and the United Kingdom, Zamindar traces the long-term repercussions of Partition on displaced populations, border communities, and state-building projects in South Asia. Neeti Nair's "A History of Partition: Violence, Borders and Memory" (2009) offers a critical examination of Partition historiography, focusing on the intersections of violence, borders, and memory. Drawing on archival materials from India, Pakistan, and the United Kingdom, Nair analyzes how Partition has been remembered, commemorated, and contested in the postcolonial contexts

of India and Pakistan, shedding light on the enduring legacies of this traumatic event.

Scholarly engagement in archival studies on Partition includes critiques, comparative studies, and thematic explorations. These analyses evaluate the accuracy of narratives presented, the efficacy of the research methodology in understanding colonial policies' role in shaping the Partition, interdisciplinary approaches to understanding the Partition's legacies, and theoretical framework and its application in understanding post-Partition memory and nationalism amongst others. This section of the thesis adds to this scholarly framework by understanding the linguistic biases and their manifestation in the archival research in Nisid Hajari's "Midnight's Furies".

### **3.1.3. "Midnight's Furies: The Deadly Legacy of India's Partition"**

Nisid Hajari's "Midnight's Furies: The Deadly Legacy of India's Partition" (2015) is a new narrative history of Partition. This book is significantly deviant from the current trends in historiography from/on South Asia as it colossally relies on archival study rather than on oral narratives. The book is acclaimed as an antidote to historical manipulation, as Hajari provides a clear and consistent thesis with assisting documentation on Partition along with exhaustive portraits of Jawaharlal Nehru, Mohammed Ali Jinnah, Gandhi, Vallabhbhai Patel, and Lord Louis Mountbatten. The author mentions that the objective of the book is to "answer a different question – not why the subcontinent was split, or who was to blame for the massacres, but how the experience of Partition carved out such a wide gulf between India and Paki-stan" (Hajari, 2015). The book provides an intricate examination of the imperative nature of Pakistan's formation and how the colonizers orchestrated communal events to prolong their dominion in South Asia. The author does not hesitate to deal with accounts of violence during Partition paralleling it with major post-colonial discourses to describe how Partition has engendered countless divisions for those who inherit it. The book has a prominent account of Kashmir whilst describing Partition as the sole reason for the future unrest in the region. "The

author chose well the origin of communal plots, which directly derived from the imperial conspiracies, and how it succeeded to keep this strategically extremely crucial land the ‘neck-bone’ for both India and Pakistan” (Thakur, 2016). The first chapter, ‘Furies’, describes the ‘Great Calcutta killings’ that followed the ‘announcement of the Direct Action Day by the Muslim League on 16 August 1946’. In the second chapter ‘Jinnah and Jawaharlal’ Hajari uses ‘interspersing personal anecdotes about Jinnah and Nehru’ to cover ‘the political background of the Indian subcontinent’. The third chapter, ‘Madhouse’ delineates the events that led to the Noakhali riots and Bihar riots, the violence, and the widening of fissures after the riots culminating in the growth of ‘Hindu Mahasabha’ and other ‘orthodox Hindu outfits’. The chapter also discusses the violence against the Sikhs in the North-West Frontier Province (NWFP), the counter-mobilization of the Sikh minority, and the efforts to establish an interim government with Nehru and Jinnah on board to control the holocaust. The fourth chapter, ‘Pakistan Murdabad’, briefs the violence and bloodshed especially in Punjab after the declaration of transfer of power and the failed negotiations over it. The fifth chapter, ‘The Indian Summer’ deals with the question of the division of military and other assets between the two nations while the sixth chapter, ‘Off the Rails’, highlights the ‘Boundary Commission Award’ and the violence post it. The seventh chapter, ‘Stop This Madness’, discusses the ‘issue of Junagadh’ and the intensified violence that reached its peak in mid-September 1947. Hajari also documents the tensions between Nehru and Patel. The eighth chapter, ‘Ad hoc Jihad’, provides details on the ‘Kashmir issue’. In the ninth chapter ‘Himalayan Quagmire’, the author describes the territorial conflicts between India and Pakistan concerning Kashmir. which pushed both the countries to the brink of war which in turn made the relations between the Muslims and Hindus in India bitter. In the ‘Last Battle’, the tenth and final chapter, Hajari portrays ‘Gandhiji’s assassination’, ‘the issue of Kashmir’, and ‘the accession of Hyderabad’. The author employs “private papers, intelligence reports, diplomatic records, first-hand accounts, and a large number of secondary sources” and presents the

collected data as an anecdote that “animates the historical characters whose human failings and idiosyncrasies enliven the historical rendition of the saga of the Partition” (Jha,2017). “The book is remarkable in the sense that it sees ‘Partition’ as a real event that took place 60 years back—nothing like ‘what if there was no Partition?’ or any other hypothetical wanderings. Moreover, the author appears keener as a historian rather than a journalist in writing this book” (Thakur, 2016).

### **3.1.4. Linguistic Bias Analysis of *Midnight's Furies***

After conducting a comprehensive computational linguistic bias analysis using the developed tool, it is evident that Hajari's writing demonstrates bias through four predominant methods.

1. misinterpretation of evidence
2. omission of significant facts
3. failure to cite evidence
4. casual explanations that mislead the reader

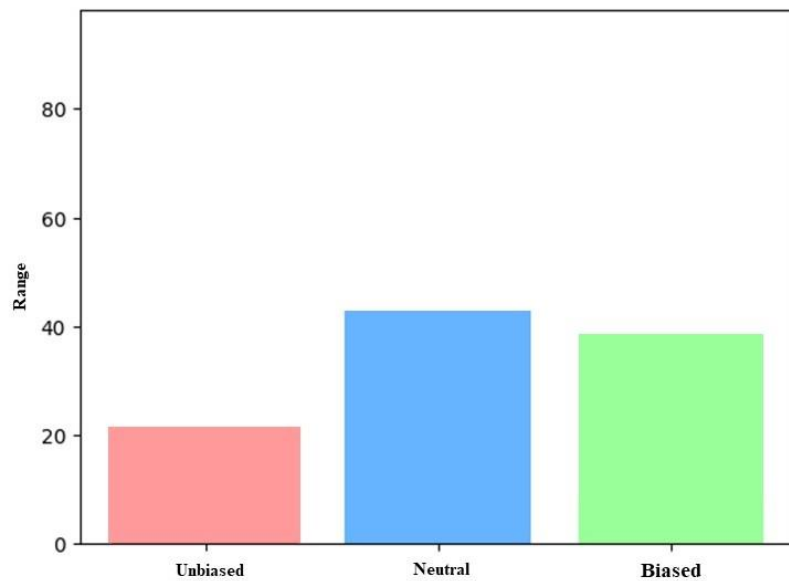
Biases in Hajari’s work have a twofold manifestation, primarily in the narrative pattern and secondarily arising out of archiving and archival study. Hajari is less proficient in conveying the profound sentiments underpinning the call for Pakistan. The absence of mention of Muhammad Iqbal, the philosophical founder of Pakistan and its initial advocate leaves the reader inadequately informed about the historical impetus behind Pakistan, consequently attributing the responsibility solely to Jinnah. Deprived of its contextual foundation, the Hajari’s description of Partition appears as a tragicomic interplay of errors, characterized by a clash of egos.

While Hajari asserts that Winston Churchill referred to Clement Attlee as "a sheep in sheep’s clothing," it was actually Ramsay MacDonald, the Labour leader, to whom Churchill directed this remark. While Hajari may characterize Satyagraha as "soul force," it's essential to emphasize that this characterization doesn't strictly adhere to the literal interpretation. Instead, Satyagraha encapsulates "the force or

insistence of truth. "Hajari's assertion that Muslim opposition to the Congress anthem "Vande Mataram" was based on the presence of numerous anti-Muslim verses may not accurately capture the root cause. Rather, objections from the Muslim community largely stemmed from the depiction of India as the goddess Durga in the original novel from which the verses were derived. Edwina Mountbatten was not the daughter but the granddaughter of Sir Ernest Cassel. Contrary to Hajari's depiction, the accepted initialism for the Hindu nationalist organization Rashtriya Swayamsevak Sangh is not "R.S.S.S." but rather the simpler "R.S.S."

A more concerning source of bias is Hajari's excessive reliance on secondary sources for narration without proper citations. For Instance, in recounting the Calcutta killings through the perspective of Nanda Lal, a bakery owner, Hajari presents the incident with vivid detail, He writes: "The air clung to his skin and sweat-stained kurta like a damp rag." The precision of description may mislead the reader into perceiving it as firsthand information, whereas, it is a reinterpretation of the narration of the event from Margaret Bourke-White's book "Halfway to Freedom."

The analysis of the book shows that 39% of biased sentences and 20% of sentences were categorized as unbiased. Consistent with the data-driven nature of the narrative 41% of the sentences were classified as neutral. Figure 3.1.4.1 shows the consolidated results of the sentence-level analysis.



*Figure 3.1.4.1: Sentence-level bias analysis of Midnight's Furies*

The major source of Hajari's bias is the slanted objectivity and framing effect bias towards Jinnah. The manifestation, reasons, and repercussions of this will be discussed further in the subsequent sections of this chapter. The second prominent bias category is confirmation bias followed by religious bias. Hajari's religious and political biases are revealed either through the lack or overloading of evidence. To cite an example Hajari write: "Each side had its uniformed fanatics. And each side had political leaders who fanned the flames with their rhetoric" (64) but he further provides evidence only to the acts by the Muslim league citing "League propagandists visited Muslim communities across northern India with photographs from Bihar of unburied skeletons and mutilated refugees, and with the charred pages of Korans they claimed had been burned by Hindu mobs" (64,65). Similarly, on multiple occasions, Hajari provides evidence only when it is favorable for his argument and this is perceived by the tool as religious and/or political bias. For instance, Hajari accuses Jinnah of involvement in funding and sponsoring war in Kashmir, a claim made without substantiated proof. Hajari writes: "There is no question that the Quaid personally approved the funding and sponsorship of a proxy war in Kashmir" (196). He similarly alleges the Muslim League's complicity in the "Calcutta

killings and Punjab bloodbath” (196). Expanding on this biased portrayal, Hajari draws parallels with another assumed allegation, suggesting that over six decades, “a succession of Pakistani leaders justify all manner of covert operations from the country’s nuclear program to its sponsorship of the Taliban” (196). Hajari extends these comparisons to contemporary times, drawing parallels between the Partition era and present-day issues such as Kashmir (205, 178), Islamic terrorism (209), and India-Pakistan conflicts. The tool flagged these instances for misinformation effect, attribution, confirmation, and selection biases.



### 3.1.5. Archival Biases in Hajari's Historiography

whether the author intends to counter the master narrative or to agree with it. In the process of building a narrative from the archival records, the author attempts to model historical texts, artifacts, and objects into a plot that is convincing to the readers. Author at the helm become intertwined within the multivocality of history. intervening archival spaces and narrative. This results in a biasedness central to the development of the narrative.

Hajari draws his work from documents, correspondences, diplomatic cables, and personal journals of participants of the 1947 Partition from “The Broadlands Archives” at Hartley Library, the National Army Museum UK, “South Asian Studies Archives” at Cambridge, “Modern Papers” at Bodleian Archives at Oxford University, “India Office records” at the British Library, “British National Archives” at Kew, “National Archives of India” at New Delhi and “U.S. National Archives”. He orders the information from these archives, especially drawing from diaries of British officials in India, and develops a narrative that views the archival materials through an Indian lens. Through his book, Hajari not only presents anew the Partition and the events causing and following it, encompassing its individuals, locales, and occurrences within the archival records but also molds a collective memory that revoiced the confirmation bias, thereby solidifying prevailing belief systems.

“It can be seen that every time an archival researcher makes a decision, they are carrying out a philosophical act, which arises from a cultural context and has cultural implications” (Pearce, 1992). The statement by Pearce underscores the profound impact of cultural biases on archival research, highlighting that every decision made by an archival researcher reflects a philosophical act embedded within a cultural framework, with far-reaching cultural implications. Indeed, when documenting human experiences, as is the primary function of most archives, there is a pervasive encoding of inherent cultural biases. These biases often stem from various factors, including gender identity,



religious beliefs, ethnic backgrounds, socioeconomic class, and other cultural norms and consensus. In line with previous research findings, much of the bias evident in Hajari's work arises from what can be termed as 'absence' – a lack or underrepresentation of certain perspectives or narratives within the archival records. One prominent example of this absence is the significant omission of women's experiences from the history of Partition. As noted by Schwartz and Cook (2002), archives, since their inception in ancient times, have systematically excluded records about or by women from their holdings. This systemic exclusion perpetuates a skewed representation of Partition, wherein the voices and experiences of women are marginalized or altogether absent, thereby distorting our understanding of the past.

Understanding the biases present in Hajari's work requires a comprehensive exploration of the foundational principles of archival practice, particularly focusing on the appraisal and cataloging processes, despite the author's lack of explicit acknowledgment of these factors. Appraisal, as a crucial step in archival management, involves the assessment of materials to determine their suitability for inclusion in the archival collection. However, this process is susceptible to biases stemming from institutional policies or the personal preferences of the archivist. As Cox (2004) aptly articulates, the act of appraising a record as archival bestows upon it a certain power, influencing not only the perception of the record itself but also shaping the perspectives of the professionals entrusted with its care and interpretation. Similarly, cataloging plays a pivotal role in the organization and accessibility of archival materials. This process entails identifying relationships between records to facilitate navigation and the creation of coherent narratives. However, cataloging extends beyond the mere identification of pre-existing relationships; it involves actively constructing these connections. Meehan (2010) underscores this aspect, emphasizing that the cataloging process is fundamentally about the creation of relationships among records. In the context of Hajari's account, the biases inherent in the archival process significantly influence the

structure and content of the narrative. The decisions made during appraisal and cataloging dictate which perspectives and voices are prioritized and represented in the narrative, potentially resulting in a skewed portrayal of Partition.

Hajari's historiography is based on letters, diaries, and speeches of Nehru, Archibald Wavell, Mountbatten, Pamela Mountbatten, Ruttie Jinnah, Gandhi, V.P Menon, Patel, and others. While a significant portion of the narrative revolves around Jinnah, Hajari doesn't use any records that are directly attributed or composed to him. As Hajari here relies on 'state-archives' for his narrative the nationalistic selective characteristic of the data favours a state-centered view of history and writing of a "history from above". There is a clear absence of Historical evidence from the leaders of the Muslim league whereas this account is Partition history prefaced with the documents from the colonizer supplemented with records from National Archives in India favoring the national ideology of Unified India. The narrative not only excludes the minorities it distorts, silences, and erases the tale of the underrepresented. This bias can be partially attributed to the process of archiving. As Derrida writes: "The archive ... produces memory, but produces forgetting at the same time ... [T]he trace is at the same time the memory, the archive, and the erasure, the repression, the forgetting of what is supposed to be kept safe." Rather than being an 'archival ethnography' that allows the author to see archives just as a site of fieldwork Hajari produces a renewed history of Partition solely based on the archives. The narrator's role here doesn't have an ethical mediation allowing them to face the narrative contradictions, rather Hajari here represents data rather than presenting it.

### **3.1.6. Bias Manifestation through 'chain of causality' of events developed by Hajari**

Narrative causality encompasses the organization of information based on the inherent cause-and-effect relationships among events. It pertains to how narratives are structured to reflect the

interconnectedness and consequential flow of events unfolding. Within this context, information is categorized as causal when its content directly leads to subsequent narrative events. whereas, Noncausal information is defined as content within a narrative that does not directly influence the progression of events and can be omitted without disturbing the coherence of the overall storyline. This category encompasses details that, while adding depth or context, are not integral to driving the plot forward or impacting the narrative trajectory. (Dahlstrom, 2010). Prior studies propose that where information is positioned within a narrative plays a substantial role in shaping its perceived authenticity within real-world contexts. This suggests that the arrangement of details can affect how credible it appears to readers or audiences, influencing their interpretation and acceptance of the narrative as reflecting reality. According to Dahlstrom (2010), data incorporated at a causal juncture within a narrative is perceived as more credible than when placed in a noncausal context. This phenomenon, termed the narrative causality effect in earlier research (Dahlstrom, 2012; Magliano, 1999; Trabasso & Sperry, 1985), holds significant implications for the manifestation of biases as well as the narrative's interpretation and its perceived truthfulness by the reader. The influence of narrative causality may intersect with various other factors in narrative persuasion and the manifestation of biases. This interplay introduces an unexplored dimension of variability that could augment the predictive capacity of linguistic biases and narrative persuasion.

Hajari's historiography is based on the narration of cause and effect. The book is an exploration of the manifold causes of events and their larger consequences. The author attributes Partition only to the immediate conflict of interest between the 'Indian National Congress' and 'the Muslim League' and also each incident of violence is attributed only to its precursor. The author records 'The direct-action day by the Muslim League' as the only cause of the Calcutta Killings and Bihar riots, the Noakhali riots, the Sikh uprising, and the Kashmir uprising are portrayed as causes and effects of the subsequent. Hajari ascribes the

present-day tumultuous and dysfunctional relationship between India and Pakistan to the historical event of Partition. He establishes a causal link wherein the enduring fissures of Partition have instigated a sequence of bilateral crises, hindering the reciprocal relationship between the two nations. He further holds the adverse India/Pakistan relationship responsible for impeding the overall progress of South Asia. Hajari outlines the interconnected causality of events stemming from Partition, leading to armed conflicts between India and Pakistan, followed by the violent uprisings and insurgency in Kashmir. Hajari argues that this chain of causality further contributes to a surge in global repercussions of radical Islamic terrorism.

The narrative focus on immediate cause and effect within Hajari's work serves to overshadow the broader contextual factors that contributed to the events surrounding Partition. By emphasizing the conflict between the Indian National Congress and the Muslim League as the primary catalyst for Partition, the narrative neglects to adequately address the longstanding communal tensions that predated colonization and were further exacerbated by colonial policies. This narrow focus on immediate political maneuvers obscures the deeper historical roots of communal unrest and fails to acknowledge the role of colonial powers in exacerbating divisions within Indian society. Similarly, the attribution of an immediate cause to the rise of Islamic terrorism(56, 196) spanning from Partition to India Pakistan conflicts oversimplifies the complex systemic injustices faced by minority communities globally. By framing Islamic terrorism solely within the context of specific incidents or geopolitical conflicts, the author overlooks the broader socio-political and economic factors that contribute to the radicalization of individuals and communities. This reductionist approach to causality not only perpetuates stereotypes but also detracts from a nuanced understanding of the underlying issues driving extremist ideologies.

While the chain of causality presented in the narrative may contribute to its persuasive power, it also results in a biased perspective that oversimplifies the multi-faceted intricacies of Partition and the

broader socio-political landscape. The categorization of these biases under framing effect bias, religious bias, confirmation bias, speculative bias, and representation heuristic bias by the tool further underscores the limitations of this reductionist approach. By failing to contextualize events within a broader historical framework and acknowledging the systemic injustices that underpin them, the narrative ultimately undermines efforts to fully grasp the complexities of Partition and its enduring impact on the region.

### **3.1.7. ‘Jinnah the Fallen Angel’ Metaphor and Linguistic Bias Manifestation**

Jinnah throughout Hajari’s description has the imagery of the fallen angel, Lucifer suggesting a symbolic association with a figure who has fallen from grace or ideals. The tool categorizes the usage of this metaphor under framing effect bias and confirmation effect bias. This metaphorical association suggests a narrative framing that positions Jinnah as a figure who has strayed from his original ideals or moral principles. By drawing parallels to Lucifer, traditionally seen as a symbol of rebellion and downfall, the author implicitly casts Jinnah in a negative light, framing his actions and decisions within a moral framework of betrayal or deviation from righteousness. Additionally, the confirmation bias comes into play as the metaphor reinforces pre-existing beliefs or assumptions about Jinnah held by the author as well as his audience. Through the lens of framing effect bias and confirmation bias, the computational tool highlights how metaphors can be wielded to influence perception and reinforce existing beliefs, ultimately shaping the narrative discourse surrounding historical figures like Jinnah.

Hajari Writes, mentioning Sarojini Naidu “Mrs. N spoke of Jinnah rather as of Lucifer ... a fallen angel, one who had once promised to be a great leader of Indian freedom, but who had cast himself out of the congress heaven” (22). Unlike Nehru Jinnah is portrayed as a man with humble beginnings, who had a moral stance and righteousness that surpassed those of his subordinates and Hajari states that Jinnah looked like ‘India’s man of destiny’ (23). According to Hajari Jinnah’s “religion

is what made him unusual among the well-heeled Congress leaders” and the author on multiple instances addresses him as ‘The Muslim’. In akin to the fallen angel Jinnah is also portrayed as a man with shaken religious faith. He is described as ‘Whiskey-drinking’, ‘chain-smoking’, and nominal Muslim. His success as in Delhi and Karachi meetings is labelled as a bruised ego and Pageantry by the author. His leadership is reduced to based “partly on bluff, partly on his image as an inveterate defender of Muslim rights” (49). Jinnah’s fall is articulated in connection with his personal and political trajectories. As the Congress and Nehru and Gandhi, ascended to prominence, Jinnah found himself marginalized, leading to a descent in both his personal influence and political efficacy. Hajari writes: “He became what he never wanted to be – a purely Muslim politician, reduced to petitioning for concessions for his community” (28). The narrative consistently intertwines failures in Jinnah’s personal life with his political decisions. The failed marriage with Rattanbhai, the 16-year-old daughter of his friend Sir Dinshaw Petit, and her mysterious death are recurrently cited by the author in tandem with Jinnah's embittered political choices. The author dedicates significant narrative space to unravel the enigma surrounding Rattanbhai's demise, probing whether it was a consequence of suicide or other factors. The narrative posits that following his estrangement from the Congress, Jinnah ceased to be the nationalist hero Rutledge had fallen in love with. Towards the midsection of the book author conveys a notion that ‘there was nothing left for Jinnah in India’ (29) paralleling his influence in league becoming negligible, a metaphorical descent, resembling Lucifer's expulsion from heaven. Hajari utilizes adjectives such as 'Distressed' (94), 'disturbed state of mind' (94), 'doesn't commit to anything' (104), 'mood grown dark' (131), 'paranoid' (142), 'unapproachable and uncompromising' (236), 'bitter and biased' (236), 'forbidding personality' (249) ‘nervous and edgy’ (48), ‘less in command of himself’(48) to articulate the intricate and layered dimensions of Jinnah's character.

The narrative undergoes two discernible shifts in its depiction of Jinnah's character. The first one post the election victory of the Muslim League where the author portrays “Even the Whiskey-drinking, chain-smoking Jinnah as a messianic figure” (46). Jinnah is portrayed to be opposing the establishment of Pakistan as an ‘Islamic theocracy’, a man who “didn’t need to pretend for the mullahs nor obfuscate to keep his flock united”. He is pictured as someone who always intended—“a multifaith democracy, just as he had once championed for India itself”. Hajari also cites the below speech by Jinnah at the opening section of ‘Pakistan’s Constituent Assembly’ to further delineate this changed image.

You are free to go to your temples, you are free to go to your mosques, or to any other place of worship in this State of Pakistan. You may belong to any religion or caste or creed—that has nothing to do with the business of the State. . . . We are starting with this fundamental principle that we are all citizens and equal citizens of one State. . . . Now I think we should keep that in front of us as our ideal, and you will find that in the course of time Hindus would cease to be Hindus and Muslims would cease to be Muslims, not in the religious sense, because that is the personal faith of each individual, but in the political sense as citizens of the State (128)

In his historiography, Hajari conspicuously shifts blame onto Jinnah whom he perceives as the most convenient figure, exemplified in his assertion that "there is little question that Jinnah was the most polarizing figure in the Partition drama" (p. 249). This statement reflects a clear attempt to assign singular responsibility to Jinnah for the divisive and tumultuous events surrounding Partition. By portraying Jinnah as the primary instigator of polarization, Hajari simplifies the complex socio-political dynamics that contributed to Partition, overlooking the multifaceted nature of the conflict. Furthermore, by labeling Jinnah as the "most polarizing figure," – which is flagged by the tool as framing effect, Hajari effectively absolves other key players and factors from

accountability, thus narrowing the scope of analysis and perpetuating a one-dimensional portrayal of Partition selectively highlighting aspects of Jinnah's leadership to fit a predetermined narrative.

The metaphorical representation of Lucifer to construct the portrait of Jinnah serves as an integral structural component of the order of discourse. Operating heuristically, it functions on the level of discursive thought, distinct from the pre-discursive impact of metaphors observed in metaphors within literary discourses. The patterned relationship between the characterization of Jinnah and Satan dramatically objectifies concepts, motifs, and themes. As a result, this metaphoric imagery communicates meaning with a dramatic potency and succinctness akin to a visual representation, demonstrating its capacity to evoke vivid and impactful understanding.

### **3.1.8. The Nehru Bias**

Hajari's narrative displays a notable bias towards an India-centric perspective, which manifests in a pronounced favoritism towards the viewpoints of Nehru, thereby obstructing the achievement of a comprehensive and impartial account of historical events. This favoritism towards Nehru's perspectives is identified by the computational tool primarily as confirmation bias and political bias. The confirmation bias is evident in Hajari's tendency to selectively present information or interpretations that align with Nehru's positions, thereby reinforcing pre-existing beliefs or assumptions about Nehru's leadership and policies. Furthermore, the computational tool categorizes these sentences as exhibiting concision bias, false consensus bias, hindsight bias, and egocentric bias. The concision bias is suggested a tendency to prioritize brevity over thoroughness in presenting information, potentially leading to oversimplification. The false consensus bias indicates an assumption that Nehru's perspectives are widely shared or accepted, overlooking dissenting viewpoints or alternative interpretations within the historical discourse. Additionally, the hindsight bias is evident through Hajari's retrospective evaluation of



Nehru's actions, as he inadvertently projects contemporary knowledge as Nehru's hindsight.

In the first pages of the book Hajari describes Nehru as "dashing", "famously handsome", and had "high, aristocratic cheekbones and eyes that were deep pools - irresistible to his many female admirers". The author also notes that "although disdainful of superficialities, he took great care with his appearance". In contrast to "irascible" and power-hungry Jinnah, Nehru is described with almost super-human qualities. The author marvels at the intellect of Nehru:

Well into middle age now, Nehru retained the same coiled energy he had exuded as a firebrand in his twenties, not long out of Cambridge, when he had packed volumes of Proust on his frequent trips to His Majesty's jails. He could still quote Shelley and Walter de la Mare by heart. Yet in his speeches he also spoke to the yearnings of millions of illiterate, grindingly poor Indian peasants. To them he represented all the possibilities they imagined for freedom. (2)

The book is close to becoming a hagiography of 'India's first Prime minister'. Hajari's Nehru is a 'true believer' (47) who has 'thrown himself into the cause' (47) of India's independence. Hajari writes:

He gave up his lawyer's suits for kurtas made of homespun cotton. He replaced the rich roasts and claret at Anand Bhavan with plain flatbread and lentils. He even got himself a charkha, a wooden spinning wheel, and heeded Gandhi's call for every Indian to spin yarn at least an hour each day to break the dependence on British-made cloth. (47)

The characterization of Nehru frequently serves as a foil to that of Jinnah. Hajari portrays their personas in contrast to each other, noting that Jinnah's "frigid demeanor was as legendary as Nehru's charm" (27). In contradiction to the ones used for Jinnah Hajari uses descriptors such as 'Flush with passion' (49), 'adored by millions' (49), 'Courageous in curbing violence', 'rising to the occasion' (133), 'suave' (250),

‘sensitive’(250), ‘handsome’ (250) ‘most sincere’(102) to describe Nehru. Nehru is characterized as someone who “quickly developed a rapport” (90) by speaking to others “in terms that they understood, and in a polished Cambridge accent” (90), and his audience expressed enthusiasm “about his statesmanship and literary sensibilities” (90). In Contrast to that of Jinnah adversities and losses in Nehru’s personal life are strategically employed to bolster a favorable portrayal of his character. Although Nehru's letters to Kamala from prison may appear mundane and obligatory, filled with inquiries about family members, requests for new reading material, and reminders for Kamala to prioritize her health, their bond is characterized as an "affectionate and respectful marriage." Hajari forces a portrayal of a deep and affectionate connection between them within their marriage despite the ordinary content of their correspondence, in contrast to his narrative attempts about Jinnah's family life. Kamala’s demise is used by the author to elevate Nehru’s marriage as well as his personality. In his profound admiration, Hajari’s characterisation of Nehru in itself has a hindsight bias Hajari's portrayal of Nehru inherently exhibits a hindsight bias, wherein the author attributes a considerable foresight to Nehru to foresee and anticipate the repercussions of significant events. The author without leaving out any chance to contrast Nehru and Jinnah compares even their correspondences, Hajari writes:

The Congress leader wrote lucid, eloquent letters; articles for international magazines; long aide-mémoire; carefully preserved diaries; and more than one autobiography. The Quaid-i-Azam Mohammad Ali Jinnah Papers— eighteen volumes (and counting) compiled by admirers in an attempt to match the voluminous collected works of Nehru and Gandhi—include only a few revealing speeches and letters. The bulk of Jinnah’s correspondence is numbingly pedestrian: one typical exchange from the spring of 1945 carefully preserves for posterity the extended back-and-forth between the leader of India’s 100 million Muslims and the Matheran Electric Supply Company

over what he insisted was an extortionate bill of 10 rupees for replacing a lightbulb. (94)

Hajari's narrative demonstrates a discernible bias stemming from his disproportionate focus on the actions and, oddly, even the inner thoughts of prominent figures such as Gandhi, Jinnah, Mountbatten, and Nehru. Moreover, Hajari's tendency to assign blame primarily to Jinnah and The Muslim League further reinforces his India-centric perspective, reflecting a bias toward interpreting Partition history through the lens of Indian nationalism.

### **3.1.9. Historiography as a Two-men issue and manifestation of Narrative Biases**

Hajari's historical interpretation of the Partition is fundamentally based on the perennial conflict between two men who are central to the events. Primarily, it unfolds as the story of righteous Jawaharlal against his arch-rival Jinnah. After the Partition, the narrative extends to encompass the ongoing rivalry between Sardar Vallabhbhai Patel and Nehru. Thus, Hajari's historical discourse revolves around the duals between key personalities. This pattern of approaching historiography as a conflict narrative is a major source of linguistic biases in Hajari's text. As Cobb suggests, historical narratives focusing on conflicts between individuals "characteristically suffer from oversimplified characters and plots, with storylines that work mainly to legitimize the self and delegitimize others, thus increasing polarization" (Cobb, 2006). The narrative is focused on describing the causes of the conflict between Jawaharlal and his rival, the essence of the discord, and the characterization of the rival, among other pertinent aspects.

Hajari's narrative demonstrates the qualities of a comprehensive conflict-supportive narrative, incorporating two main narrative structures that extend across time, traversing from the past to the present and projecting into the future (Bar-Tal, 2007, 2013). This narrative framework not only delves into historical events but also considers their ongoing implications, shaping perceptions and expectations regarding

the conflict's resolution and its repercussions for future generations. Hajari mainly employs two narrative frameworks in describing conflicts.

1. "The concept of collective memory"(Paez & Liu, 2011) in this context encompasses the narrative that intricately describes the roots and development of a conflict, aiming to create a cohesive and meaningful depiction of Partition. This narrative framework forms a comprehensive understanding of the conflict's trajectory over time.
2. "The ethos of conflict," (Bar-Tal, 2007a, 2013; Oren, 2009) as articulated in narrative form, delineates the prevailing societal mindset, offering insights into the present circumstances and environment while also establishing a path forward with defined objectives for the future. This narrative construct not only reflects the collective attitudes and beliefs about Partition but also serves as a framework for understanding its dynamics and aspirations in relation to it.

Tool categories framing effect bias, speculative bias, and misinformation effect in sentences that Hajari uses for

1. Rationalizing the conflict and narrating its developmental trajectory
2. delineating the threat of Partition and its aftermath
3. Characterizing the opponent's objectives as unwarranted and irrational
4. delegitimizing the opponent
5. crafting a glorified image of 'unified India'
6. victimizing the Indian side and INC
7. Fostering sentiments of patriotism and unity
8. Expressing a yearning for peaceful coexistence

Hajari's Narrative here is constructed on selective supportive sources that favour the Indian story of the conflict marginalising contradictory

information. The narrative further uses a framing language contributing to a linguistically biased historiography.

### **3.1.10. Religious bias**

In Hajari's narrative, a substantial portion, 38% of biased sentences, are categorized by the computational tool as exhibiting religious bias, particularly towards the Muslim community. This bias manifests in various forms within the narrative, as identified by the tool's sub-categories. One prevalent form of religious bias is coverage bias, wherein certain aspects of the Muslim community's experiences or perspectives are disproportionately highlighted or marginalized compared to other groups. Speculative bias is another sub-category identified by the tool, indicating a tendency to make unfounded assumptions or conjectures about the beliefs, motivations, or actions of individuals or groups within the Muslim community. The tool also identifies gatekeeping bias (selective inclusion or exclusion of information or perspectives), framing effect bias, and misinformation effect bias.

Hajari is quick to judge that the “men who are crying for a Muslim state were neither observant Muslims nor oppressed” (10). When attributing violence to a Hindu mob, he associates them with their political affiliations as Congress or Rashtriya Swayam Seva Sangh (abbreviated as RSS by Hajari). But when the violence is initiated by the Muslim League, as after the direct-action day and other similar instances, Hajari uses the religious tag ‘Muslim’ as the identifying marker. Hajari creates a portrayal that Muslims believed in British rule rather than in a unified India at the outset of the book. Following Partition, Hajari portrays that the community's allegiance is directed towards Pakistan, revealing the author's confirmation bias. The computational tool flags Hajari's rationale for Violence from Hindus citing slaughtering of cows, and peasant oppression as biased and It's crucial to recognize that there's often a tendency to depict Muslims as the initiators of violence, overlooking the context that may have led to

their actions. Instances of violence perpetrated by Muslims are frequently viewed in isolation, without acknowledging the underlying factors or historical background that may have contributed to such behavior. Hajari uses descriptors such as 'long bearded' (51,52), 'raggedly dressed crowd' (53), and 'forerunners of Taliban' (53) to demarcate individuals without explicitly referring to their religious identity, which is categorized as a biased perception by the tool. The author is swift to mention Maulana Abdul Kalam Azad's religion 'himself a Muslim' when depicting the conflict between him and Patel regarding the Delhi riots against Muslims.

"If one were to characterize a dominant strand in historical writing on the immediate coming of Partition it would instead focus on the politics of elite conflict in late colonial India in the 1930s and 1940s, and to the elite manipulation of religion (and popular politics) as the central key to Partitions coming" (Gilmartin, 2015). Though Hajari doesn't portray Partition as "the culmination of a civilizational clash between Hinduism and Islam," he Projects "Hinduism" and "Islam as distinct, internally coherent yet mutually opposing systems". While Hajari doesn't view Partition solely through the lens of a clash between Hinduism and Islam, he does present Hinduism and Islam as separate and internally cohesive belief systems that are depicted as mutually contradictory. He avoids characterizing Partition solely as the result of a clash between these religions, but still underscores the notion that Hinduism and Islam are distinct entities with their own internal consistencies. With an evident nostalgia for colonization throughout the book, he fails to accentuate the colonizer's manipulation and communization of religion to serve their exploitative purposes. Similarly, another of Hajar's biased perceptions leaves the role of Hindu communalism unaccounted for, in his depiction of Partition violence.

### **3.1.11. Portrayal of Women by Hajari and Linguistic Biases**

The historical memorialization of Partition is significantly patriarchal. Even amid the surge of the oral history wave, the 2015

narrative persistently directs its attention only towards men. Hajari's treatment of women within the historical discourse remains conspicuously elementary, consistently tethering their significance to male counterparts. With the sole exception of Sarojini Naidu, the women featured in the narrative are consistently contextualized in relation to prominent men. Utilising the LIWC feature identification of the tool it was found that in this work with detailed and comprehensive accounts of violence, instances of violence against women are mentioned on a mere 16 occasions. Women of Hajari are irrepressible and flirtatious dinner party companions. The author describes Sarojini Naidu as follows:

Naidu was a big, bawdy, irreverent woman—a scintillating raconteur and dinner party companion, especially compared to the often dour Indian politicians. Not even Gandhi escaped her lacerating wit: she impishly called the Mahatma “Mickey Mouse” for the way his ears stuck out from his bald head. (23)

Another woman featured in the narrative is Rattanbai, Jinnah's wife, whose characterization primarily serves to underscore the flaws and setbacks in Jinnah's life. The author characterizes her as follows:

he had a sixteen-year-old daughter—a sinuous beauty named Rattanbai, or “Ruttie.” Jinnah would have been hard-pressed to ignore her presence. She wore gossamer-thin saris that clung to her body and had a ready, flirtatious laugh. One prim memsahib described her as “a complete minx.” (41)

The other two women Mountbatten's wife, Edwina, and Padmaja Naidu, Sarojini Naidu's 'buxom' (49) daughter with whom the author suspects Nehru to be having a 'steamier affair' (49) are portrayed as admirers of Nehru's heroics. The only mention of Nehru's wife is about her death. The characters of Edwina and Padmaja are strategically employed to illuminate Nehru's susceptibility to the influence of women, though the trait is not overtly presented as a flaw in the narrative. The author describes Edwina as:

Mountbatten's wife, Edwina, the svelte and fabulously wealthy daughter of King Edward VII's banker, Sir Ernest Cassel, came along as well. A photograph showed the attractive vicereine seated regally during the recital, her host—the erstwhile leader of Asia—curled at her feet like a cat. (102)

Hajari terms the Noakhali allegations as 'overhyped stories about rapes and abductions' (71). The only time he engages with the subject of the Indian woman is in describing Gandhi's idealization of feminine virtue which the author terms as a 'late Victorian obsession' (28)

To Gandhi, the ideal Indian woman was cast in the mold of Sita, the god Ram's blameless wife, who first walked on hot coals then meekly went into exile when her chastity had been questioned. Gandhi urged the Hindu women of Noakhali to remember "the incomparable power of Sita."<sup>29</sup> He wanted them to commit suicide rather than submit to their Muslim ravishers: they should "learn how to die before a hair of their head could be injured." Perhaps they could "suffocate themselves or . . . bite their tongues to end their lives," he advised. Told that such methods were impracticable, the Mahatma suggested the next day that they drink poison instead. (71)

The complex history of 'nationalist and patriarchal fetish on women's sexuality' is reinforced without criticism in the narrative. The inadequacy of representation becomes particularly apparent when considering the extensive gendered violence that occurred, a dimension that is underrepresented, apart from a few token references in fiction. Hajari's work serves as compelling evidence of the deficiencies and gaps within the academic study of the 1947 Partition, particularly in understanding its social and cultural dynamics, regardless of the emergence of the Oral history movement.

### **3.1.12. Linguistic indicators of biases**

In this work, one notable linguistic indicator of bias is the consistent utilization of bigrams, which are pairs of consecutive words



frequently occurring together in the text. Interestingly, these bigrams often refer to members or groups aligned with the opposing side of the author's perspective, suggesting a tendency to associate certain terms with a particular bias. Moreover, the tools employed in this study have identified nouns as a primary indicator of word-level bias within the text, a departure from the findings of other texts analyzed. This suggests that the author's choice of nouns, whether in reference to individuals, groups, or concepts, plays a crucial role in shaping the biased nature of the narrative. Additionally, emotional words, regardless of their polarity, have emerged as another noteworthy indicator of bias within the text. Furthermore, the high prevalence of biased metaphors, accounting for close to 70% of all metaphors employed by the author, underscores the pervasive nature of bias throughout the work. Comparative sentences within the text were also flagged for bias, indicating a consistent pattern of biased language usage throughout the work. Additionally, almost all instances of words such as 'what if,' 'could,' and 'may' were flagged for speculative bias, indicating a tendency on the part of the author to engage in speculative or hypothetical reasoning without sufficient evidence or justification. These further underscores the author's inclination towards presenting speculative assertions as factual or authoritative statements, thereby contributing to linguistic biases. Transitional phrases such as 'due to,' 'because of,' 'owing to,' and 'consequently,' as well as signal words such as 'cause,' 'reason,' 'lead to,' 'result in,' 'trigger,' and 'ramification,' were categorized as biased. These words and phrases are commonly used to develop cause-and-effect descriptions within the narrative.

## **3.2. SECTION -II**

### **3.2.1. Testimony, Memory, and Representation in Oral Histories**

Novelist David Lodge has provided a definition of history, wherein he describes it as 'the verdict of those who weren't there on those who were'. "In the best dynamic of an interview, interviewees reverse the equation, trying to explain to those of us who weren't there

how things really were”. (Ritchie,2014) “And the very act of the oral histories, in their long, slow, unfolding and the different qualities (long interviews, minimal interruption) enacts a different pattern of communication and exchange”. (Colton and Ward, 2005). Although oral history has only recently gained prominence in historical methodology, it can be regarded as one of the earliest forms of historical inquiry, predating organized writing methods. Initially, its primary purpose was to uncover factual information. However, as it has evolved, oral history has shifted its focus towards shaping historical narratives rather than merely gathering facts. This adaptation has been prompted by criticisms regarding its effectiveness and has led to a re-evaluation of its methodologies. oral history has transitioned from being a tool primarily used for gathering empirical data to becoming a means of actively shaping historical narratives. The oral narrative presents a distinct lens through which to examine history, offering an alternative perspective that diverges from traditional written accounts. These narratives often intertwine seamlessly in terms of temporal sequencing, blurring the rigid timeframes typically imposed by historical discourse. This occurs because individuals often anchor their memories based on personal experiences or milestones rather than the specific events that historians use to delineate historical periods. As a result, oral narratives transcend and intersect with the disciplinary boundaries of conventional historical narratives. They challenge the linear and compartmentalized nature of traditional historiography by weaving together personal recollections, emotions, and interpretations of past events. “They offer us a way of turning the historical lens at a somewhat different angle, and to look at what this perspective offers” (Butalia, 13).

In their book "Oral History: An Interdisciplinary Anthology," David Dunaway and Willa Baum outline the evolution of oral history through four generations of practitioners. The first generation, led by Allan Nevins and Louis Starr, viewed oral history as a method to gather the untold stories of prominent individuals, primarily for future historians, research purposes, and the creation of oral biographies. The

second generation emerged in the mid-1960s with the establishment of basic archives. They expanded the scope of oral history beyond elites to include marginalized and non-literate individuals, aiming to empower historically disenfranchised communities. Rooted in the social history movement, their work intersected with feminist and educational movements, pioneering new approaches to historical inquiry. The third generation, which emerged in the 1980s, emphasized the process of conducting oral history interviews. This generation responded to the challenges posed by new technologies such as computerized research aids, which increased the capital investment required for professional oral history collections. Concurrently, they navigated the ideological landscape marked by extreme conservative movements and liberal countercultures. The fourth generation of oral historians marked a shift by embracing technological advancements such as video and cassette recorders, as well as computer technology. They recognized that oral interviews and their construction represent history in themselves, shaped by the interactions between interviewer and narrator, contemporary trends, and cultural conventions. This generation placed a different significance on the utility of oral history as a tool for historical inquiry, acknowledging its role in shaping narratives within a historical context.

While the value of oral and life histories is indisputable, there are problems with collecting and using them as historical sources which historians like Ron Grele trace under the categorization of record management and evaluation. One of the most prominent yet less studied problems is the linguistic and cognitive biases in oral history narratives. It is less addressed due to what Michael Frisch has criticized as ‘Anti-History’ which is viewing “oral historical evidence because of its immediacy and emotional resonance, as something almost beyond interpretation or accountability, as a direct window on the feelings and . . . on the meaning of past experience”. Oral testimonies inherently involve the act of remembering, and while individual memories vary in their reliability, thoroughness, and truthfulness, interviews often contain inaccuracies, vague information, and occasionally even deliberate

fabrications. Bias is an inherent aspect of any piece of writing, including interview transcripts. However, it is possible to mitigate bias by acknowledging its presence and actively incorporating this awareness into the composition of a text. By recognizing and addressing bias, writers can provide readers with a clearer understanding of the interview process and its potential limitations, thus enabling a more insightful interpretation of the text.

### **3.2.2. Oral History Narratives in Partition Studies**

The adoption of oral history has increasingly become prevalent within Partition Studies starting from the 1990s. This surge in popularity stems from the desire to delve into 'history from below,' thereby altering perceptions of Partition. Instead of centering solely on the actions of 'great men' in history, this approach broadens the narrative to encompass the experiences and perspectives of ordinary people. Oral history, emerging as a novel methodological tool, has played a crucial role in facilitating this paradigm shift. Feminist historians such as Ritu Menon, Kamla Bhasin, Anam Zakaria, and Urvashi Butalia were among the pioneering figures who ventured into crafting narratives centered around women's encounters during the Partition era. They embarked on the endeavor of documenting and interpreting the experiences of women amidst the tumultuous events of Partition in works like "Borders and Boundaries" and "The Other Side of Silence". This academic scholarship marked a significant departure from the prevailing narratives of Partition, which predominantly focused on the intricacies of decolonization, the key players involved, and the geopolitical dimensions of power transfer. By shifting the perspective and embracing a grassroots approach to history, these feminist scholars and activists illuminated unsettling realities surrounding Partition. They shed light on issues such as abduction, sexual violence, and the perpetuation of patriarchal norms within the societal fabric during a time typically associated with the euphoria of independence. Through their work, Butalia, Menon, and Bhasin challenged the boundaries of discourse, prompting critical questions about the agency and representation of

marginalized voices. They delved into the complexities of power dynamics and the ability of marginalized communities, often referred to as the "subaltern," to articulate their experiences and narratives within dominant historical narratives. But in contrary Paola Bacchetta suggests that, their accounts “reflect a different kind of subaltern writing that challenges established subaltern writing, which . . . continues to marginalize women.” (Bacchetta, 2000)

Starting from the late 1990s, there has been a noticeable surge in fascination with the Partition, particularly intensifying during milestone anniversaries such as those in 2007 and 2017, and thus a growth in the number of oral history accounts published through various platforms including social media. This growing interest is largely driven by both producers and consumers within a burgeoning heritage industry dedicated to commemorating the event. Though memory can be an unreliable witness, “it also provides a window into the human dimensions” of history poignantly capturing the psychological scars of Partition. Most of the oral histories were collected using crowdsourcing protocols to collaborate with the public in documenting oral histories. The expansion of interdisciplinary research and advancements in technology have played a pivotal role in refining both the remembrance and representation of the Partition, particularly through the utilization of new media. These explorations have rightly extended beyond the traditional focal points of Punjab, reaching regions like Sindh (formerly the North West Frontier Province), Assam, and broader North India. Moreover, there has been a shift away from solely focusing on capital cities such as Delhi and Karachi, towards exploring provincial spaces like Hyderabad.

Firsthand memories, recounted by survivors, have transcended geographical boundaries and have been carried by the diaspora, spreading beyond the confines of the subcontinent. The intergenerational accounts of refugees have added depth to these narratives, demonstrating the enduring impact of Partition even across generations. The dispersed South Asian diaspora, particularly in the UK

as well as in North America, has been instrumental in driving the current surge of interest in Partition studies, “leveraged by the market might (predominantly of Indians, rather than Pakistanis/Bangladeshis), with an accompanying need to connect fractured histories through tropes of memory, myth, and nationalism” (Vidre, 2022).

The proliferation of initiatives spearheaded by dedicated professionals, such as the Partition Education Group based in the UK, has opened up new avenues for engaging with history and heritage in ways that are both inclusive and localized. Concurrently, works like journalist Kavita Puri's "Partition Voices," stemming from the acclaimed three-part BBC Radio series commemorating seventy years since Partition, and featuring recorded oral histories from the diaspora, have resonated widely with the general public, exemplifying the concept of public history. Similarly, the BBC's highly regarded television documentary, "My Family, Partition, and Me: India 1947" (2017), delves into the compelling narratives of four British families, capturing the personal and poignant aspects of the Partition experience. The utilization of dramatized personal journeys to explore family histories has contributed to the increasing popularity of examining history through individualized accounts, albeit at the expense of diminishing the significance of structural and state actors. However, these individual-focused approaches pale in comparison to two prominent initiatives based in the United States. One of these initiatives, spearheaded by Harvard University, involves the Lakshmi Mittal and Family South Asian Institute. This project, titled "Looking Back, Informing the Future: the 1947 Partition of British India," is dedicated to preserving and uncovering diverse narratives surrounding what is often regarded as the "largest migration in history." The project characterizes its interview collection as a comprehensive mixed-method analysis, involving "300 trained volunteers" who conducted "2,396 interviews using semi-structured questionnaires" in the native languages spoken across India, Pakistan, and Bangladesh. Tarun Khanna, affiliated with the Harvard Business School, led the project and reflects on the challenges

associated with employing quantitative analysis on qualitative memories. He emphasizes that memories are inherently subjective and malleable constructs, and highlights the limitations of using natural language processing techniques to translate and transcribe oral data collected in multiple languages into English. In this endeavor, the project represents a departure from earlier approaches taken by Butalia, Menon, and Bhasin. Instead of solely relying on qualitative analysis, the project utilizes manual coding, employing linear regression to analyze demographic data pertaining to the narrators. The findings reveal that 68 percent of the narrators were men, 78 percent belonged to the majority community, and 81 percent identified as middle/upper class.

The 1947 Partition Archive, initiated by physicist Guneeta Singh Bhalla in 2010, stands as a remarkable endeavor that presently boasts a collection of 10,200 oral histories sourced from fourteen countries and spanning over thirty-six languages, all compiled with the aid of dedicated volunteers. Describing itself as a "crowdsourced community-based archive," the project's mission is encapsulated in its call to action: "share, collect, donate." The initiative not only provides training but also actively encourages volunteers who have undergone its workshops to record and submit interviews via its user-friendly online platform. Stanford University Libraries serve as the archival repository for 4,000 of these interviews, complemented by a scholarship program aimed at facilitating the recording of interviews in the native languages of South Asia. While fifty-one of these video-recorded interviews are accessible for online streaming, it's worth noting that certain interviews may require a Stanford affiliation for viewing. Reflecting on Tarun Khanna's earlier concerns, it's important to note that twenty-seven of the interviewees were aged ten or younger in 1947, adding a unique perspective to the narrative. Additionally, twenty-four interviews were originally conducted in non-English languages and subsequently translated for wider accessibility. In terms of demographics, the project reveals intriguing insights: 60 percent of the interviewees are male, with over 90 percent hailing from middle/upper-class backgrounds,

encompassing roles such as landlords, civil servants, and industrialists. Furthermore, four interviewees have connections to princely or aristocratic families, underscoring a departure from earlier trends in subaltern history.

Prior to the emergence of Partition initiatives originating from the United States, the Centre for the Study of Developing Societies in New Delhi launched its own Partition project with financial support from the Ford Foundation. Directed by Ashis Nandy and Rajmohan Gandhi, this project commenced in 1998 and amassed approximately 1,300 interviews conducted primarily across India, with some coverage extending to Pakistan and Bangladesh.

The year 2017 witnessed Partition gaining considerable attention within the South Asian community, particularly after receiving widespread coverage in the UK. This newfound prominence prompted physician Binita Kane and other professionals to take notice of the issue “to demonstrate that South Asian history is also British history”. The initial proposal to designate August 17 as Partition Commemoration Day in the UK evolved into the establishment of an annual South Asian Heritage Month, spanning from July 18 to August 17. The primary aim of this initiative was to elevate awareness and appreciation of South Asian heritage throughout the UK. However, there are concerns that such an approach to commemoration may risk oversimplification, resembling a superficial “food and festival” approach to multiculturalism. Each of these initiatives strives to maintain an apolitical stance, seeking to distance itself from the inherently political processes that precipitated Partition and its enduring repercussions, including the rise of Hindu nationalism in contemporary times. However, this detached approach bears a resemblance to collecting testimonies related to the Holocaust without adequately contextualizing them within the broader framework of Nazism or the Nakba.

In a parallel effort, The Arts and Cultural Heritage Trust (TACHT) took a significant step by establishing a Partition Museum in



the border city of Amritsar. Housed in the former colonial-era town hall, this museum was made possible through the support of the government of Punjab. Though TACHT is now establishing a second museum in Delhi, there exists an implicit risk associated with this state-endorsed heritage creation. It can potentially align with the interests of majoritarian nationalism in India, thereby inadvertently perpetuating and strengthening conflictual relations with neighboring states. The museum, both on its website and in its physical layout, integrates oral history excerpts through short videos of selected interviews, supplemented by artifacts and newspaper archives. However, in such hybrid public-private spaces, curatorial decisions hold the power to reduce oral history to mere excerpts, a phenomenon described by Ted Svensson as "fragments of memories." Visits to these public history spaces, adorned with emotive labels like "Hope" and "Refuge," may function as performative acts of remembrance, echoing what Raphael Samuel refers to as "theatres of memory."

Another noteworthy initiative is the Museum of Material Memory, a digital repository established in 2017 by Aanchal Malhotra. This innovative project traces family histories and social ethnography through heirlooms, collectibles, and objects of antiquity. Malhotra's book "Remnants of Partition" (2019) emerged from this endeavor, featuring accounts of twenty-one objects accompanied by interviews. These concise narratives offer easily shareable and digestible insights into the Partition experience, resonating with global audiences across various social media platforms amidst a dispersed diaspora. In Aanchal Malhotra's exploration of Partition through personal narratives and objects, a number of critical observations arise. Her focus primarily rests on elite individuals and their experiences, as well as the material artifacts associated with them. This approach has been critiqued for its failure to provide a comprehensive understanding of the diverse and often marginalized perspectives within the Partition narrative. Malhotra's interviews predominantly feature individuals from privileged backgrounds, such as poets, landlords, intellectuals, and civil servants,

reflecting a homogenous pool of participants. Their accounts and the objects they share are far from ordinary, often representing the elite class of 1947 and contemporary society. This focus on the elite overlooks the experiences of the masses who were unable to carry nonessential belongings during the turmoil of Partition. Moreover, Malhotra's uncritical exploration of nostalgia and cliched narratives of loss and displacement fails to critically engage with the complexities of the political landscape and the varied impacts of Partition on different groups and regions. Her admiration for the refined English accents of her interviewees further highlights a disconnect from the realities of the broader population, where English was not widely spoken. Additionally, the inclusion of oral interviews with colonial British figures without acknowledging the power dynamics of colonial relationships raises concerns about the portrayal of equality in suffering. This overlooks the asymmetrical nature of imperial relations and the injustices inherent in colonial structures. Furthermore, Malhotra's approach has been criticized for objectifying Partition narratives and reducing them to exhibits of memorialization within the public domain. The intersubjective element of trust, essential in oral history, is often compromised by factors such as gender, caste, religion, and language, which can influence the dynamics between interviewer and interviewee. Overall, Malhotra's work highlights the challenges and limitations of oral history in capturing the diverse experiences and perspectives of Partition, particularly when the focus remains on elite voices and nostalgic narratives.

In Pakistan, where the memory of Partition is often framed more as the genesis of a homeland rather than a narrative of loss, systematic efforts to collect oral histories have been relatively limited. One notable initiative is the Citizens Archive of Pakistan, conceived by the Academy Award-winning documentary filmmaker, Sharmeen Obaid Chinoy. This archive provided the multigenerational stories that formed the basis of Anam Zakaria's book, "The Footprints of Partition." In contrast, there is a palpable sense of urgency driving the US-based projects affiliated with

Harvard and Stanford. Historian Priya Satia describes this as "a race against time," highlighting the pressing need to capture as many oral histories as possible before the generation that directly experienced Partition passes away. While the drive to collect and preserve firsthand accounts is undeniably commendable, it is not without its methodological challenges. Chief among these challenges is the realization that the task is far from straightforward or easily achievable.

Katja Müller identified that many of the interviewees for the 1947 Partition Archive were descendants of individuals who experienced Partition firsthand. Motivated by a desire to document the memories of their aging family members, these volunteers embarked on the task of conducting interviews. This phenomenon is not uncommon, as it is often the case that researchers are drawn to study subjects with which they have a personal connection. However, these interviews also serve as a blend of history, memory, and the process of remembering, a common characteristic observed in oral histories focused on well-documented historical events (Vidre, 2022). Oral historian Lynn Abrams has pointed out the complex nature of memory, highlighting its symbiotic relationship with the public commemoration of the past. This underscores the intricate interplay between personal recollections and collective memory in shaping historical narratives.

The aspect of trust in oral history is vital, particularly in documenting the accounts of a society that is fragmented. Various factors, including gender, age, geographical location, religion, caste, social class, and language, often dictate the dynamics of the conversation, determining who leads or dominates it. Religious, caste, and regional identities, easily identifiable through names, can either hinder or enhance trust between the interviewer and interviewee. Oral historians are well aware of the challenges involved in identifying potential interviewees, leading many to rely on personal networks of friends and family. This reliance necessitates a heightened level of self-awareness and reflexivity.

While oral histories undoubtedly offer invaluable insights into past experiences, their utilization as historical sources present several challenges that historians must navigate. One of the most significant hurdles is the presence of various forms of bias embedded within these narratives. Bias can manifest in numerous ways, stemming from the subjective perspectives, memories, and interpretations of the individuals recounting their stories. Whether influenced by personal beliefs, cultural norms, or societal pressures, biases inherent in oral histories can shape the narratives in subtle yet impactful manners. It's essential to acknowledge that complete elimination of bias from any form of storytelling is unattainable. However, addressing bias involves recognizing its existence and accounting for it when constructing an oral history narrative. This section of the chapter attempts to understand the manifestation of different patterns of linguistic biases in oral history narratives of Partition through a close reading of Urvashi Butalia's "The Other Side of Silence" using the developed computational tool.

### **3.2.3. The Other Side of Silence – Urvashi Butalia**

*The Other Side of Silence* is an oral historical account by Urvashi Butalia, first published in 1998. This is a pioneer work in terms of its extensive engagement with people's history of Partition. Butalia documents stories across borders, of victims and perpetrators, men and women, families and villages, Dalits and children, and develops them into engaging narratives of ordinary witnesses of Partition. The oral narratives within her work offer compelling and immersive storytelling, standing in contrast to the factual data found in official records. They unfold the experience of violence and trauma and its drawn-out effects on individuals, families, and communities. The work is an example of an outstanding feminist study. Butalia is involved with her subject on a deeply personal level partly because it is the history of her family too. Although Butalia does not delineate her objectives, this account may help the reader to understand Partition as an event from a humanitarian perspective. Often when people have lived through trauma of this magnitude they prefer to forget; Butalia quotes the author Krishna Sobti

who wrote: "Partition is difficult to forget but dangerous to remember." Yet in each communal killing in the present-day subcontinent, the echo of Partition reverberates.

The narratives in her book suggest that in moments of crisis and despair, notions of honor and pride become so thwarted that the protection of these is perceived to justify any means. And so it comes to be that cutting the throats of one's daughters is preferable to them being raped by people of another religion, and that to reject from society those women who did get abducted is preferable to having to live with the "illegitimate" children of any cross-religious liaisons. Here, violence is perpetrated and then covered up in such a way that much can be said through silence. Butalia writes:

. . .it was when I went back over our conversation that it struck me that that awkward silence, that hesitant phrase was perhaps where the disappearance of the two sisters lay hidden: in a small crack, covered over by silence. I realized then that in this silence lay the many hidden histories of Partition, the histories that have always hovered at the edges of those that have been told, the histories that describe the dark side of freedom (106)

Although the author's primary focus is what happened to women, we also hear stories of what happened to children and people from Scheduled Castes. Urvashi Butalia brings to our attention the wider prospects of agency, autonomy, and notions of purity by narrating how the issue of abducted women and children is handled post-Partition India as well as Pakistan. Butalia's account in fact is Partition history seen from the eyes of the marginalised.

Subjective and embodied voices of the ordinary can create narratives beyond the standard history. Oral History is essentially the study of human memory. Naturally, it comes with its own drawbacks, the greatest of which is that we cannot fully rely only on memory; there will always be gaps and sometimes discrepancies, for human memory is not a diligent recorder, it is a pathological liar too. Memory is not

unmediated, it often depends on who remembers, when, with whom, and how.

How can we know that, four or five decades after the event, the stories are not simply rehearsed performances, or they are told differently for different people, perhaps tailored to suit what the person thinks the interviewer wishes to hear? How do we reach beyond the stories into silences they hide, how can we assume that speech, the breaking silence, is in itself a good thing? (12)

#### **3.2.4. Linguistic Biases in *The Other Side of Silence***

The inherent bias in the accounts provided by survivors is anticipated in the oral history narratives. In juxtaposition with various narratives, discernible memory biases manifest prominently in the majority of these stories. The devised tool lacks the training to discern categories of memory biases beyond the false memory effect and its subcategory, childhood amnesia. Consequently, the tool typically classifies sentences under memory biases along with available subcategories. Notably, a significant 72% of biased sentences in the text are attributed to memory biases. However, it's essential to acknowledge that linguistic biases also coexist within these sentences. This underscores the complexity of bias detection within oral history narratives and highlights the need for a methodological framework for simultaneously understanding both memory and linguistic biases in textual analysis.

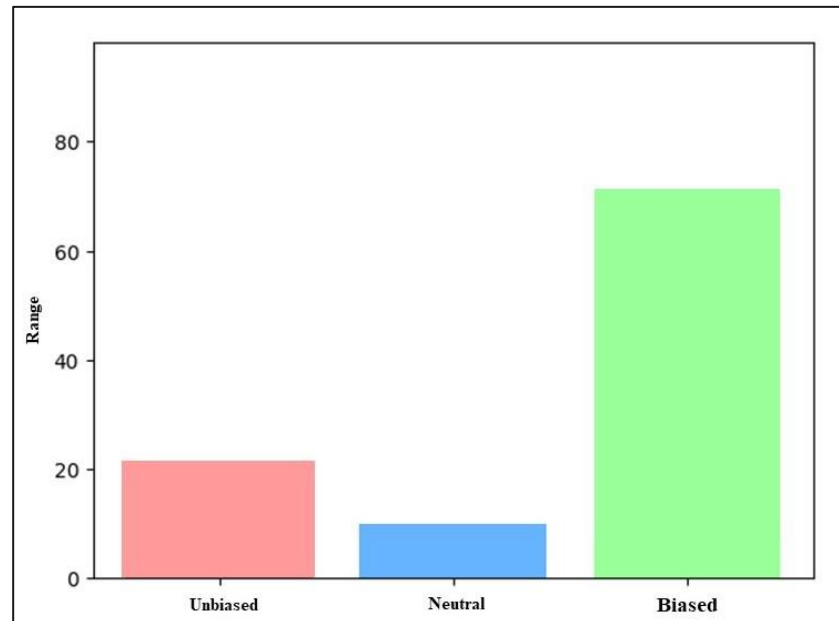
The linguistic biases in the text manifest primarily as Butalia occasionally acquiesces to the biases of her interviewees, particularly yielding to certain entrenched patriarchal notions. For instance, individuals like Bir Bahadur Singh and Basant Kaur, whom she interviewed, demonstrate a reluctance to acknowledge women as capable of violence. In response, Butalia appears to align with these biases by either implicitly or explicitly downplaying the agency of women involved in violent acts. This is evident in her conscious effort to reframe the actions of women who coerced others into suicide,

relocating them into the more comfortable and symbolic realms of sacrifice (as suggested by Bahadur Singh and Basant Kaur) or victimhood (as implied by Butalia herself). Such linguistic biases contribute to the perpetuation of gender stereotypes and undermine the agency and complexity of women's experiences during Partition. The tool identified instances of such bias under multiple categories, including representation heuristic bias, confirmation bias, and memory bias. This suggests a tendency within the text to rely on simplified representations, confirming pre-existing beliefs or narratives, and potentially distorting recollections or interpretations of events.

Another bias evident in Butalia's narrative lies in her framing of violence as an external phenomenon, perpetuated by individuals or groups perceived as outsiders. By attributing violence solely to external forces, Butalia implicitly absolves members of affected communities from responsibility, portraying them as victims rather than active participants. Throughout the explorations, Butalia's concerns about the relationship between nation-building and violence come to the fore. This perspective overlooks the complexities of inter-community tensions and the role played by internal dynamics in fueling violence during Partition. Furthermore, Butalia's characterization of violence as originating from outside the boundaries of family and community reflects a bias towards preserving the integrity of these social units. Butalia explains, "As long as violence can be located somewhere outside, a distance away from the boundaries of family and the community, it can be contained". By framing violence as an external intrusion, which the tool identifies as framing effect bias, she reinforces the notion of communal cohesion and solidarity, potentially overlooking instances where violence may have been perpetrated within these same social structures. Moreover, Butalia's emphasis on locating violence as belonging to the 'other' reinforces a binary division between 'us' and 'them'.

Figure 3.2.4.1 shows the range of biased, unbiased, and neutral sentences in *The Other Side of Silence*. There is a higher number of

biased sentences as there were inherent biases in the testimonies falling into the categories of both memory and narrative biases.



*Figure 3.2.4.1.: Range of Linguistic Biases in The Other Side of Silence*

In Butalia's narrative, the tool identifies egocentric bias in the focus on individual experiences potentially overshadowing broader historical contexts. Though the tool doesn't directly suggest a patriotic bias, the LIWC features tags the glorification of certain nationalist figures while downplaying the contributions of others. The tool also identifies Confirmation bias highlighting selective emphasis on evidence and Selection bias through selective inclusion resulting also in gatekeeping bias. The framing effect is apparent in the narrative's structure and emphasis on specific themes or motifs, influencing readers' perceptions of the events described. The most evident bias in Butalia's text is childhood amnesia. The stories in the entire text are told by adults who are recollecting what happened to them as children. "Six, eight, nine, eleven, ten ... these were the sorts of ages they had been then" (258). The reliability of their memory, the childhood fabrications in the narratives, the accuracy of their recollection, the Zeigarnik effect in the stories and the underplays of the adult mind on those stories are highly disputable. Butalia writes:





stereotypes, and expectations can influence the recollection process, shaping the narrative in ways that may diverge from objective truth. Thus, while oral histories offer valuable insights into personal experiences, they must be approached with caution due to the inherent unreliability and susceptibility to bias inherent in memory recall.

### **3.2.5. Bias Manifestation Through Author's Interventions**

When attempting to document oral histories recording their experiences without judgment is very important. It is hard to do so because we are all furnished with preconceived notions. The moment an interviewer arrives at an interview with an already made-up mind about what he or she may achieve during the next few hours, it means they had fallen prey to judgment and bias. The work of an oral historian as per definitions is to record the past in the way in which it happened—not to contort, vilify, or justify it. But in *The Other Side of Silence*, the author is at the center of her narrative. Interviews were conducted to support a specific argument the author was making. When presenting these interviews to the audience, the author has chosen to narrate them in a particular way, adding her own storytelling elements. The questions asked by the interviewers, and interventions by other people are removed. Thus, the transcript of the interview itself is not unmediated. When converting interviews into written text, a significant amount is inevitably lost. This includes the nuances of tone, pauses, and emotional expressions, as well as the body language of the participants. Additionally, the deliberate influence of the interviewer in shaping the direction and content of the interview is often not fully captured in written transcripts. These elements are crucial for a comprehensive understanding of the interview context and the nuances of the conversation. Most of the time Butalia's narrative revolves around the myth of intellectual honesty and she admits to having "silenced those parts that needed to be kept silent" (37) being afraid of the implications it would have had.

### **3.2.6. Selection of Interviews and Resultant Biases**

Another notable bias apparent in Butalia's work is the arbitrary criterion she employs for selecting interviews to include in her narrative. In her own words, Butalia acknowledges that she has only included a fraction of the interviews she conducted, opting instead to include stories that held personal significance or resonated with her on a deeper level. She writes: "...But of the number I spoke with, I have included only a fraction...I have included the stories that meant the most to me, stories of people with whom I have formed real friendships or stories to which I keep returning again and again" (14). This selection process, driven by subjective factors such as personal attachment and emotional connection, introduces a bias that stems from the author's failure to recognize the limitations of her approach within the context of historical scholarship. By prioritizing stories based on her own subjective criteria rather than objective historical relevance, Butalia inadvertently introduces a bias that may skew the representation of events and perspectives within her work. Another significant aspect contributing to selection bias in Butalia's work is the absence of interviews from Pakistan, except for that of her uncle. Additionally, there is a notable lack of direct interviews with individuals from the Muslim community, which further compounds the limitations of her selection process.

### **3.2.7. Dual Perspectives: Butalia's Use of Personal Narratives in Historical Inquiry and Doubts on its Efficiency**

*The Other Side of Silence* is part of an ever-growing corpus of work that approaches history from the angle of personal narrative, while ongoing discussions within historical discourse persist regarding the legitimacy of this methodological approach. The diversion from traditional methodologies by prioritizing personal narratives over the conventional focus on "facts and dates." This departure from the norm has drawn both curiosity and criticism, which Butalia acknowledges throughout her work. Despite her established reputation as a feminist scholar, Butalia finds herself in the unexpected position of defending her

chosen methodology and, by extension, feminist historiography as a whole. The first chapter of Butalia's book is dedicated to elucidating and justifying her approach, often tinged with a sense of apology. She frequently questions the legitimacy of her voice and methodology, emphasizing her identity as not strictly a historian but rather someone who engages with history through personal and political lenses. She admits, "I am not a historian. History is not my subject." Instead, she explains her entry into this field as a result of political engagement, particularly with issues surrounding communalism, and her deep-rooted belief in feminism. she writes then on page 71:

"I am not a historian and have neither the capability nor indeed the interest to explore these [historical] questions. I am concerned instead with the consequences of Partition for people then, and its ramifications now, in their lives. My focus here is on the small actors and bit-part players, whose lives, as the lives of all people, were inextricably interwoven with broader political realities".

Throughout the text, Butalia's focus remains steadfast on the human dimensions of historical events, particularly the Partition of India. She emphasizes her interest in the "small actors" and "bit-part players" whose lives intersected with broader political realities, eschewing grand narratives in favor of nuanced individual experiences. Urvashi Butalia's extensive focus on justifying her methodology in "The Other Side of Silence" could inadvertently manifest biases and because of the repetitive and self-explanatory tone. While her repeated emphasis on defending her chosen approach may serve to address potential criticisms and maintain transparency with readers, it also raises concerns about the potential reinforcement of confirmation bias and the limited scope of analysis. This overt self-justification might compromise the perceived objectivity of her work and alienate certain readers, potentially leading to misinterpretations of her intentions.

### **3.2.8. The one-sidedness of Butalia's Narrative**

A notable lacuna in Butalia's work lies in its unilateral perspective. Aside from her uncle, Urvashi Butalia lacks information or interviews from Pakistan. Butalia contends, "It is one of the tragedies of Partition that researchers working on this major event in the history of the subcontinent can only have access to both the countries – India and Pakistan - if they belong to a third country" (22). Despite occasional references to the prevailing animosity toward the Muslim community, Butalia herself has not conducted any interviews with individuals from the community. While addressing Muslim suffering, she refrains from direct interviews and relies instead on interviews conducted by others, such as Peter and Satti. These individuals, in their film, interviewed Murad, a Tonga driver, who articulated the nuanced reality: "class is not so easily dismissed after all," and emphasized that "landlords go to the landlords and the poor go to the poor."

Acknowledging the inherent narrative essence of social history, Butalia acknowledges her subjective role as a narrator, highlighting the interpretative element inherent in constructing any historical narrative. She contemplates the decision to incorporate her perspectives and interpretations into the narrative, adopting a method that intertwines oral testimonies with her own insights. The book's thematic organization, with headings such as 'Blood,' 'Facts,' 'Women,' and others, allows the interviews to transcend individual anecdotes, offering a comprehensive exploration of various themes. This narrative strategy, while potentially blending specific details, transparently conveys the subjective nature of historical interpretation, presenting an approach that openly acknowledges the author's influence and viewpoint in her historiography. The absence of interview questions and the interviewer in the text exempts reflections on the interview dynamics, emphasizing the predominant role of the author/narrator, who holds sole responsibility for representing the intertwined narratives. The author's narrative approach involves omitting questions and interventions, and molding interviews, thereby losing subtleties such as intonations,

pauses, and body language in the textual rendition. Her effort to challenge the idea of formal history as absolute on one hand and personal stories or oral histories as less deserving of formal documentation on the other is a recurring contrast throughout the book.

### **3.2.9. Linguistic Indications of Biases**

In analyzing Butalia's text for linguistic biases, it becomes evident that certain patterns emerge, particularly in the use of emotive language, defensive language, and the deployment of euphemisms or loaded terms. Most of the sentences in Butalia's text that were categorized as linguistically biased had an Emotive language, employed to evoke strong emotional responses signaling subjective interpretations. The biased sentences also had a prominent defensive language, mostly in the case of Butalia's methodological justifications and self-explanations, aimed at justifying actions or deflecting criticism. Additionally, the use of euphemisms or loaded terms to frame events has revealed the highest range of biases. Furthermore, the frequent use of the word 'one' in Butalia's text is observed as also seen in the word cloud, which sometimes serves as a bias inducer by generalizing experiences or perspectives.

### **3.3. Language of Describing Violence in *The Midnight's Furies* and *The Other Side of Silence***

In Hajari's work, when describing violence, three narrative logics emerge:

#### **1. Reduction and Oversimplification of the Victim:**

This narrative logic involves minimizing the complexity and humanity of the target or victim of harmful actions. It often entails portraying the victims in a one-dimensional manner, reducing them to mere stereotypes or symbols rather than individuals with agency and humanity. By oversimplifying the victims, the perpetrators may justify or rationalize their actions, framing the violence as necessary or justified. This reduction and oversimplification serve to dehumanize the

victims, making it easier for the perpetrators to distance themselves emotionally from the harm inflicted upon them.

## 2. Portrayal of Harmful Action as Unavoidable or Beyond Control:

This narrative logic involves framing the harmful actions as inevitable or beyond the control of the actor. By presenting the violence as predetermined or driven by external forces, the perpetrators may absolve themselves of responsibility or culpability for their actions. This narrative tactic can serve to deflect blame and rationalize the violence as an unavoidable consequence of circumstances or historical forces beyond the perpetrator's control. It can also serve to normalize or naturalize the violence, making it seem like an inevitable part of the social or political landscape.

## 3. Permission to Harm:

In this narrative logic, the actor is portrayed as having explicit or implicit permission to inflict harm upon others. Hajari offers a justification for their violence by inflaming the horrors previously faced by the perpetrator. Legitimize or sanction violence against certain groups or individuals. By framing the actor as having authorization or justification for their harmful actions, this narrative logic can serve to reinforce power dynamics and hierarchies, perpetuating cycles of violence and oppression.

In stark contrast to Hajari in Butalia's work, violence is portrayed through three distinct narrative lenses:

### 1. Intimidation Narrative:

This narrative lens portrays violence as spontaneous, emotional, and lacking in premeditation or control. In this approach, violent acts are depicted as erupting suddenly, often driven by intense emotions such as anger, fear, or frustration. Butalia provides vivid and detailed descriptions of the violence, emphasizing its graphic and visceral nature. Through this narrative, readers are confronted with the raw brutality and

immediate impact of the violent acts, which may leave a lasting impression. By focusing on the emotional intensity and unpredictability of violence, Butalia highlights its disruptive and chaotic nature, challenging any notions of rationality or justification.

## 2. Moral Narrative:

In this narrative framework, the emphasis is not solely on the violence itself or its function within a social or economic context, but rather on the ethical concerns surrounding the act. Butalia explores the moral implications of violence, questioning the ethical boundaries crossed and the consequences for both perpetrators and victims. This approach invites readers to reflect on the broader ethical and philosophical questions raised by acts of violence, such as issues of justice, accountability, and responsibility.

## 3. Survivor Narrative:

The survivor narrative centers on the experiences and perspectives of those who have endured violence and its aftermath.

Butalia here allows her interviewees to share their stories, struggles, and resilience in the face of adversity. Butalia highlights the agency and resilience of survivors, portraying them not just as passive victims, but as individuals who actively navigate and confront the challenges posed by violence.

### **3.4. The language of the Holocaust in narrating Partition in *The Midnight's Furies* and *The Other Side of Silence***

The integration of oral narratives into the study of Partition focuses on examining testimonies, recollections, and the lasting impact of generational trauma. This approach is influenced by methodologies and investigations found in the fields of memory studies and genocide studies. This borrowing is logical, as it enriches the discourse by providing a broader vocabulary and facilitating a deeper comprehension of the Partition. Comparing and contrasting with other traumatic events, such as the Holocaust, is insightful as it offers valuable insights into the



complexities of human experience during times of conflict and displacement. By analyzing these parallel narratives, scholars can gain a more comprehensive understanding of the various dimensions of the Partition, shedding light on marginalized voices and overlooked aspects of history.

In "The Other Side of Silence," Butalia was among the first to draw parallels between the Holocaust and Partition. She made reference to James Young's 1988 publication, "Writing and Rewriting the Holocaust," which explored the representation and reinterpretation of Holocaust narratives. This connection highlights Butalia's pioneering efforts to link the experiences of mass violence and displacement in different historical contexts. Young asked, "How can we know the Holocaust except through the many ways in which it is handed down to us? ... we know the Holocaust through its literary, fictional, historical, and political representations and through its personal, testimonial representation," because "how people remember those facts, and how they represent them" are equally important. Butalia employed Young's framework as a model for her own investigation into the overlooked narratives within the Partition story. She found this approach particularly beneficial because it allowed her to spotlight narratives that were frequently sidelined, depicting the experiences of individuals who were typically invisible in mainstream historical accounts. These narratives frequently centered around everyday experiences, yet they carried profound importance, particularly in relation to the perspectives of women, children, and individuals from marginalized communities, such as those belonging to scheduled castes. Each story, despite its apparent ordinariness, provided invaluable insights into the diverse and often overlooked experiences of these groups during Partition. By adopting this methodology, Butalia aimed to amplify the voices of marginalized groups and provide a more inclusive portrayal of the Partition's impact on diverse communities. The analysis of LIWC features identifies the transformation in language and usage is evident in "The Other Side of Silence," where there is a noticeable shift from portraying individuals

solely as victims to emphasizing their resilience as survivors. Additionally, the term "account" is replaced with "testimony," highlighting the significance of personal narratives in bearing witness to historical events. Similarly, the focus moves from mere recollection of events to a more deliberate act of remembrance, acknowledging the enduring impact of memory. Furthermore, the portrayal of suffering evolves from simply acknowledging pain to recognizing the lasting psychological effects of trauma. This linguistic transition mirrors the terminology commonly employed in the field of memory studies, particularly in relation to the Holocaust, highlighting the broader comprehension of Partition and its enduring impacts on both individuals and communities.

Similar to the trajectory observed in the field of memory studies, Butalia's approach to oral history also aligns with concepts such as Marianne Hirsch's "generation of post-memory" and Ananya Kabir's idea of "post-amnesia." In this framework, the initial generation represents the primary repository of memories, while subsequent generations inherit and preserve the transgenerational impact of trauma. Within Butalia's work, this notion parallels Hirsch's argument concerning the transition of Holocaust memory guardianship and the evolution of transmitted knowledge into historical narrative or myth. The usage of various terms such as "absent memory," "inherited memory," "belated memory," "prosthetic memory," "received history," and now "post-memory" by the author reflects the complexity of these inherited histories. Butalia Writes:

“over years, its [Partition’s] memories have become more complex, acquired more nuance and layers, and been seen differently, depending on the particular circumstances of the moment of remembering” (Butalia, 2020).

In "The Midnight Furies" by Hajari, the analysis of LIWC features identify that there are explicit references to the Holocaust era, and the author draws parallels between the events of Partition and those

of the Holocaust. Hajari's comparison extends to various aspects, including likening the rule of the Indian National Congress (INC) to Hitler's regime, drawing parallels between the violence of Partition and that of the Holocaust, and equating Jinnah's call for Pakistan to Hitler's advocacy.

Moreover, Hajari incorporates terminology from Holocaust literature to describe the brutality of Partition violence. For instance, he uses phrases such as "Death Camps," "Death March," "Genocide," "Ghetto," "Dehumanizing," and "Displacement." By employing this language, Hajari underscores the severity and inhumanity of the violence and upheaval experienced during Partition. These terms evoke images and connotations associated with the Holocaust, highlighting the magnitude of suffering and loss endured by individuals and communities affected by Partition violence.

It is noteworthy that while contemporary interpretations of Partition often draw upon the language and frameworks of memory and genocide studies, there is relatively less integration of Partition scholarship within accounts of genocide. Concepts such as ethnic cleansing and genocide, commonly associated with events like the Holocaust, may not always be as applicable in understanding the complexities of Partition. Despite the endeavors of researchers such as Ian Talbot, Anders Hansen, Ishtiaq Ahmed, and Paul Brass to integrate the Punjab's ordeal into the wider discourse on genocide, prominent works on genocide like "Centuries of Genocide" and "Explaining Ethnic Cleansing" do not delve into the topic of Partition.

This discrepancy raises questions about the appropriateness of retroactively applying the language and frameworks of holocaust studies to the violence of Partition in Punjab in 1947. Unlike many instances of genocide where perpetrators and victims are clearly delineated, the violence during Partition was often characterized by fluidity, with individuals sometimes occupying both roles as victims and perpetrators. Additionally, the motivations behind the violence during Partition were

multifaceted, stemming from a complex interplay of historical, political, and social factors. As Butalia Writes in *The Other Side of Silence*

“So much depends on who remembers, when, with whom, indeed to whom, and how. But... the way people choose to remember an event, a history, is at least as important as what one might call the facts of that history, for after all these latter are not selfevident givens; instead, they too are interpretations, as remembered or recorded by one individual or another”

## References

- Blee, K. (2012). Interviewing activists. In D. Snow, D. della Porta, B. Klandermans, & D. McAdam (Eds.), *Blackwell Encyclopedia of Social and Political Movements* (Vol. II, pp. 623–26). Oxford: Blackwell.
- Blee, K. M., & Taylor, V. (2002). Semi-structured interviewing in social movement research. In B. Klandermans & S. Staggenborg (Eds.), *Methods of Social Movement Research* (pp. 92–117). Minneapolis: The University of Minnesota Press.
- Blouin, F. X. Jr., & Rosenberg, W. G. (2011). *Processing the Past: Contesting Authority in History and the Archives*. New York: Oxford University Press.
- Butalia, U. (2000). *The Other Side of Silence: Voices from the Partition of India*. Duke University Press.
- Butalia, U. (2018). *The Other Side of Silence: Voices from the Partition of India*. Duke University Press.
- Butalia, U. (Ed.). (2015). *Partition: The Long Shadow*. Oxford University Press.
- Clemens, E., & Hughes, M. D. (2002). Recovering past protest: Historical research on social movements. In B. Klandermans & S. Staggenborg (Eds.), *Methods of Social Movement Research* (pp. 201–27). Minneapolis: The University of Minnesota Press.
- Cobb, S. (2006). A developmental approach to turning points: "Irony" as an ethics for negotiation pragmatics. *Harvard Negotiation Law Review*, 11(147), 147-197.
- Collins, L., & Lapierre, D. (1975). *Freedom at Midnight*. Simon & Schuster.
- Cox, R. J. (2004). *No Innocent Deposits: Forming Archives by Rethinking Appraisal*. Oxford: Scarecrow Press.
- Dahlstrom, M. F. (2010). The role of causality in information acceptance in narratives: An example from science communication. *Communication Research*, 37, 857-875.
- Dahlstrom, M. F. (2012). The persuasive influence of narrative causality: Psychological mechanism, strength in overcoming resistance, and persistence over time. *Media Psychology*, 15, 303-326.
- Derrida, J., & Prenowitz, E. (1995). Archive fever: A Freudian impression. *Diacritics*, 25(2), 9. <https://doi.org/10.2307/465144>

- Dill, B., & Aminzade, R. (2007). Historians and study of protest. In B. Klandermans & C. Roggeband (Eds.), *Handbook of Social Movements Across Disciplines* (pp. 267–311). London: Springer.
- Frisch, S. A., Harris, D. B., Kelly, S. Q., & Parker, D. C. W. (2012). Introduction: Taking the road less traveled. In S. A. Frisch, D. B. Harris, S. Q. Kelly, & D. C. W. Parker (Eds.), *Doing Archival Research in Political Science* (pp. 1–32). Amherst, NY: Cambria Press.
- Gilmartin, D. (2015). The historiography of India's Partition: Between civilization and modernity. *The Journal of Asian Studies*, 74(1), 23–41. <https://doi.org/10.1017/S0021911814002134>
- Hajari, N. (2017). *Midnight's Furies: The Deadly Legacy of India's Partition*. Amberley.
- Iggers, G. G. (2005). *Historiography in the 20th Century: From Scientific Objectivity to the Postmodern Challenge*. Middletown, CT: Wesleyan Press.
- Jha, S. (2017). Book review: Nisid Hajari, *Midnight's Furies: The Deadly Legacy of India's Partition*. *Indian Journal of Public Administration*, 63(2), 313–318. <https://doi.org/10.1177/0972262917701018>
- Kelly, R. (2012). Archival research in composition studies: Re-imagining the historian's role. *Rhetoric Review*, 31(4), 461–478. <https://doi.org/10.1080/07350198.2012.711201>
- Khan, Y. (2007). *The Great Partition: The Making of India and Pakistan*. Yale University Press.
- Magliano, J. P. (1999). Revealing inference processes during text comprehension. In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative Comprehension, Causality, and Coherence: Essays in Honor of Tom Trabasso* (pp. 55–76). Mahwah, NJ: Lawrence Erlbaum.
- McCullagh, C. B. (2000). Bias in historical description, interpretation, and explanation. *History and Theory*, 39(1), 39–66. <https://doi.org/10.1111/0018-2656.00116>
- Meehan, J. (2010). Rethinking original order and personal records. *Archivaria*, 70, 27–44. <https://archivaria.ca/index.php/archivaria/article/view/13294>
- Nair, N. (2009). *A History of Partition: Violence, Borders and Memory*. Cambridge University Press.
- Pandey, G. (2001). *Remembering Partition: Violence, Nationalism, and History in India*. Cambridge University Press.

- Pearce, S. (1992). *Museums, Objects, and Collections: A Cultural Study*. Leicester: Leicester University Press.
- Rose, S. K. (1999). Preserving our histories of institutional change: Enabling research in the writing program archives. In S. K. Rose & I. Weiser (Eds.), *The Writing Program Administrator as Researcher: Inquiry in Action and Reflection* (pp. 107–18). Portsmouth, NH: Boynton/Cook.
- Schwartz, J., & Cook, T. (2002). Archives, records, and power: The making of modern memory. *Archival Science*, 2(1), 1–19.  
<https://doi.org/10.1007/BF02435628>
- Sengupta, D., & Bose, S. (Eds.). (2019). *The Partition of Bengal: Fragile Borders and New Identities*. Cambridge University Press.
- Talbot, I., & Singh, G. (2006). *Divided Cities: Partition and Its Aftermath in Lahore and Amritsar, 1947-1957*. Oxford University Press.
- Trabasso, T., & Sperry, L. L. (1985). Causal relatedness and importance of story events. *Journal of Memory and Language*, 24, 595-611.
- Zamindar, V. F.-Y. (2007). *The Long Partition and the Making of Modern South Asia: Refugees, Boundaries, Histories*. Columbia University Press.

## Chapter 4

### **‘Master Narratives in (Re-En) Action’: 1947 Partition in School Textbooks and Social Media Narratives**

*“Hindustan had become free. Pakistan had become independent soon after its inception but man was still a slave in both these countries -- slave of prejudice ... slave of religious fanaticism ... slave of barbarity and inhumanity.”*

— Saadat Hasan Manto

Textbooks are the fundamental and most extensively utilized educational resource across various disciplines and academic tiers, offering structured and organized knowledge essential for comprehensive learning. Their role extends beyond mere dissemination of facts; textbooks serve as roadmaps, guiding students by providing context, analysis, and exercises to deepen understanding. In classrooms worldwide, they are revered companions, empowering educators and learners alike with a shared foundation for intellectual exploration. Textbooks hold significant importance in education, serving as conduits not solely for disseminating factual information, but also for imparting ideas and cultural norms. The textual and visual content encountered by children in educational settings plays a crucial role in shaping their attitudes and beliefs as they transition into adulthood. These textual and visual elements not only convey concepts but also constitute integral components of the educational journey, influencing the formation of ideas and perspectives. Textbooks exert considerable influence in educational environments as a result of numerous factors. Initially, they are meticulously chosen for official adoption by duly authorized entities or governmental bodies, thus conferring upon them a sanctioned status. Furthermore, textbooks are introduced to students under the auspices of authoritative figures, namely their educators, imbuing them with a sense of legitimacy and credibility. Additionally, textbooks possess a universal reach, as they are mandated reading material for every student, thereby amplifying their impact. Perceived by many as repositories of absolute truth, textbooks assume a position of paramount importance in shaping



the educational landscape. Moreover, owing to the limited experiential reservoirs of children, they are particularly susceptible to the influences propagated through educational materials. Given the pivotal role textbooks play in the educational milieu, the onus for ensuring their quality is substantial. Moreover, in light of the uncritical acceptance often accorded to textbook content, the imperative for rigorous critique by scholars and educators is heightened significantly. The phenomenon of distortion and bias within textbook authorship is not a novel occurrence, rather it dates back to the inception of historical documentation. Despite contemporary endeavors to prioritize factual accuracy and mitigate perceived credibility deficits, these issues persist unabated. Thus, the necessity for fresh investigations aimed at Clarifying this persistent dilemma becomes evident as biased content within textbooks serves to distort the perception of reality for children belonging to both the dominant and oppressed societal groups. In such inquiries, due consideration must be afforded to both formal and informal educational encounters in historical pedagogy, situated at the confluence of collective remembrance and scholarly historiography.

LaRoque (1975, p. 64) introduces six categories of inequitable treatment discernible within educational textbooks. These encompass:

1. Linguistic Partiality: Manifesting disparity through linguistic constructs.
2. Obsolescence Omission: Failure to incorporate contemporary information, thereby perpetuating outdated assertions.
3. Historiographical Exclusion: Neglecting pivotal facets of historical narratives.
4. Blame Bias: Accentuating deficiencies of the 'other' while embellishing the merits of one's own cohort, often intertwined with personal biases.
5. Superficial Recognition: Granting perfunctory or token acknowledgment to the contributions of minority groups.

6. Unilateral Narrativization: Propagating singular perspectives devoid of nuanced representation or counterbalancing viewpoints.

The entire media landscape has been profoundly impacted by the emergence of new information and communication technologies following the broader accessibility of the Internet to the general population. The utilization of social media platforms as a primary source of news dissemination is witnessing a notable surge, particularly among the demographic of younger users. Nevertheless, the reliance upon platforms such as Facebook, Twitter, and analogous mediums for news consumption is not devoid of inherent challenges, as evidenced by the doubling share of social media engagements originating from unreliable news sources observed between 2019 and 2020. In light of the dynamic nature of social media platforms, news items transcend mere informational content, becoming deeply entrenched within the interactive discourses of users. This phenomenon holds particular significance in instances of inaccurate or overtly misleading information, given that user engagement significantly influences the uncritical propagation or skepticism thereof. Empirical studies have underscored the influence of biased social media coverage on individual decision-making processes, yet the extent to which users are cognizant of encountered biases and their subsequent responses remains a topic of inquiry. This aspect assumes heightened relevance as user reactions not only serve to contextualize social media reporting for fellow users but also possess the potential to either alleviate or exacerbate the impact of biased narratives. Watts et al. argue that biased content generation on social media, even if factually accurate, presents a greater challenge to democratic systems compared to outright misinformation because of its widespread influence.

This chapter is divided into two sections where the first section titled “Partitioned Histories: Portrayal of 1947 Partition in Select School Textbooks from India and Pakistan” studies the linguistic manifestation of biases in chapters on the 1947 Partition in school texts prescribed in

India and Pakistan while the second section titled “Recollection and Transmedial Reiteration: Twitter and the New Epistemology of 'Subjugated knowledges' about Partition” studies the Tweets on the declaration of the Partition Horrors Remembrance Day using the developed computational tool.

## **1.1.SECTION - I**

### **1.1.1. Partitioned Histories: Portrayal of 1947 Partition in Select School Textbooks from India and Pakistan**

Textbooks are the most widely used educational media that facilitate a structure and order to the teaching and learning process. According to historical records, they existed as long as the formal schooling system and systems of writing have existed in the forms of clay tablets, bound sheets of papyrus, scrolls, parchment, vellum, or modern print textbooks. The print textbooks used for mass schooling were first produced in Europe and they expanded their reach to encompass the global populace, following the patterns of colonization and cultural and technological adaptations. When the colonies attained Independence and Self- governance, they replaced the colonial textbooks with the ones reflecting their national ideologies. The traditional genre of history textbooks also emerged in this context during the early twentieth century. With the rise of a national form of education, they were endorsed to develop the notion of citizenship, loyalty, and collective identity. They also became the agents in the determination and assimilation of official and collective memory.

Textbook plays a key role in Indian and Pakistani educational systems owing to the peculiar ways in which the process of teaching and education is perceived. Krishna Kumar describes this as ‘textbook culture’ (Kumar,2005) where teaching, learning, and assessment are based only on the prescribed textbooks. Both countries have a nationalistic approach in teaching history to develop an identity agnate to the national ideology. Pakistan has maintained a steadfast commitment to forging a distinct identity for itself, deliberately

distancing from its historical connections to United India. The Partition, widely regarded as one of the most profoundly impactful events in the histories of both India and Pakistan, has been depicted in a rather subdued manner within the school textbooks of both nations. Students on either side of the border encounter fragmented depictions and divergent narratives of a shared historical past. India perceives the Partition as a "crisis," viewing it as a rupture in the "secular" fabric of society (Kumar, 2001). Conversely, Pakistan commemorates the Partition, interpreting it as a moment of "liberation" from the Hindu majority.

Lately, several analysts have undertaken commendable efforts to analyze the contents of historical instructional materials to elucidate the intended and expressed goals and intentions of policymakers, textbook compilers, and educational authorities. Indian historians and educationalists like Neeladri Bhattacharya, Romila Thapar, Bipan Chandra, Krishna Kumar, and others and Pakistani scholars like Mazhar Abbas, Pervez Hoodbhoy, Abdul Hameed Nayyar, Rubina Sajol, Ayesha Jalal and several others have written extensively on the textbooks from India and Pakistan. Works such as "The Contested Terrain of Textbooks: A Case Study of India and Pakistan" by Krishna Kumar analyze the comparative portrayal and remembrance of key historical events in the textbooks of India and Pakistan. Similarly, edited volumes like "Imagining the Nation: History Textbooks in India and Pakistan" by Sugata Bose and Ayesha Jalal, as well as "Textbooks, Identity Politics and Social Cohesion in India, Pakistan and Sri Lanka" edited by Meenakshi Thapan, provide a comprehensive examination of various aspects of these textbooks, including their content, language, and pedagogical methods. Sylive Guichard (2012), Devika Mittal (2022), and Kusha Anand (2023) have researched the formation of historical narratives and nationalism linked with the textbooks used in India, meanwhile, Dhananjay Tripathi (2018) and Joshi S (2010) have examined the depiction of India in Pakistani textbooks, and vice versa.

Meenakshi Chhabra (2015) has studied the memory practices associated with Partition in Indian and Pakistani texts.

The absence of studies scrutinizing the portrayal of the Partition in school textbooks in both India and Pakistan, alongside inquiries into possible linguistic biases embedded within these educational resources, represents a significant gap in both textbook research and Partition scholarship within these nations. This void inhibits a comprehensive understanding of how historical events are narrated and interpreted within educational contexts, as well as the potential impact of such narratives on shaping collective memory and national identities. Identifying this gap, this section analyses Punjab Textbook Board Social Studies and Pakistan Studies textbooks from Grade IV to Grade VIII, and Grade IX to Grade X, respectively, and NCERT textbooks from Grade VI to XII using the Semi-Automated tool to identify and mitigate narrative and memory biases in textual data. With the aid of computational analysis, this section aims to understand

1. How the divided historical memory of Partition is described in the select textbooks?
2. How linguistic patterns of biases are manifested in these textbooks?
3. How textbooks are perceived in understanding the history of Partition?

This section of the chapter does not aim to advocate for a fixed and/or singular interpretation of Partition in either of the two nations while disregarding diverse narratives. Instead, it seeks to highlight the existence of overarching narratives, with textbooks serving as just one example of these narratives. The "omnipotence" attributed to textbooks can lead to the belief that their contents are beyond questioning. It often implies that the only valid knowledge is that which is what is reiterated in textbooks, while other avenues of knowledge acquisition and education are relegated to a secondary status. Children are conditioned to perceive history as solely what is written in their history textbooks, with no regard for the historical experiences of their grandparents

(Pathak, 2002). Central to this discussion is the importance of questioning, redefining, and reevaluating the limitations of knowledge that present themselves as dominant narratives, immutable identities, and impartial reflections of truth (Giroux, 1992).

### **1.1.2. Teaching History in School Classrooms**

“History education is a site where teachers and pupils as members of distinct generations engage with textbooks and other materials as specific forms of memory texts that guide what should be passed on to the younger generation” (Macgilchrist et.al, 2015). Teaching history involves the dynamic interplay between recounting past events and engaging with collective memory, facilitating the conveyance of knowledge, shaping identity, and fostering the advancement of historical analytical skills. Within the context of this chapter, the term "history" pertains to a documented account of bygone events rather than the lived reality of those times. Scholars and educators in the realm of textbooks and education acknowledge that the act of teaching history is intricate and multifaceted citing that:

1. Learning objects in history often escapes immediate understanding as they are situated within the contemporary context of the present.
2. The past continually expands, and the historical domain perpetually enlarges.
3. The field of history is consistently influenced by digital transformations.
4. A crucial role of history is to stimulate historical thinking, which is described as an 'unnatural act' (Wineburg, 2001).
5. History serves the dual functions of mediating and shaping individual and social identities.
6. The complexity intensifies when history is imparted to schoolchildren, demanding attention to their distinct cognitive requirements.

When academic disciplines make the transition to school texts, an epistemological "sameness" overtakes what are distinct fields of study (Brown, Collins, & Duguid, 1989; Wineburg, 1991). Thus, was coined the pejorative term "textbookese." According to the American Historical Association, quality history textbooks should distill the available knowledge on major subjects within the discipline. Additionally, they should offer exercises tailored to the appropriate level for students. While embracing a variety of methodologies is encouraged, textbooks at every educational level should adhere to particular standards. The primary principle, which serves as the foundation to achieve this, is to ensure that the content enriches the teaching of history effectively while maintaining scholarly rigor.

### **1.1.3. History of School Textbooks in India and Pakistan**

Upon gaining independence, both India and Pakistan inherited a centralized and authoritative system for formulating educational policies and molding the curriculum, as control shifted from the British colonial administration to nationalist leaders in these two countries. By the late 19th century, the colonial state had initiated the use of published textbooks in the educational system, alongside some expansion of primary and secondary education. The rise of what Krishna Kumar has labeled as the 'textbook culture' during colonial India resulted in a scenario where educators frequently prioritized ensuring that the material contained within officially sanctioned textbooks was thoroughly addressed in the classroom. In essence, the textbook became the cornerstone of the curriculum, overshadowing other potential sources of knowledge and pedagogical approaches. The textbook was considered a powerful source of knowledge, to the extent that some ambitious students expressed a desire to "grind the texts into a pulp and extract the knowledge out of them and drink it" (Tandon, 1968). Naturally, this emphasis on textbooks also made the textbook industry a profitable business for publishers.

#### **1.1.4. History of NCERT Textbooks in India**

The print textbook culture in India took root during the 19th and 20th centuries, particularly under colonial rule. British officials such as Alfred Lyall and Vincent Smith, along with colonial historiographers, played a pivotal role in shaping these early textbooks. The colonial-era textbooks were characterized by an explicit division between 'self' and 'other,' with the former being portrayed as superior and the latter as negative and inferior (Kumar, 2022). This division reflected the ideological interests of the British and their interpretations of Indian history.

The significance of textbooks as political tools persisted in post-independence India, adopting the colonial ways of engaging with them. In 1961, the National Council of Educational Research and Training (NCERT) was founded with the mandate to supervise the production of textbooks nationwide. The principal objective of NCERT was to develop exemplary textbooks that could be uniformly recommended for use in schools across India, ensuring that all citizens had access to the same syllabus, texts, narratives of the past, and information (Bhattacharya, 2009).

This standardization effort extended to the state level with the establishment of State Councils of Educational Research (SCERT) to ensure uniformity and integration. The Education Commission's report (1964-66), commissioned by the Government of India, emphasized the need for a common textbook. It envisioned a set of books authored by the best scholars, translated for accessibility to every school, with the belief that such a shared educational foundation could significantly contribute to national integration (Education Commission, 1964-66, 230).

The new textbooks produced by NCERT in the 1960s and 70s, written by eminent historians like Satish Chandra, Bipan Chandra, R. S. Sharma, and Romila Thapar had mainly two objectives:

1. they attempted to decolonize the existing textbooks.



2. they tried to embody the secular, democratic, and socialistic ideologies after the communal carnage during the Partition

The efforts of homogenisation of textbooks were countered in multiple ways like

1. insertion of local and community histories
2. Transculturation through vernacularization
3. Private publishers producing their versions of the textbooks

To date, there have been three consecutive sets of NCERT textbooks: the initial series, which is the above-discussed series, was developed in the late 1960s and throughout the 1970s. Subsequently, new textbooks were introduced from 2002 to 2004, coinciding with the ascendance of the coalition government led by the Bharatiya Janata Party (BJP) at the national level. This second series of textbooks sparked significant controversy. Following the publication of the second set of history textbooks, the Indian History Congress (IHC) conducted an analysis and brought out a catalog of discrepancies. In reference to the textbook on Ancient India for class VI (written by Lal et al. 2003), the IHC made the following observation: ‘It is peculiar that neither the caste (varna/Jati) system nor the dharma shastra texts are referenced in the chapter on Hinduism’ (IHC 2003: 24); and concerning the one on Medieval India for class XI (Jain 2002): ‘It is noteworthy that this book does not contain any mention of the treatment of Dalits or of untouchability under the caste system, as if the Dalits have no place in our legacy’ (IHC 2003: 70–1).

Following the victory of the Congress-led coalition in the 2004 general elections, a third series of textbooks was implemented. These specific educational materials have been utilized in schools since 2008. NCERT periodically revises and updates its textbooks to incorporate changes in educational content, pedagogy, and curriculum guidelines. As a result, there are multiple editions of NCERT textbooks for different subjects and classes. Typically, NCERT textbooks go through significant

revisions every few years to ensure that the content remains relevant, up-to-date, and in line with evolving educational standards and requirements and new editions are published to reflect these changes.

#### **1.1.5. Controversies Waged Against NCERT History Textbooks in India**

“Production of Textbooks became a site of negotiation and political battles” (Bhattacharya, 2009) in India since the inception of a homogenized educational system. The major disputes and controversies around textbooks were raised around the questions of

##### **1. temporality and periodization:**

Historians crafting textbooks in the 1960s leaned towards organizing Indian history into ancient, medieval, and modern eras, but colonial orientalist promoted a periodization framework characterized by three epochs: Hindu, Muslim, and British rule. However, beyond the issue of periodization, there were deeper underlying problems. In the 1970s, the textbooks consistently criticized and dismantled the dualistic notion of a "glorious past for Hindu civilization" juxtaposed with a "Muslim period (1200-1800) characterized by degeneracy and tyranny". This stance sparked controversy, particularly among the Hindu right.

##### **2. Origin and Indigeneity:**

The issue of origin and indigeneity is intertwined with the assertions of pure Aryan descent within the discourse of Hindu historical politics. The timeline of the Aryan past and the Indus Valley civilization frequently sparks debate.

- 3. communal history**
- 4. Controversies Over Historical Interpretations**
- 5. Regional and State-Specific Issues**
- 6. Inclusion of Controversial Figures**
- 7. Cultural and Religious Sensitivities**

These controversies often revolve around issues related to the content, historical interpretations, and political ideologies presented in the textbooks. NCERT has periodically revised its textbooks in response to these issues to maintain objectivity, accuracy, and inclusivity in the educational content.

#### **1.1.6. Changes in NCERT History Textbooks Across Ages**

The domain of historical education in India, particularly concerning teaching materials for school students, has been a battleground between historians with different ideological leanings since the 1960s. Following the emergency period in India, a coalition government took power in New Delhi in 1977. Reacting to pressure from this coalition, one of the first moves made by this administration was the removal of NCERT school textbooks from distribution, a decision that sparked protests. Subsequently, in 1999, when the Hindu right-wing political faction assumed power, textbooks became a focal point of intense controversy once again. Specific passages, particularly those referring to Ancient India, were expunged due to concerns about offending Hindu sensitivities. Entire textbooks were recalled and substituted with a fresh set of materials that, for the first time in Indian educational history, integrated numerous communal concepts that had previously been restricted to popular media or imparted in communal educational institutions.

Despite these ongoing disputes, the content of history textbooks remained largely unchanged for nearly four decades, with only subtle alterations. Yet, during a short interval from 2002 to 2004, when the Bharatiya Janata Party (BJP) commanded a majority, there was an effort to "saffronize" the educational programs and learning materials. This "saffronization" aimed to promote a Hindu nationalist interpretation of history, aligning with the strategies of Hindu nationalists on the right-wing, sometimes referred to as Hindutva. However, concerns arose regarding the Modifications declared by NCERT in the school syllabi in January 2002, particularly the perceived disregard for the examination

of the medieval era in Indian history. The allocation of space and attention to ancient and medieval India in the curriculum reflected a perceived disparity in importance. Additionally, A unit within the social science curriculum focusing on the characteristics, dissemination, and core principles of "major religions" omitted "Islam". These decisions were seen as politically and ideologically motivated. Furthermore, even when the chapter "Advent of Islam" was added in the social science textbooks for the subsequent grade, it was positioned within a section alongside the history of West Asia. This placement was justified by Murli Manohar Joshi, the 'Minister of Human Resources and Development' (HRD), who argued that Islam "emerged from that region - its history is intertwined with the history of Arab civilization, its dissemination, and its emergence." The BJP government was voted out of power in 2004, and the subsequent 'United Alliance (UA) Government' took steps to reverse the textbooks. In 2004, textbooks resembling those used before the 2002 changes were brought back into circulation.

In tandem with UA government backing, NCERT spearheaded a substantial overhaul of the school curriculum in 2005. This comprehensive revamp saw the introduction of fresh textbooks in 2008, breaking a nearly four-decade-long tradition. The 2008 curriculum emphasized employing disciplines such as sociology, ecology, linguistics, and literary analysis as avenues for self-discovery within the framework of interpersonal relationships. It aimed to provide students with opportunities to explore differing perspectives rather than simply memorizing information from the textbook. In this framework, history textbooks for secondary education were tailored to foster students' abilities to form their own interpretations of historical events.

The 2005 revision of the curriculum aimed to bring forth a nuanced perspective on the concept of a nation. It sought to highlight the diverse array of historical narratives and the history of everyday life. While these updated materials refrained from offering a drastic critique of the nation-state, they did pose challenges to the idea of a singular,

unproblematic national identity in varied patterns. Specifically, in the lower secondary levels (grades 6, 7, and 8), the curriculum restructured the study of "Ancient, Medieval, and Modern India" into three segments titled "Our Pasts I, II, and III." This renaming was intended to underscore the existence of numerous historical perspectives and narratives that collectively shape our understanding of the past. It emphasized the absence of a sole linear narrative or a unified national history that individuals could easily adhere to. A significant aspect of these textbooks was their approach, which seamlessly transitioned from exploring specific local histories to delving into the broader histories of entities and regions. This approach wasn't merely about acknowledging local variation; it was about recognizing that the overarching narrative of the nation unfolds on the backdrop of these local sites. The objective was to recognize that diverse communities and regions possess unique historical narratives that cannot be fully encompassed within the overarching national narrative.

Further in 2022-23, NCERT has dropped key history developments from textbooks including chapters and sections on 'The Mughal empire' (Class XII), 'Gujarat riots' (Class XII), 'Rise of Popular movements' (Class XII), 'Hindu extremists and Mahatma Gandhi' (Class XII), 'controversies regarding Emergency' (Class XII), 'Era of One Party Dominance' (Class XII), 'The Cold War Era' (Class XII), 'US Hegemony in World Politics' (Class XII), 'Central Islamic Lands' (Class XI), 'Clash of Cultures' (Class XI), 'Industrial Revolution' (Class XI), 'India After Independence' (Class VIII), 'Democracy and Diversity' (Class X), 'Popular Struggles and Movements' (Class X), and 'Challenges to Democracy' (Class X). The reasons suggested for these changes are that of reducing content load due to COVID and the transition phase to National Education Policy (NEP) 2020.

The trajectory of historical education in India, particularly reflected in NCERT textbooks, has been marked by a series of ideological and political battles spanning several decades. From the

controversies surrounding the portrayal of historical events during different political regimes to the more recent revisions purportedly justified by stating to be aimed at aligning with the objectives of the National Education Policy (NEP) 2020 and accommodating challenges posed by the COVID-19 pandemic, the landscape of historical education continues to evolve. The continued progression of these educational resources underscores the imperative for rigorous examination and critical analysis of historical narratives, necessitating a concerted effort to ensure that pedagogical materials embody a comprehensive and multifaceted comprehension of India's historical trajectory.

#### **1.1.7. History of Textbooks in Pakistan**

During the initial Education Summit held in Pakistan soon after its formation, the education minister articulated the government's preference for an educational system that would prioritize the Islamic character of the newly established nation, created as a home for the Muslim population of the region. Despite variations in governance and their principles over time, this focus on accentuating the Islamic ethos has remained a foundational aspect of the educational curriculum in Pakistan.

In 1959, under the administration of General Ayub Khan, the establishment of the "National Commission for Education" decentralized the management of primary education, assigning greater authority to provincial governments across Pakistan. This initiative resulted in the consolidation of subjects like history and other social sciences, which were merged into "Social Studies" for grades 1 to 8 and "Pakistan Studies" for grades 9 to 12. Despite the purported decentralization, the central or federal government in Pakistan continues to directly supervise educational policies and curricula.

Textbooks undergo development within the oversight of provincial "textbook boards." Despite their official authority, these boards find their autonomy constrained due to their role as "ideological gatekeepers," as observed by Nayyar and Salim. Beyond mere

administrative oversight, they hold the crucial responsibility of meticulously filtering educational content to align with specific ideological frameworks. This duty extends beyond factual accuracy, delving into the realm of ideological conformity, wherein classroom materials must adhere strictly to predefined national and ideological standards. Thus, the textbook boards serve as custodians, diligently ensuring that only content deemed ideologically acceptable according to prevailing norms and beliefs is disseminated within the educational system.

The turning point in the evolution of school textbooks in Pakistan towards a more overtly ideological stance is often traced back to the 1977 coup orchestrated by General Zia-ul-Haq. This juncture marked a significant shift in educational policy and content. The transformation gained further momentum with Zia's introduction of the New Education Policy, a sweeping initiative that sought to redefine the educational landscape of the nation. Central to this policy was the emphasis on cultivating 'a deep and unwavering loyalty to Islam' as a cornerstone of the school curriculum. Under Zia's regime, the educational framework underwent a profound overhaul, with the

infusion of religious ideology permeating various facets of the curriculum. Subjects ranging from history to literature were reshaped to align more closely with Islamic principles and values. The influence of Zia's policies extended well beyond his tenure, leaving an indelible mark on Pakistan's educational system. Although Benazir Bhutto may not have been as fervently committed to the Islamization of education as General Zia-ul-Haq, critics assert that her administration's attempts to eradicate ideological content from the educational system were neither meticulously orchestrated, nor were they given paramount importance. Furthermore, these efforts were not subjected to thorough examination or rigorous scrutiny. As a result, despite Bhutto's leadership, the ideological underpinnings of the educational curriculum persisted to a significant extent. Ideological forces deeply entrenched in the education system successfully resisted attempts to alter the curriculum or

textbooks. Under the leadership of Nawaz Sharif, the government extended significant support, openly advocating for the Islamization of the school curriculum. When General Pervez Musharraf's administration assumed power, it promised a comprehensive overhaul of the education system. Despite this assurance, during Musharraf's tenure, the long-term objectives articulated in Nawaz Sharif's National Education Policy continued to exert influence, guiding the trajectory of the educational system. As a consequence, there was a notable absence of comprehensive reform in the infrastructure responsible for the production of textbooks. This continuity in policy direction highlights the enduring impact of ideological agendas on educational reforms in Pakistan, underscoring the challenges in effecting substantial change within entrenched institutional frameworks. Following the recommendations of the Sharif Commission Report, Textbook Boards were entrusted with a crucial mandate: overseeing the comprehensive process of planning, compiling, and producing textbooks catering to students from Grade I to Grade XII. This responsibility extended beyond mere logistics; it also encompassed ensuring that the educational content aligned effectively with the prevailing government policies and directives. To execute this task efficiently, the boards convened a team of seasoned authors tasked with the meticulous compilation and creation of the textbooks. However, the journey from conception to publication was far from straightforward. Once the initial drafts were prepared, they underwent rigorous scrutiny, revision, and enhancement by a panel of additional subject matter experts. This iterative process aimed to refine the educational materials, enhancing their quality and relevance to meet the evolving needs of the student population. But the scrutiny didn't end there. Before final approval and dissemination, the "ideological" content within the textbooks underwent a meticulous review by the National Review Committee within the Federal Ministry of Education. This committee, comprising esteemed scholars and experts, ensured that the content aligned closely with the ideological framework set forth by the government. Finally, upon receiving the committee's endorsement, the textbooks were published by the Provincial Government and designated



as the official curriculum materials. This designation conferred upon them a significant status, cementing their role as the primary educational resources for students across the province. Thus, the process from conception to publication involved multiple layers of review and approval, underscoring the thorough attention paid to shaping the educational landscape through the production of textbooks.

#### **1.1.8. Portrayal of Partition in NCERT (Indian) Textbooks**

The initial series of NCERT history textbooks expressed feelings of sorrow and displacement about the 1947 Partition, highlighting how difficult it was to leave behind one's homeland. The chapter mentioning Partition in the initial series reads as follows: "The dream of Indian unity had been shattered and brother had been torn from brother." Within that chapter, the Indian National Congress was hailed for its fervent patriotism, while culpability for communalism and the Partition of India was assigned to the British authorities and Mohammad Ali Jinnah.

Preserving the sovereignty of India was a prime focus of the narration. In this particular series of textbooks, the depiction of Partition and its associated collective memory as an entrenched historical truth serves to stifle the development of independent thinking among children. By presenting a singular narrative, divergent perspectives, and critical analysis are discouraged, constraining students' ability to engage with the complexities of history. Furthermore, only one sentence within the entire lesson acknowledged the profound brutality experienced by ordinary people during Partition. The chapter reads as follows, "[E]ven at the very moment of freedom, a communal orgy, accompanied by indescribable brutalities was consuming thousands of lives in both India and Pakistan." The absence of oral histories and the solitary mention of the horrors of Partition accentuate the inadequacy of the portrayal presented in these textbooks. By neglecting firsthand accounts and diverse perspectives, students are deprived of the opportunity to engage with the lived experiences and realities of Partition. Furthermore, the implication within the text that the communal intent of the Muslim

League was the sole instigator of violence oversimplifies the complex socio-political dynamics at play during that period. Such a reductionist perspective fails to acknowledge the multifaceted factors contributing to violence and division, thereby perpetuating a biased and incomplete understanding of history.

In the third set of textbooks, a new lesson titled "Understanding Partition - Politics, Memories, and Experiences" was incorporated for higher secondary school students. The content within this section is derived from various primary historical manuscripts, alongside oral testimonies and prevalent media forms such as literature and cinema. Diverging from the earlier versions, in this textbook, Partition is not portrayed merely as a communal conflict instigated by the Muslim League and Jinnah. Instead, the chapter delves into the historical intricacies, and aftermath of the events surrounding Partition, examining various social and political catalysts that precipitated the event. The chapter reads as follows,

Yet it would be incorrect to see Partition as the outcome of a simple unfolding of communal tensions. ... Communal discord happened even before 1947 but it had never led to the uprooting of millions from homes. ... Partition was a qualitatively different phenomenon from earlier communal politics, and to understand it we need to look carefully at the events of the last decade of British rule

Further, for the first time, a school textbook incorporates personal anecdotes from Hindus, Muslims, and Sikhs pertaining to Partition. Within these accounts, individuals assume roles as instigators, victims, and saviors across all three communities. The aim of the chapter is to cultivate within students a disciplined comprehension of the complexity and diversity of historical occurrences. The chapter's objective is articulated as follows:

This chapter will examine the history of Partition: why and how it happened as well as the harrowing experiences of ordinary

people during the period 1946-50 and beyond. It will also discuss how the history of these experiences can be reconstructed by talking to people and inter-viewing them; that is, through the use of oral history.

Although introducing new perspectives about Partition is not spelled out as one of the outcomes, potentially, teachers could use the multiple perspectives presented in the chapter to raise critical questions about existing memory narratives of Partition in their classrooms. While the topic of Partition is rarely addressed in current Indian textbooks, other than the above-mentioned chapter, the tenth-grade textbook "Democratic Politics II" briefly touches upon it within the context of communalism in Chapter Four. This section discusses the adverse

effects of communal politics and refers to the suffering experienced by minorities in India and Pakistan during the Partition era. However, in other sections of the book dealing with political science, there is a deliberate effort to avoid direct mention of Pakistan. This cautious avoidance of teaching about Partition and Pakistan, despite the undeniable significance of Partition in Indian history and the substantial border India shares with Pakistan, might be explained as an attempt to handle a sensitive and controversial element of India's history and its current political dynamics. Nevertheless, the somber and unsettling narrative of Partition within Indian history warrants attention within school classrooms. These conversations are essential for acquainting young minds with the history of Partition in a manner that provides impartial and precise historical insights. Avoiding detailed discussions about Partition perpetuates viewpoints that foster cultural animosity. Neglecting this topic entirely raises concerns that knowledge about Partition history may be conveyed to younger generations through informal channels such as peer discussions, media coverage, films, television programs, literary works, and even biased political sources, all of which are abundant on this subject. Although these sources are vital for educating and influencing people's perspectives on intricate

subjects, they cannot supplant the structured acquisition of knowledge through educational materials and classroom instruction.

#### **1.1.9. Portrayal of Partition in Punjab Textbook Board (Pakistani) Textbooks**

In M.D. Zafar's Pakistan Studies textbook in the chapter on Partition and creation of the nation of Pakistan, the 'two-nation theory' is portrayed as natural, asserting that Hindus and Muslims have always been distinct nations within the subcontinent, to argue that the textbook states as follows: "the Ideology of Pakistan which developed through a long period of 1000 years was materialized in 1947 when . . . a new Muslim State, Pakistan, appeared on the map of the world". Contrary to implying that the demands of the Muslim League were somehow supported by colonial administrators, Zafar depicts the Congress and the British as colluding to deprive Pakistan of its rightful position, a conspiracy that was only foiled by "the grace of Allah [and] the wisdom and sagacity [of] Quaidi-Azam [the title given to Jinnah in Pakistan] and his followers". Pakistani youth are educated with the belief that Pakistan is unique; it acknowledges a shared history with India but has consistently retained its identity as a distinct nation. Consequently, the prevailing narrative asserts that Pakistan's existence predates the Muslim League's insistence on an independent Muslim homeland. While this "original" assertion finds limited acceptance among historians outside of Pakistan, it nonetheless serves as the cornerstone of Pakistan's historical narrative, as portrayed in Pakistan Studies textbooks

A thorough discussion of the events surrounding the Partition of 1947 is provided due to their significance in shaping the national narrative of the emergence of Pakistan as a liberated state from its precursor, India. The establishment of Pakistan is briefed as follows neglecting the events preceding it.

The establishment of an independent state in the name of Pakistan in the subcontinent is the most important event of the

twentieth century. There was a strong ideology at the root of this incident ... The ideology of Pakistan is based on the Islamic ideology. Beliefs, worship, rule of law, brotherhood, equality and justice are elements of Pakistan's ideology

The text further reads as follows:

As the days of gaining independence from the British approached, the Hindus strongly opposed the idea of the establishment of Pakistan. Under these circumstances, the unparalleled leadership of Quaid-e-Azam Muhammad Ali Jinnah compelled the British and the Hindus to accept the demand for the establishment of Pakistan. However, the Hindus did not wholeheartedly acknowledge the existence of Pakistan

The Punjab text board book describes the event of Partition as follows justifying the advocacy for the Two- Nation Theory citing the reasons for demanding a separate nation based on religious beliefs.

After winning their freedom, they (Muslims) wanted to establish a government in which they could live in accordance with Islam, where every law would be in accordance with Qur'an. But they knew that the Hindus were in a majority in India. After the British leave, they would not let an Islamic state be established here. They would establish a rule of the Hindu law rather than that of the law of the God. In this law, Muslims would be treated as untouchables. They feared that after getting rid of the slavery of the British, they would become slaves of the Hindus. The Muslims wished to have a true freedom, in which only Muslim would rule, there would be the rule of the Law of Allah.

The excerpt from Pakistani textbooks illustrates a perspective that highlights instances of violence during historical events. Specifically, it portrays a narrative where Muslims in Pakistan are depicted as providing assistance to those wishing to leave, while people in India are accused of committing atrocities against Muslim refugees during the Partition.

Additionally, the mention of the 1857 War of Independence reflects a narrative in Pakistani textbooks that suggests Hindus employed deceptive tactics to persuade the British that only Muslims were involved in the revolt. This portrayal aligns with the idea of the "Two-Nation Theory," a concept that emphasizes the irreconcilable differences between Hindus and Muslims in terms of cultural practices, history, language, cuisine, music, and attire. The chapter argues that the "Two-Nation Theory" forms the basis for the creation of "Muslim Pakistan" and "Hindu India."

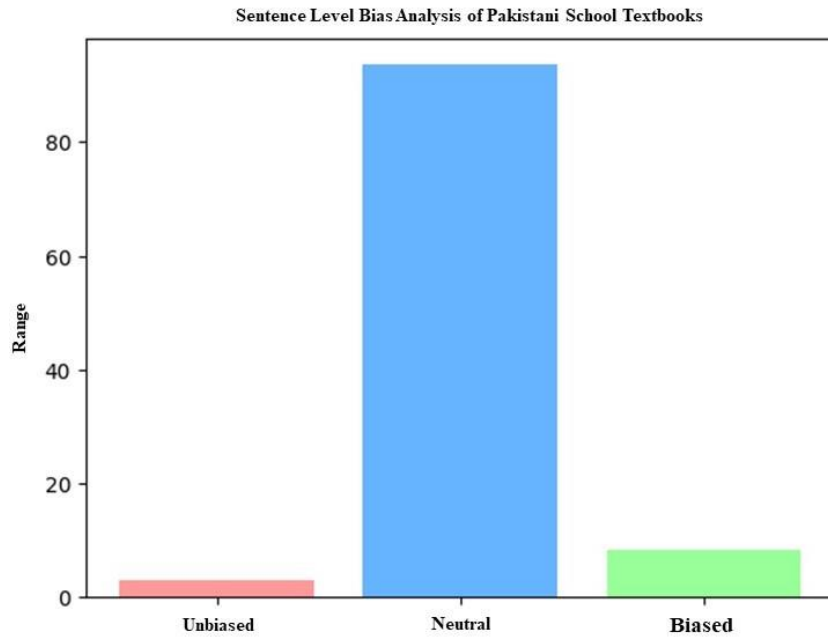
It's noteworthy that the textbooks appear to avoid elaborating on the differences between the two communities, possibly out of concern that such details might reveal commonalities that could challenge the foundational premise of the "Two-Nation Theory." This highlights the strategic use of historical narratives in shaping and reinforcing particular ideological perspectives, which in this case, underpins the distinct identity and separation of Muslim Pakistan and Hindu India. Educational materials, including textbooks, play a pivotal part in molding national narratives and identities, and the chosen narratives can significantly influence how individuals perceive historical events and their broader cultural context.

Analyzing linguistic biases within the textbooks mentioned above is indispensable for scholarly inquiry on Partition. Linguistic biases wield significant influence over the portrayal of historical events, thereby shaping students' comprehension of intricate historical narratives. By examining the linguistic constructs employed within these textbooks, the biased portrayal of Partition that perpetuates specific ideological or political inclinations can be identified. Furthermore, linguistic biases can affect the inclusivity and fidelity of narratives in the textbooks, especially in the case of Partition of 1947. Linguistic biases often mirror underlying cultural hegemony, underscoring the exigency of scrutinizing language constructs within educational materials.

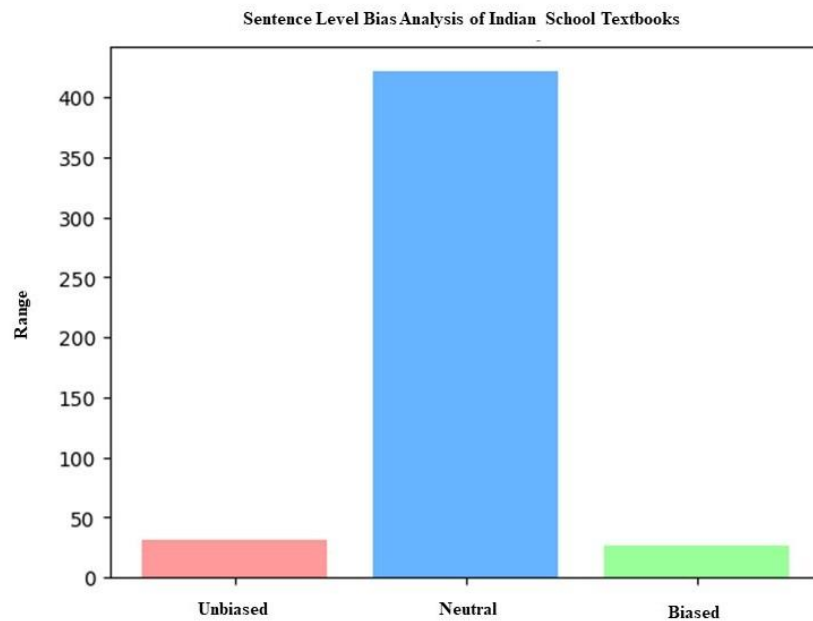
#### **1.1.10. Linguistic Bias Analysis of Chapters on Partition in Indian and Pakistani School Textbooks**

There have been considerable changes in the nature and content of history textbooks in both India and Pakistan. They provide students with a distended, linear, and non-controversial compilation of facts in well-produced illustrated materials. Despite these meliorations, inaccurate commissions, omissions, and more subtle forms of linguistic biases are identified by the tool in these textbooks. Unfairness exists here through the use of language, neglect of certain information, and othering by enabling only one side of the event.

Figures 4.1.10.1 and 4.1.10.2 show the range of biased, unbiased, and neutral sentences in chapters on Partition in Pakistani and Indian books respectively. There is no significant difference in the ratio of these three categories in both texts. There are a high number of neutral sentences as these texts mostly deliver factual information about the events. The analysis of sentences in chapters on Partition in Pakistani and Indian books reveals a comparable distribution of biased, unbiased, and neutral statements across both texts. This finding suggests that there isn't a significant difference in the ratio of these three categories between the Pakistani and Indian narratives. A notable observation is the prevalence of neutral sentences in both sets of texts. This abundance of neutral statements can be attributed to the predominantly factual nature of the information presented regarding the events surrounding Partition. Given the historical significance and complexity of Partition, authors likely prioritize providing objective accounts and factual details, resulting in a higher frequency of neutral statements. While biases may naturally exist in historical narratives, the relatively balanced distribution of biased statements across both sets of texts suggests a concerted effort by authors to mitigate overt biases and present a balanced perspective on Partition.



*Figure 4.1.10.1: Sentence-level bias analysis of Chapters on Partition in Pakistani (Punjab Board )School Textbooks*



*Figure 4.1.10.2.: Sentence-level bias analysis of Chapters on Partition in Indian (NCERT) School Textbooks*

The tool shows a high rate of positive sentences in the textbooks suggesting a Linguistic positivity bias. This bias is due to the writer's attempt to confirm earlier societal perceptions of Partition delineating objective circumstances and affective states. This in turn results in a high rate of confirmation, which is confirmed by the results of the tool. The narration in both the textbooks is victim-centered and an affect-inducing



language is used in describing hardships. Both narratives entail the other the responsibility for the event, which is suggestive of the negative attribution bias in the results.

It is important to note that linguistic valence can significantly influence younger adults engaging with educational content to both intellectually engage with and precisely remember its content. Tool identifies through the LIWC feature analysis that NCERT textbooks specifically have a higher passive voice construction. Previous research infers that content presented in passive voice constructions is less likely to be remembered by readers. (Turner & Rommetveit, 1968). Further, the studies of Waller and Henley et al have found that When using passive voice constructions, individuals are less inclined to attribute responsibility to perpetrators for their violent acts and to be emotionally affected by the narratives (Waller, 2002; Henley et al., 2005). The bias-inducing pattern of both textbooks is a result of a specific historical narrative pattern that Mark Phillips (2013) terms as 'historical distance'. According to Philips Historical narratives can evoke either a sense of approximation, fostering "warm encouragement," or a sense of distancing, fostering "deliberate estrangement," depending on factors such as demeanor, timeframe, and authors' "normative calls for action" (Phillips, 2013,). Through this, writers have the ability to influence readers' emotional reactions, subtly instilling their ideological values into readers' perceptions (Phillips, 2013). The word clouds show the prominent bias-inducing words in both textbooks. It is important to note that most of the biased words are nouns which refers to nation, religion and political affiliation. But it is also significant that there is a very less precisely religious biased sentences, precisely 2% in Pakistani texts and .8%in Indian texts. Similarly there is 1.2% politically biased sentences in Pakistani texts whereas there is 1.6% in Indian Texts.



crafted to reinterpret Partition in a way that “abstract causal phenomena present a more intricate version of reality” (Coffin,2004) By doing so, the textbooks try to reveal the intricate forces at play behind historical events, giving readers a deeper understanding of the complexities involved in the Partition story.

The linguistic tools reflect this causal pattern of narration through the higher number of connectors, conjunctions, and causal verbs. This linguistic process of developing causation involves the strategic use of conjunctions such as "because" or "so," as well as nouns like "reason," "effects," and "response," alongside verbs such as "make" and "lead to," and prepositions including "for," "through," and "from," all of which contribute to constructing causal meanings within the narrative. Moreover, broader organizational strategies within the text, such as introductory generalizations and concluding summaries of information, also play a significant role in shaping causal reasoning. The structural arrangement of clauses within the text, including the way in which they are connected and positioned in proximity to each other, further contributes to this construction of causal relationships. Causation can be explicitly signalled as an inter-clausal relation through connectors and conjunctions. However, it can also be conveyed within a single clause through the utilization of causal verbs and other linguistic elements that impart agency to the narrative.

These Textbooks impose a linear or teleological understanding of causality, framing events as leading inevitably to certain outcomes or interpretations, resulting in a temporal bias that obscures contingencies in historical development. Further, the language used to describe causality attributes more significance to themselves while shifting blame to the other and/or certain actors or factors resulting in attribution Bias and consequently confirmation bias and ethnocentric bias as identified by the tool. In an effort to make Partition more accessible to school students, textbooks oversimplify causal relationships which in turn causes the simplification bias.

### **1.1.12. The Language of ‘Othering’ in Partition Textbooks and Manifestation of Biases**

A peculiar amalgamation of reflexive nationalism, as elucidated by Samaddar (1998, pp. 74–75), and reflexive otherness has emerged as a pivotal defining characteristic within the narrative especially the geopolitical perspectives presented in both Indian and Pakistani textbooks. Geopolitical vision, as defined by Dijkink (1996, p. 11), encompasses any conception pertaining to the interrelation between one's own and foreign territories, entailing sentiments of (in)security or (dis)advantage, and often invoking notions of collective mission or foreign policy strategy. Dijkink further posits that a geopolitical vision necessitates the delineation of a distinction between 'Them-and-Us,' alongside an emotional attachment to a specific locale, thereby closely intertwining with sentiments of national identity. This intricate interplay of reflexive nationalism and reflexive otherness within educational materials not only shapes perceptions of the self and the other but also plays a crucial role in constructing and reinforcing geopolitical visions. Such visions, rooted in notions of collective identity and territoriality, inherently involve delineating boundaries between the self and the other, often engendering feelings of (in)security or (dis)advantage. Moreover, these geopolitical perspectives serve as frameworks through which individuals perceive and interpret the external world, influencing foreign policy orientations and shaping broader national narratives.

The ensuing excerpt from the Class VI Social Studies textbook serves as a lucid illustration of this phenomenon.

“Children of Pakistan - The children of different regions of Pakistan are quite similar. All these children know the national language Urdu and can read and write it ... . “Children of India - In India, the Muslim children generally wear Shalwar Kameez or Kurta Pyjama. While Hindu children wear Dhotee. Children in North India like Chapati whereas the children of

Southern India like to eat rice. Muslim children like eating meat while Hindu children eat pulses and vegetables.”

Partition and conflict with the "separated other" is the organizing principle on which a variety of exclusions and inclusions were based in textbooks from India. Hindu vs its Other, the Muslim; unified India/Bharat vs its 'national other' Pakistan; each of which is defined by the existence of its other. The Sentences that display the language of 'othering' in these chapters on Partition are the major sources of confirmation bias, political bias, speculative bias, and geographical bias identified by the tool.

#### **1.1.13. Geographical Bias in Textbooks**

An initial attempt to overhaul Indian history textbooks from a Hindu nationalist viewpoint emerged in 1992 following the electoral victory of the BJP in India's most populous state, prompting a comprehensive revision of the history curriculum in secondary schools. This initiative resulted in the publication of the textbook titled "High School Itihaas," which garnered substantial criticism from liberal media outlets upon its release. The textbook opens with an assertion of the concept of 'greater India' (Vishaal Bharat), positing it as a naturally constituted entity. It then proceeds to bemoan the contemporary fragmentation of the country into three distinct political entities – India, Pakistan, and Bangladesh. A parallel notion of the 'natural nation' is discernible in Pakistan Studies textbooks. Within a Pakistan Studies volume intended for advanced-level students, M.D. Zafar communicates to students that while the formal establishment of the modern state of Pakistan took place in 1947, the 'spirit' of Pakistan has persisted 'for centuries.' This perspective underscores the assertion that 'Pakistan wasn't... born on August 14, 1947. It is as old as history, and nature has endowed it with a unique unity.' Furthermore, the fundamental notion that the entire subcontinent constitutes a unified entity also underpins the historical perspective of R.S. Sharma, the author of "Ancient India," a publication by the NCERT during a period characterized by its

liberalism. While elucidating the 'significance of Ancient Indian History' to students, Sharma articulates:

The ancients strove for unity. They looked upon this vast subcontinent as one land. . . . Our ancient poets, philosophers and writers viewed the country as an integral unit. . . . The kings who tried to establish their authority from the Himalayas to the Cape Comorin and from the valley of the Brahmaputra to the land beyond the river Indus in the west were universally praised.

The approach embraced by Indian nationalists reflects the sentiments echoed within the textbooks, as they steadfastly maintain the concept of India as an inherently evolving nation. Even in contemporary discourse, prevalent portrayals of India depict it as an organic entity, with its unity disrupted by the tumultuous 'tragedy' of Partition. In contrast, a comprehensive examination of Pakistani textbooks reveals a deliberate and concerted effort to cultivate a distinct regional identity that diverges from India's adherence to the overarching narrative of the subcontinent's natural geographical unity. Given the early emphasis placed on fostering an Islamic identity in Pakistan's educational curriculum, it is not surprising that students were encouraged to perceive their country's geography primarily through the prism of the Muslim world. This emphasis was further accentuated in textbooks produced during the Zia-ul-Haq era, where a conspicuous focus on Pakistan's Islamic identity became pronounced. For instance, in a Social Studies textbook tailored for Class Six, the opening chapter titled 'Location of Pakistan' underscores Pakistan's pivotal position within South Asia. However, aside from a perfunctory acknowledgment of neighboring countries such as India, China, and Russia, the chapter predominantly delves into situating Pakistan within the broader Islamic framework.

Subsequent volumes within the series authored by the same individual also echo this thematic trajectory, with the initial chapter titled 'Pakistan and the Muslim World.' The ensuing chapters delve into various facets of the Muslim world, including its physical characteristics, climate,

societal conditions preceding the advent of Islam, the impact of Islam on society, and the resurgence of the Muslim community. Even in Muhammad Arshad's foundational textbook on Pakistan Studies, Pakistan's geography is depicted through a lens that accentuates its connection and affinity with the Muslim world. The chapter reads as follows:

Bharat lies on Pakistan's eastern border . . . to its east is another chain of Muslim countries, Bangla Desh [sic], Malaysia, and Indonesia . . . further West, Pakistan is linked with the Middle East countries. To the West of these countries are situated the Arab countries of Arabia. This group of Muslim countries form a continuous Muslim block.

While these descriptions are not inherently inaccurate, it's not difficult to discern the geopolitical implications of such portrayals, particularly evident when the textbook forthrightly posits that, owing to its geographical location, "Pakistan occupies a pivotal position among the Muslim nations spanning across Southeast and Southwest Asia. Consequently, Pakistan is poised to wield considerable influence in addressing the intricate political dilemmas confronting the Muslim world". This assertion underscores a broader narrative of Pakistan's geopolitical significance within the Muslim community, suggesting that the nation's strategic placement endows it with a unique vantage point to engage in regional and international affairs. It implies a sense of responsibility and leadership thrust upon Pakistan as a central player in the Muslim world underscoring the country's aspirations for regional leadership and its desire to assert its influence on the global stage. This portrayal not only reflects a sense of pride in Pakistan's identity but also reinforces the country's geopolitical ambitions and its perception of its role in shaping the political landscape of the broader Muslim community.' It's quite evident that the emphasis is on connections with other Muslim nations and people, essentially 'spiritual' neighbors, rather than geographical ones. 'Bharat' (India) indeed occupies a prominent

position in discourses regarding Pakistan's contemporary state, albeit typically portrayed in an adversarial light. The segment titled 'Pakistan's Relation with Bharat' opens with the assertion: 'Following independence, Pakistan's primary existential threat emanated from Bharat.' In light of David Harvey's proposition regarding the centrality of territorial and regional conceptions in shaping consciousness, identity formation, and political subjectivity, the divergences evident in the conceptualization of the 'geo-body' within Indian and Pakistani educational materials carry profound implications. The nationalist perspectives entrenched in the pedagogical approaches guiding students to interpret their historical narratives not only limit the breadth of students' historical understanding but also actively instill the belief among learners from both India and Pakistan that they inhabit distinct and separate realms. This perpetuation of divergent national narratives fosters a sense of otherness and perpetuates the perception of mutual estrangement between the two nations, further complicating efforts toward reconciliation and fostering mutual understanding.

By situating the histories of these two nations within differing conceptualizations of geography, school textbooks facilitate, and even encourage, the construction of disparate narratives from a shared historical foundation. Perhaps the true tragedy lies in the fact that these textbooks endeavor to teach young individuals, who may be studying just around fifty miles apart, for instance, in Lahore and Amritsar, to perceive themselves not only as heirs to distinct pasts but also as residents of entirely separate worlds.

#### **1.1.14. Selective Amnesia of Conflicts and Violence and Linguistic Biases**

I was in high school when the anti-Sikh carnage in Delhi took place, the Shah Bano case was threatening to erupt and even a quick glance at a newspaper would raise troubling questions. But our textbooks did not equip us to ask those questions. There were intense conflicts in society, yet the textbooks presented a vision



of Arcadian harmony. In an age even more marked by information than our age of pre-TV innocence was, the yawning gap between textbooks and society will generate an even more corrosive scepticism. (Mehta, 2006)

It is not expected that textbooks will provide graphic and sensationalized descriptions of violent acts. Furthermore, the content provided is tailored to suit the age group of the students, resulting in variations in presentation across different educational levels. However, the examination of how Partition and violence are depicted in textbooks predominantly centers on those designed for students in Class XII, typically aged seventeen to eighteen. Within the first set of textbooks, the narrative predominantly revolves around a chronological sequence of political events stemming from decisions made by key political figures. In the third series of textbooks, riots are primarily portrayed as consequential outcomes of historical progression or the nation-state's evolutionary trajectory. The depiction of violence in both sets of textbooks is quite brief and abstract, providing a general overview rather than delving into specific details. As identified from the LIWC associated with the tool, the terms used to narrate the violence are 'hate-torn' / 'pain and sadness'/'senseless communal slaughter'/'communal frenzy' / 'cruelty'/'communal orgy'/'indiscriminate brutality'.

However, the textbook of the first series, written by Bipan Chandra, emphasized a shared accountability for violence borne by Hindus and Muslims:

the elation of coming independence was marred by the large-scale communal riots during and after August 1946. The Hindu and Muslim communalists blamed each other for starting the heinous killings and compete with each other in cruelty. Mahatma Gandhi, engulfed in gloom at this total disregard of elementary humanity and seeing truth and non-violence cast the winds, toured East Bengal and Bihar on foot to check the riots.

In contrast, the corresponding textbook of the second series attributes the responsibility for the pre-Partition violence primarily to the Muslim League. Despite the potentially traumatic and foundational nature of the Partition violence, narratives within both the initial and subsequent sets of textbooks portray it as an aberration and deviation from the envisioned trajectory of the nation. Within the Class XII textbook of the initial series, the history of modern India culminates shortly after achieving independence. Bipan Chandra briefly touches upon significant events such as the population movement between India and Pakistan, the ensuing violence, Gandhi's assassination, and the drafting of the Constitution in the final two pages of the textbook. The concise treatment of these events suggests that they are not meant to detract the reader from the principal narrative -the realization of the nation-state. The textbook concludes with the following sentence: "With confidence in their capacity and their will to succeed, the people of India now set out to change the face of their country and to build a just and good society and a secular, democratic and egalitarian India." (Chandra 1971, 2004: 273) The narrative depicted in the textbooks of the second series adopts a more subdued tone, recognizing that the ultimate goal of complete attainment was not entirely realized with the attainment of independence. The envisioned nation-state remains a work in progress. As highlighted by Mittal (2003: 261), "[e]ven today, about a third of Kashmir territory remains illegally occupied by Pakistan, which it calls as [sic] Azad Kashmir." This acknowledgment reflects an awareness of the ongoing challenges and unresolved issues that persist within the national framework. When examining the portrayal of violence in textbooks, it is pertinent to observe that both secular and Hindu nationalist narratives briefly touch upon the violence associated with Partition. However, the Hindu nationalist perspective exhibits less hesitancy and places greater emphasis on the significance of the movements and individuals who employed violence in the struggle for independence. This emphasis reflects a particular ideological stance that

seeks to valorize the role of those who resorted to force as a means to achieve national liberation.

The killing and uprooting of millions of people receive no more than a few lines of clerical mention in the textbooks of both countries. Between the paragraph about the Mountbatten Plan and the Indian Independence Act, most Indian school historians find no space to devote to the mass killing, rapes, and forced migration that took place between 15 August 1947 and the assassination of Mahatma Gandhi five and a half months later (Kumar, 2001).

### **1.1.15. Understanding the Conceptualisation of History Textbooks in Remembering Partition**

There are two distinctive approaches to the conception of history textbooks among textbook researchers. The first categorization where textbooks are a paramount instrument in crafting the broad social consensus, is conceived as the ‘autobiographies of nations’ (Jacobmeyer, 1998). The second category that are objects of vehement political grapple over the architecture of a ‘common-sense reality’ (Apple, 2000) marking the frontiers for what can be legitimately said. This category of textbooks does not perpetuate unbiased knowledge rather they patronage the construction of a ‘selective tradition’ (Williams, 1961). Here, textbook authors assume the role of a ‘truth giver’ (Crismore, 1984) or a ‘surveyor’ (Myskow, 2017) speaking with an anonymous authoritative voice excursively constructing a ‘disinterested impartiality’ (Myskow, 2017) building a rhetoric of surety. Similarly, In the context of teaching and learning Partition, textbooks can be perceived as:

#### **1. Textbook as a Memory Practice**

Textbooks, which have been defined as ‘weapons of mass instruction’ (Ingrao 2009), are imperative in ‘sustaining mnemonic hegemonies’ (Christophe, 2019). In the case of Partition Textbooks are a site of remembering but “memory (whichever concept we use from the current range including collective memory, cultural memory, social memory, connected memory, prosthetic memory, multi-directional

memory, traveling memory, and entangled memory) is a site of political contestation, subject formation, power struggle, knowledge production, and community-building” (Macgilchrist et.al, 2015). Remembering Partition implies politicized interpretations of past conflicts and is evident in personal and collective spaces. The memory of Partition is passed down from one generation to the next, through family anecdotes, broader societal narratives, and textbook and classroom enactments, creating a master memory narrative. Discussions, disputes, and conflicts regarding the creation of school textbooks encompass a battle for the production and regulation of collective memory. “The art of memory for the modern world is both for historians as well as ordinary citizens and institutions very much something to be used, misused, and exploited rather than something that sits inertly there for each person to possess and contain” (Said, 2000)

## 2. Textbook as a ‘Boundary Object’ (Star et.al, 1989)

The notion of 'boundary objects,' as elucidated by Susan Leigh Star and James Griesemer in their 1989 article, proposes that the establishment and administration of these entities are essential for fostering and sustaining coherence among interconnected social spheres (Star et al., 1989). Partition textbooks can be conceptualized as 'boundary objects' acting as hypothetical linkages between different types of knowledge, diverse regions, and imagined communities. As Binnenkade (2005) notes, textbooks function as mediators between worlds of thought and practice, accommodating various traditions of needs and duties.

In the context of Partition textbooks, viewing them as 'boundary objects' is particularly apt. These textbooks not only navigate the boundaries between hegemonic and non-hegemonic knowledge but also exert influence over these boundaries. They function as discursive catalysts, shaping conversations across time, space, and the politics of inclusion, exclusion, and identity. By presenting historical narratives, they contribute to the construction of collective memory, influencing

how individuals and communities perceive the events of Partition and their implications.

The term 'boundary objects' emphasizes the dynamic and interactive nature of these educational materials. They play a role in mediating different perspectives, historical interpretations, and cultural contexts, serving as a bridge between diverse social worlds. Understanding Partition textbooks in this way acknowledges their multifaceted impact on shaping collective understanding and memory, transcending the boundaries that exist in the realms of knowledge, regions, and imagined communities.

### 3. Textbooks as National Ideology

[T]he teaching of history, more than any other discipline, is dominated by textbooks. And... textbooks are boring. The stories that history textbooks tell are predictable, every problem has already been solved or is about to be solved. Textbooks exclude conflict or real suspense. They leave out anything that might reflect badly upon national character (Leowen, 1995).

It's essential to recognize that textbooks are created, designed, and written by actual individuals with motives, as Lowen writes about American school textbook authors “Something about the enterprise of writing a high school American history textbook converts historians into patriots” (Leowen, 1995). Textbooks introduce students to the prevalent socio-economic order within its intricate filiation with power, making them active sites of ideological and socialistic discourse favorable for the nation-state. Several preordained factors decide the political economy of textbook publishing including national curriculum frameworks, teaching pedagogies, and assessment protocols.

Factual knowledge undergoes a two-stage process of transformation into approved curriculum knowledge intended for teaching in schools, such as :

1. The selection of factual historical knowledge that favors and/or contradicts the national official knowledge is selected, diluted, and reproduced according to the cultural, economic, and social workbench of the Nation.
2. According to the demands and needs of the textbook publishers which are mostly economical rather than intellectual, the selected text is further altered, streamlined, and/or succented.

Past is often presented as a “morality tale” where students are effortlessly discouraged to raise questions making them believe that ‘history is facts to be learned’ (Loewen, 1995). Textbooks mostly emphasize citizenship, patriotism, abiding by rules and authority, and respect for individual rights where pupils are not encouraged to deviate from the general lifestyle of the state.

In both India and Pakistan, history textbooks were noted as particularly significant in fulfilling the function of national cohesion. The 1969 Report of the Committee on School Textbooks emphasized the pivotal role of history textbooks in fostering 'national unity,' guiding the selection of episodes to be incorporated into the renewed national narrative.

In view of the fact that in teaching Indian history in the past, during the British rule, stress has often been laid on religious differences and conflict, it is necessary that we should now highlight those situations – and they are legion – where people of all religious faiths have worked together in unity and cooperation. [ . . . ] This would, obviously, involve a creative and purposeful reinterpretation of history and a judicious selection of historical truths. The interests of national unity and the needs of a modernizing society should be the primary consideration in our choice and presentation of material. (Government of India, Ministry of Education 1969: 12)

The portrayal of Partition within official textbooks in both India and Pakistan typically aims to showcase the nation and its people in a

positive light. However, beneath this surface narrative, subtle or sometimes explicit messages can be identified, reflecting certain political agendas—a phenomenon referred to by Jackson (1968) as the "hidden curriculum." This hidden curriculum, embedded within the official knowledge presented in textbooks, serves as a conduit for specific ideologies to be transmitted to the younger generation, influencing their understanding of which political forces are to be embraced or marginalized within the collective consciousness of the nation (Taylor, 2004). Taylor's (2004) concept of the "social imaginary" is particularly insightful when examining school textbooks' portrayal of Partition. These textbooks play a pivotal role in shaping the collective understanding of "common practices and a widely shared sense of legitimacy" around Partition that are deemed essential by newly established governments. In essence, textbooks serve as potent political instruments employed by authorities to mold societal norms, values, and historical narratives, thereby influencing the formation of a nation's identity and collective memory.

#### *4. Textbooks as reinforcements of Homogeneity*

"How a society selects, classifies, distributes transmits, and evaluates the educational knowledge it considers to be public reflects both the distribution of power and the principles of social control". (Bernstein, 1971). This perspective is particularly relevant when examining Indian and Pakistani textbooks, as they are deeply rooted in the societal, philosophical, and political influences of dominant factions. The primary purpose of these textbooks is to establish and perpetuate cultural uniformity, fostering common perspectives and shared historical recollections. The ultimate goal is to create a collective hegemony, wherein influential groups guide society intellectually and morally, gaining governance not through force but by leading a society in which subordinate sects unknowingly endorse and adhere to prevailing cultural standards. School textbooks serve as instrumental tools in disseminating and reinforcing these predominant cultural norms.

‘Textbooks offer an obvious means of realizing hegemony in education ... Within history texts ... the omission of crucial facts and viewpoints limits profoundly the ways in which students come to view history events.’ (Griffin & Marciano, 1979 p. 35). This omission, along with ambivalence or incongruity in texts, is indicative of the 'politics of incorporation' (Apple, 2000). This political process serves to stabilize hegemony by silencing the voices of excluded minorities, authoritatively delineating and imposing categorical distinctions on students in their mediated encounters with the past. The strategic selection and presentation of narratives in textbooks are driven by the powerful elite's decisions, aiming to emphasize certain perspectives, neglect or exclude others, and sometimes reinterpret, dilute, or distort historical events to align with popular belief systems. This control over educational content contributes to the shaping of collective memory and influences societal perceptions of history, reinforcing the prevailing power structures and cultural norms.

Through the process of selective tradition, what is considered legitimate knowledge is constructed, and this knowledge is elevated to a position of authority, territory, and resources. This is achieved by including certain elements in the textual narrative while excluding others, effectively granting official approval to one group's cultural knowledge. The process of shaping social representation, historical recollection, and collective identity in this manner results in the establishment of cultural rifts within society. Often, what remains overlooked are the rich and diverse discourses and narratives that could arise from alternative historical perspectives. As Apple suggests, this oversight extends to what he terms 'mentioning,' wherein only limited and isolated aspects of the history and culture of less powerful groups find their way into the educational texts. Consequently, the full spectrum of experiences and voices within society is not adequately represented, perpetuating cultural gaps and marginalizing certain narratives. Anderson (1991) posits that textbooks play a crucial role in fostering an imagined community by promoting inclusivity. However, social



imaginaries frequently emerge through deliberate acts of segregating or silencing certain individuals or groups within society. Balibar (2002) emphasizes the significant influence wielded by the state in shaping national identities. In his view, exclusion becomes an inherent aspect of constructing a national community of citizens. This exclusionary process often involves marginalizing or overlooking the contributions and experiences of particular segments of the population, thereby reinforcing divisions within society. Althusser (1971) argued in "Ideology and Ideological State Apparatuses" that schools, like other state institutions such as the church and the military, function to ensure submission to the "ruling ideology or the mastery of its practice". According to Foucault (2003), schools and educational institutions are viewed as "apparatuses of domination," while Deacon (2006) characterizes them as a "disciplinary response to the need to manage growing populations."

## **4.2.SECTION – II**

### **4.2.1. Recollection and Transmedial Reiteration: Twitter and the New Epistemology of 'Subjugated knowledges' about Partition**

Since its establishment in 2006 X, formerly known as Twitter, has been a notable platform for the dissemination of real-time information and news, and social networking. The platform has been crucial in shaping global conversations and instigating political movements and digital activism. India ranks as Twitter's third-largest market, boasting 24 million users, following the United States and Japan. In India Twitter has become both, a tool for help as well as a tool for propaganda. While during COVID people have used Twitter for crowdsourcing help, bypassing the conventional lines of communication, the platform has also been used to communalize and polarise societies. Political and ideological groups have social media wings for controlling the narrative that affects them. They actively propagate hateful rhetoric through organized 'troll farms' inducing panic or staging hoaxes. But the misuse of Twitter does not only happen

around organized groups. The potential anonymity that Twitter can offer enables individuals to spread hate. Twitter, not being oblivious to this has set up a 'Trust & Safety Council' along with active monitoring of accounts "to ensure that people feel safe expressing themselves on Twitter". Irrespective of this in India, propagandists have devised an effective strategy for disseminating their message on Twitter without incurring any costs primarily by manipulating the trending column through targeted hashtag campaigns. At least ten hashtags trending on Twitter in India are the outcome of organized campaigns. These campaigns provided individuals with tweet templates and encouraged them to post identical tweets to boost their visibility. Over 50% of these tweets were duplicates, with many appearing to be directly copied from the provided tweet templates.

It is also alarming how the Government of India removes tweets under "Section 69A of the Information Technology Act" which provides provision for blocking public access to information when "the Central Government or any of its officers specially authorized by it in this behalf is satisfied that it is necessary or expedient so to do, in the interest of sovereignty and integrity of India, defense of India, security of the State, friendly relations with foreign States or public order or for preventing incitement to the commission of any cognizable offense relating to above" ( Information Technology Act (IT Act) 2000). Over the past four years, the government's practice of blocking hashtags, Twitter accounts, and innocuous content under significant legal provisions such as the "sovereignty and integrity of India" has raised valid concerns about the transparency and inherent secrecy of the blocking process. Between 2017 and 2020, Twitter, on average, received a Section 69A order approximately once every three months. The Lumen database filings provide a general overview of what the government seeks to censor on Twitter. Complete hashtags like #KashmiriUnrest or #KashmirNow were subjected to blocks, and tweets addressing India's cow vigilantes were directed to be taken down. In reporting about the 1947 Partition Twitter has become a 'third

Space' of cyber-spatial perception where 'subjugated knowledge' of the event is irradiated and continuously transmitted, negotiated, and appropriated. Twitter thus becomes an archive and site of multidirectional memories.

There are growing studies on biased content on Twitter and computational ways of mitigating them recently. In 2020 Mozafari et.al proposed a “BERT model for Hate speech detection and racial bias mitigation” in Twitter (Mozafari et.al, 2020), following which Tankard et.al used sentiment analysis to define an approach to analyze an individual's tweets to identify potential biases (Tankard et.al, 2021). Chen et.al used neutral social bots “to probe distinct biases emerging from platform mechanisms versus user interactions” on Twitter (Chen et.al, 2021). 2023 saw much more attention towards bias analysis on social media platforms as Xiao et.al published their methodology for “detecting political biases of named entities and hashtags” on Twitter using “the Polarity-aware Embedding Multi-task learning (PEM) model” (Xiao et.al, 2023). Similarly, Alba et.al suggest establishing an analytical framework for identifying biases or prejudices in the content extracted from Twitter messages (Alba et al., 2023).

However, there are only a few research that studies Indian Twitter space. In a 2022 study, Dash et al. analyzed the role of influencers in amplifying “political polarization” in India, by quantifying the partisan engagement of Twitter handles (Dash et.al, 2022). Similarly, Bhat et.al proposes a qualitative textual analysis of tweets to study mob censorship in India (Bhat et.al, 2023), and Borah and Singh investigate political polarization in India by employing network analysis and content analysis methodologies, focusing specifically on Twitter data. (Borah &Singh, 2022).

Identifying the above-mentioned gap, this section of the chapter aims to understand the nature of discourses and the manifestation of biases in Tweets about the 1947 Partition on 14 and 15 August 2021 around “Partition Horrors Remembrance Day”. The study with the help

of the computational tool further attempts to comprehend the Perception of Twitter as a site of contention and false memory formation, a mnemonic medium, and a Creator and manipulator of collective consciousness. The section also looks into how recollection occurs within a social network's collective memory and how Twitter's transmedial structure Aids it.

Following the ethical guidelines this section keeps the identity of Tweep anonymous and focuses on the discourse. As Twitter's API has been under a paywall since February 2023, the study couldn't be extended further or couldn't access the history of the tweets from 2021.

#### **4.2.2. Partition Horrors Remembrance Day**

"Partition Horrors Remembrance Day" was announced on August 14, 2021, by the Prime minister of India stating that

"Partition's pains can never be forgotten. Millions of our sisters and brothers were displaced and many lost their lives due to mindless hate and violence. In memory of the struggles and sacrifices of our people, 14<sup>th</sup> August will be observed as Partition Horror's Remembrance Day" (Narendra Modi, Twitter, August 14, 2021)

The Ministry of Culture, Government of India has also published a detailed report briefing the significance of remembering Partition. However, this short document has significant biased content and misinformation. The report terms 'British Government as a mediator rather than an arbitrator' in Partition almost justifying the colonizer. The document has chosen newspaper reports from 1947 highlighting the sufferings of Hindus and it barely mentions the mass killings of non-Hindus. On page 13 of the report the Newspaper reports titled 'Muslims Pass Through Amristar Safely' and 'Hindu Village in Ambala Distt Attacked' are placed alongside. Similarly, on pages 11, 16, and 19 there are news articles on attacks on Hindus and Sikhs, disregarding the violence inflicted on Muslims on either side of the borders, rather there is an attempt to portray them as the perpetrators of the violence. This

disparity is clear in portraying the abuse and abduction of women and children. While highlighting the pain of displacement, the report cherishes the nostalgia of ‘unified Bharat’.

After the declaration and publishing of this report, Pakistan’s Spokesperson Zahid Hafeez Chaudhri rebutted saying “No modern state is so much in contradiction with itself as the Indian state – the so-called ‘largest democracy’”. As this report was to be circulated amongst educational institutions in India there were also minor disagreements against it across the nation and a few states refused to carry forth it further.

#### **4.2.3. Portrayal of Partition in Tweets**

Declaration of Partition Horrors Remembrance Day saw a rise in the number of Tweets around Partition. A lot of Twitter users from India were motivated to share their familial histories around Partition. There were dedicated accounts like @1947Pod, @PartitionIB, and @PartitionStoriR created around the declaration to share the stories of migration, violence, and displacement suffered by families. There are long threads where oral histories, newspaper reports, and photos from Partition are shared and a large number of users were engaging with the same.

Almost 37% of the collected tweets appreciated the decision to commemorate Partition Horrors Remembrance Day. Tweeps thanked the prime minister and expressed the feeling that this should have been done immediately after the transfer of power in 1947.

“Thank You Modi India needs to recognize the wounds inflicted first to provide a healing touch to them or they turn into ulcers that bleed our conscience Like this, we also need to recognise the horrors of the Brahmin genocide in Maharashtra in 1948 & the Sikh genocide of 1984” (Twitter user, 14 August 2021)

13.7% of the collected tweets expressed disapproval towards this decision citing that there shouldn't be glorification of violence and an attempt to sector people any further. While 9% of tweets blamed Congress, the Muslim League, and the left for Partition 1.5, 9.8, and 2.6 percentage blamed Azad, Nehru, and Gandhiji respectively. The following tweet against Gandhi amongst many others saying "In this Amritkal, let take a pledge that you will never call someone's Father as your father who is solely responsible for the Partition of your Maa Bharati (India)!" is one of the tweets that has received the most number of likes.

Most of the tweets attempt to shift the blame of the division on individuals or political parties that aren't favorable for them, in doing so they even justify the colonizer reducing their role to just that of an agent who carried forth the plan. One of the tweeps reasons out the event as follows:

The map proposed for the 'Pakistan Commonwealth Nations' by Muslim League under Jinnah at the time of Partition. If Sardar Patel & RSS were not there, Nehru & Gandhi might not have any problem with this map. Idea of Pakistan was incubated in Cambridge Univ, UK by Choudhry Rahmat Ali who envisaged separate Muslim homeland but never settled in Pakistan & spent most of his life in the UK and died insolvent

And many boot licking leftists say it was Savarkar's idea to divide the country

#PartitionHorrorsRemembranceDay (Twitter user, 15 August 2021)

A very few tweets precisely 1.5% express the pain of Partition and hope that India and Pakistan can exist in co-dependence and unity as in the following tweets "I do not necessarily disagree with your statement but, it was India that allowed this separation to happen. We too feel the pain of this Partition to this day, let us be what we were meant to be, Sister Nations!". While 7% of the total collected tweets express the nostalgia

for a unified India wanting to “reclaim Pakistan” (Twitter user, 14 August 2021), 1.3% believes that Partition was a right decision. This solidarity with Partition is expressed in most of the tweets citing the oppression faced by minorities, especially the Muslims in India. One of the tweets is as follows “Every time we see the Indian oppression of minorities, especially the Muslims, only one thought crosses my mind: “#JinnahWasRight”. Similar to the calls for reclaiming Pakistan there are tweets urging to liberate Kashmir as well. While 19.3% of tweets recall the sufferings of Hindus during Partition significantly citing the treatment of minorities in Pakistan, 1.8% of tweets mention the horrors faced by Muslims, and 0.3% about other minorities. A huge 30% of tweets were flagged for hate speech and islamophobia during the computational analysis having content like “Some are still staying in India and singing tunes of Pakistan”, “Some people are still here sir please proper Partition kar do”, and “Jihadi\$, Kafir\$ & liberal\$ who’s Grandparents have missed their last train to Pakistan during Partition is trying to make India look as Pakistan in global map”. A few tweets even expressed their dislike towards the current central Government for the declaration of Partition Horror’s Remembrance Day calling it one amongst the several acts to please the minority. One of the tweets on that topic reads further as “Ask Modi to do his duty instead of appeasing the 2nd largest community, totally illegal residents since Partition, under the garb of Minority & granting them over Rs45000 crores under 300 schemes for them”. The table consolidates the topics, associated hashtags, and the number and percentage of tweets under each topic among the collected tweets.

<b>No</b>	<b>Topic</b>	<b>Associated Hashtags</b>	<b>No. of Tweets</b>	<b>% of Tweets</b>
<b>1</b>	Thanking the Government for the declaration	#PartitionHorrorsRemembranceDay #ThankyouModiji #Modi #NarendraModi #Partition #RemembranceDay #PartitionHorrorsRemembranceDay #YesPunjab #JaiShreeRam #happyindependencedayindia	223	37
<b>2</b>	Against the declaration	#PartitionHorrorsRemembranceDay	82	13.7
<b>3</b>	Blaming Congress, the Muslim League, and left	#Partition #PartitionHorrorsRemembranceDay	54	9
<b>4</b>	Justifying the colonizer	#PartitionHorrorsRemembranceDay	6	1
<b>5</b>	The wound of Partition	#Partition #PartitionHorrorsRemembranceDay	126	21
<b>6</b>	Sharing familial memories of Partition	#Partition #PartitionPodcast #PartitionHorrorsRemembranceDay	48	8
<b>7</b>	Pakistan and India as sister nations	#IndiaPakistanUnity #Leadership #India #Pakistan	9	1.5



8	Partition as the right decision	#ThankyouJinnah #Jinnahwasright #Qaid_e_Azam_Muhammad_Ali_Jinnah	8	1.3
9	Horrors faced by Hindus during Partition	#Partition #PartitionHorrorsRemembranceDay	116	19.3
10	Horrors faced by Muslims during Partition	-	11	1.8
11	The current state of Hindus in India	#FreeHinduTemples #PartitionHorrorsRemembranceDay	33	5.5
12	The current state of Muslims in India	#IndianMuslimsAzadiBaqiAbhi #KashmirBleedsGreen	29	4.8
13	Minorities in Pakistan	#PartitionHorrorsRemembranceDay	36	6
14	Horrors faced by other minorities	-	2	0.3
15	One Nation one law	#OneNationOneLaw	6	1
16	Against Azad	#PartitionHorrorsRemembranceDay	9	1.5

17	Against Nehru	#PartitionHorrorsRemembranc eDay	59	9.8
18	Against Gandhi	#PartitionHorrorsRemembranc eDay	16	2.6
19	Call for unified Bharat	#AkhandBharat	42	7
20	India is heading for another Partition	#PartitionHorrorsRemembranc eDay #Partition	29	4.8

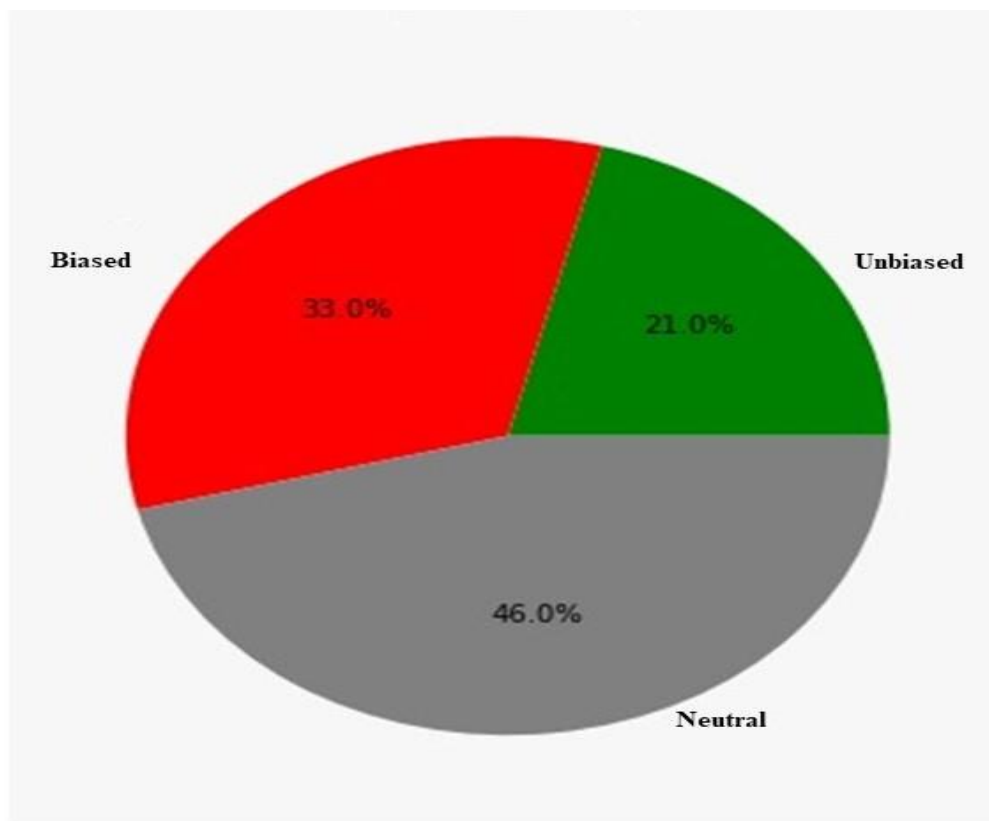
*Table 4.2.3.1: Topic Categories of Analysed Tweets*

#### **4.2.4. Linguistic Bias Analysis of Tweets**

Bias analysis of content posted on Twitter poses a twofold challenge, occurring both at the level of the data source (where users share personal opinions or subjective perceptions of events) and within the algorithms designed for Twitter analytics. Previous studies indicate that these algorithms exhibit noticeable signs of bias in varied categories. This phenomenon can be attributed to the inadvertent incorporation of biases by AI software engineers during the algorithmic design process, thereby perpetuating a problem that AI is intended to mitigate. Given that bias is an inherent aspect of human nature, it inevitably influences the content of tweets, the ranking algorithms for Twitter content, and the algorithms developed for their analysis. Notably, Twitter has publicly acknowledged the presence of bias in its algorithms, acknowledging a tendency to favor specific political ideologies over others (Belli, 2021).

Figure shows the percentage of biased, unbiased, and neutral tweets in the select content, identified through computational analysis. Most tweets that share accurate factual information are flagged as

neutral, while others are classified between the biased and unbiased tags.



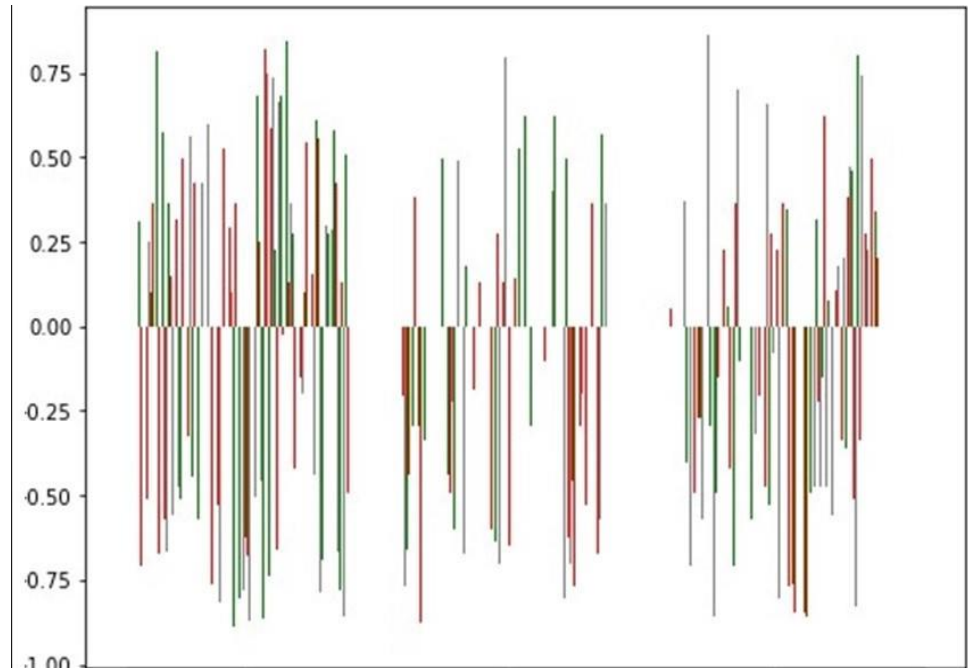
*Figure 4.2.4.1: Bias Analysis of Tweets*

The following tweet from August 14 was posted and cross-posted multiple times with or without minor changes

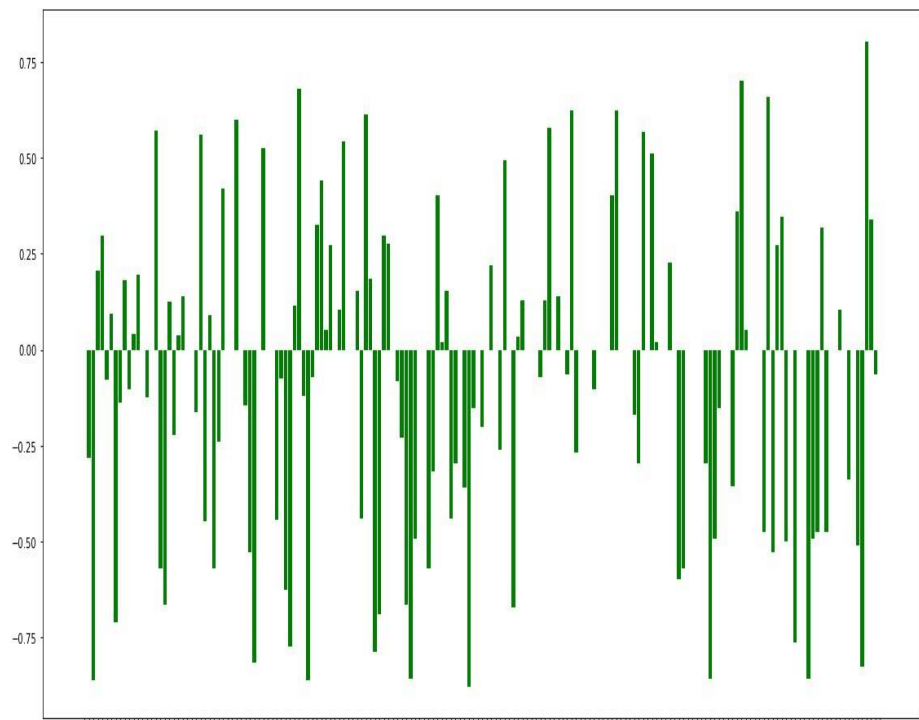
Thank you Modi Ji for recognising the unforgettable pain & struggle with the #PartitionHorrorsRemembranceDay. The nefarious seeds from that day still continue to inflict pain to Bharat Mata. August 14 will always remind us of these seeds—and give us strength (Twitter user, August 14, 2021)

This content was tweeted from 124 handles as comments and primary tweets including verified Twitter handles. In a few instances, words were removed, added, or replaced mostly with synonyms. It is also important to note that while expressing their disagreement towards the declaration of Partition Horrors Remembrance Day, either as a primary tweet or as replies the language used by tweeps was polite and apologetic. For instance, phrases like ‘I my humble suggestion’ and ‘I

request to' were used. Figures show the pattern of biases in sentence and tweet levels of granularity.



*Figure 4.2.4.2.: Sentence level Linguistic Bias Analysis of Tweets*



*Figure 4.2.4.3.: Tweet level Linguistic Bias Analysis*

This analysis further supports earlier research (Spinde et al., 2023), indicating a correlation between the degree of hateful comments or reactions to a tweet and the inherent bias present in the primary tweet. This correlation implies that the level of hostility in comments can serve as a proxy for gauging the bias within a tweet, providing a potential avenue for future research. Political biases are frequently manifested in these tweets through entities and hashtags. Here Hashtags are “both text and metatext, information and tag, pragmatic and metapragmatic speech” (Rambukkana, 2016). Nouns and noun phrases, acting as inductors, play a pivotal role in introducing political bias across a wide spectrum of tweets.

Bias is introduced to tweets through sentiment-charged language mostly using negative emotion words. It has been observed that first-person singular pronouns and words expressing anxiety and feelings are common in case of a biased sentence in a tweet. Similarly, most of the biased tweets highlight a group identity and have an innate artefact of self-reporting. Figure 4.2.4.4 shows a word cloud of the prominent words in biased, unbiased, and neutral categories of tweets.



whereas it allows us to investigate social perceptions of the everyday. Twitter in the context of the 1947 Partition can be perceived as:

1. Mediated sites of contention

Tweets about Partition are sites of reciprocal political, ideological, and religious contentions. Tweets here are sites that are visited, debated, contested, and challenged by the user according to their belief systems. They reflect the hardening religious identities and frequent political conflicts in symbolic, societal, and physical dimensions. Most of the Analysed tweets reinforce images of Partition that appeal to the majority with little or no critical historical appraisal. This not only disregards the past but also shapes a perspective of the present that assumes the role of the predominant and all-encompassing historical view. Once a tweet is posted, it becomes detached from any specific context and lacks a fixed meaning, making it susceptible to reinterpretation and manipulation by various actors in society. Consequently, the tweet persists as a contested site in the public domain, eliciting both shared experiences and differences within our social and political landscape. Despite Twitter offering a direct link to constituents, its horizontal and non-hierarchical nature leads to what is known as "context collapse" — the flattening of diverse audiences into a singular entity. Meyrowitz initially formulated the concept of context collapse to describe the absence of contextual distinctions in electronic media. Most of the tweets on Partition are written with an assumed familiarity about the audience in a vernacular that they will understand according to the perceptions of the tweep. Thus, a strange sense of community is assumed by the tweeps, which in turn makes them comfortable expressing their biases. This perception of togetherness where people are flexibly connected through a myriad of feelings about Partition ranging from anger, grief, resilience, and joy is interconnected with the predisposition towards the proclaimed nationalism.

## 2. Mnemonic medium

Twitter as a memory site has not only memorialized Partition but has also mobilized meaning and developed society's sense of self, which has in turn redirected and shaped the shared public memory. The distortions in memory that are perpetuated through platforms like Twitter contribute to the creation of a unique realm of memory. This realm is characterized by the dynamic interaction between individuals, groups, and historical narratives, leading to the formation of new collective memories. As users engage with diverse perspectives, share personal accounts, and challenge dominant narratives, they actively participate in shaping the collective memory of events such as the 1947 Partition. In this process, individuals expand their ability to adopt novel forms of memory by incorporating previously marginalized or suppressed narratives into the collective consciousness. By acknowledging and amplifying diverse voices and experiences, users on Twitter contribute to a more inclusive and multifaceted understanding of history. Moreover, the interactive nature of Twitter facilitates the formation of new collectives based on shared experiences, interests, or identities. These collectives may transcend geographical boundaries and traditional social structures, bringing together individuals with diverse backgrounds and perspectives. Through collective engagement and collaboration, users on Twitter forge connections and alliances that contribute to the construction of a shared memory landscape. However, it's essential to recognize that the realm of memory on Twitter is not without its challenges and complexities. The platform's algorithms, echo chambers, and viral dynamics can amplify certain narratives while silencing others, leading to further distortions and fragmentation of memory. Additionally, the proliferation of misinformation and disinformation poses risks to the integrity of the historical discourse of Partition on the platform. There is an intermedial adaptation and reconfiguration of the memory of Partition from a diachronic perspective when posted on Twitter. There exist two major memories of the 1947 Partition on Twitter: the "memory of denunciation" and the



“memory of praise.” The first one focuses on dislocation, violence, and loss of land while the second is around the creation of a nation.

### 3. Site of false memory formation

When inaccurate information is conveyed through Twitter there is a higher possibility of formation of false memories when compared to other formal sources. Previous studies suggest that this is not due to deeper or accurate encoding of information but it is due to the informal nature of the tweets that mimicked spontaneous thoughts. (Mickes et.al, 2013). Thus, tweets are expected to have a high rate of false memory formation and integration rate. The unverified inaccurate testimonials about Partition on Twitter thus become sites of memory distortion and false memory formation. The results from the tool show that 8.3% of biases in the collected tweets can be categorized as false memories. It was also observed that while instilling false memory tweeps usually associate it with real-world common-knowledge statements, thus making it believable for the reader.

### 4. Creator and manipulator of collective consciousness

The tweet's capacity to transcend temporal and spatial boundaries mobilizes collective memory, where the tweet itself serves as a locus for validation and dissent among followers. This perpetuates the myths of the past of United India and Partitioned India and Pakistan while articulating explicit forms of ideologies, introducing tensions and conflicts for minorities. Our analysis confirms the formation of an ‘Artificial genealogy of thought’ (Amini & Anwar, 2022) that is the content from earlier tweets creates a confirmation bias and this affects the content of the later tweets in the thread. This way, confirmation bias results in the formation of an artificial genealogy of thought that is far-fetched from truth. Further this self-perpetuating knowledge results in the formation of a rapidly evolving collective consciousness.

#### **4.2.6. Recalling the Past Within a Social Network's Collective Memory: History in the Digital Age**

On Twitter, the process of transferring and subsequently reinterpreting events from the past often entails the remediation of other media forms. Many tweets essentially involve presenting content from another medium, such as a photo, poster, or newspaper article, or sharing links to more comprehensive online media content. This content can range from written articles with images sourced from various journals to YouTube videos or Facebook posts. Consequently, a distinct paradigm of intermediality emerges that of copresence. In this paradigm, diverse media transmitting memory materials can coexist synchronically within a single tweet.

Twitter has altered the way we write and understand history. It is rewriting history in 140 words at a time. It in itself has become an easily searchable 'born digital' archive of millions of personal diaries. A growing trend on Twitter now is 'historical tweeting' where trending past events are tweeted as if they are live. These tweets include old newspaper cuttings, images, and other historical sources and Twitter is thus recreating a long gone past. This trend was followed in the tweets after the declaration of Partition Horrors Remembrance Day where the role of history readers, who are tweeps here has also changed dynamically. Elissa Frankle wrote, "In the history museum of the future, curators' work will be driven by our audiences' curiosity and their preference for inquiry over certainty." Writing about Partition on Twitter is based on this growing preference for co-creating historical knowledge over consuming it. But Several historians and digital humanists have criticised this citing the dangers in sharing historical events and artifacts without proper authentication, and even who has the right to be the sharer of history.

## References

- Alba-Cepero, J., Cabezas-Puerto, M., López-Batista, V.F., Moreno-Montero, Á.M. (2023). Bias Analysis on Twitter. In: de la Iglesia, D.H., de Paz Santana, J.F., López Rivero, A.J. (eds) *New Trends in Disruptive Technologies, Tech Ethics and Artificial Intelligence*. DiTTEt 2022. *Advances in Intelligent Systems and Computing*, vol 1430. Springer, Cham.  
[https://doi.org/10.1007/978-3-031-14859-0\\_12](https://doi.org/10.1007/978-3-031-14859-0_12)
- Ali, M. (2002). History, ideology and curriculum. *Economic and Political Weekly*, 37(44), 4530-4531.
- Anand, K. (2023). Education and National Identity Construction in India and Pakistan. In *Teaching India–Pakistan Relations* (pp. 18–51). <https://doi.org/10.2307/j.ctv2tsxm4.6>
- Apple, M. (2000). *Official Knowledge: Democratic Education in a Conservative Age*. New York: Routledge.
- Aziz, K.K. (1993). *The murder of history in Pakistan*. Lahore: Vanguard Books.
- Belli, L. (2021). Examining algorithmic amplification of political content on Twitter. [Online] Available:  
[https://blog.twitter.com/en\\_us/topics/company/2021/rmlpoliticalcontent](https://blog.twitter.com/en_us/topics/company/2021/rmlpoliticalcontent).
- Bhattacharya, N. (2009). Teaching history in schools: The politics of textbooks in India. *History Workshop Journal*, 67(1), 99–110.  
<https://doi.org/10.1093/hwj/dbn050>
- Binnenkade, A. (2015). Doing Memory: Teaching as a Discursive Node. *Journal of Educational Media, Memory & Society*, 7(2), 29–43. Retrieved from <http://www.jstor.org/stable/44319954>
- Borah, A., & Singh, S.R. (2022). Investigating political polarization in India through the lens of Twitter. *Soc. Netw. Anal. Min.*, 12, 97.  
<https://doi.org/10.1007/s13278-022-00939-z>
- Chen, W., Pacheco, D., Yang, K.C., et al. (2021). Neutral bots probe political bias on social media. *Nature Communications*, 12, 5580.  
<https://doi.org/10.1038/s41467-021-25738-6>.
- Chhabra, M. (2015). Memory practices in history education about the 1947 British India Partition. *Journal of Educational Media, Memory, and Society*, 7(2), 10–28.  
<https://doi.org/10.3167/jemms.2015.070202>
- Christophe, B., & Thorp, R. (2019). Textbook Memories of the Cold War: Introduction to Part One. In B. Christophe & P. Gautschi (Eds.), *The Cold War in the Classroom: International*

- Perspectives on Textbooks and Memory Practices (pp. 41–48).  
Palgrave Studies in Educational Media.
- Coffin, C. (2004). Learning to write history: The role of causality.  
*Written Communication*, 21, 261–289.
- Crismore, A. (1984). The Rhetoric of Textbooks: Metadiscourse.  
*Journal of Curriculum Studies*, 16(3), 279–296.
- Dash, S., Mishra, D., Shekhawat, G., & Pal, J. (2022). Divided we rule:  
Influencer polarization on Twitter during Political Crises in  
India. *Proceedings of the International AAAI Conference on Web  
and Social Media*, 16, 135–146.  
<https://doi.org/10.1609/icwsm.v16i1.19279>
- Frankle, E. (2011, July 28). More Crowdsourced Scholarship: Citizen  
History. Blog post. Center for the Future of Museums.  
<http://futureofmuseums.blogspot.com/2011/07/more-crowdsourced-scholarship-citizen.html>
- Guichard, S. (2012). The Indian nation and selective amnesia:  
representing conflicts and violence in Indian history textbooks.  
*Nations and Nationalism*, 19(1), 68–86. doi:10.1111/j.1469-  
8129.2012.00556.x
- Henley, N. M., Miller, M., & Beazley, J. A. (1995). Syntax, semantics,  
and sexual violence: Agency and the passive voice. *Journal of  
Language and Social Psychology*, 14(1-2), 60–84.
- Hoodbhoy, P. A., & Nayyar, A. H. (1985). Rewriting the history of  
Pakistan. In *Islam, politics and the state: The Pakistan experience*  
(pp. 164-177).
- Horner, M. (1975). *Guidelines for Creating Positive Sexual and Racial  
Images in Educational Materials*. New York: Macmillan  
Publishing Company.
- Ingrao, C. (2009). Weapons of Mass Instruction: Schoolbook and  
Democratisation in Multiethnic Central Europe. *Journal of  
Educational Media, Memory, and Society*, 1(1), 180–189.
- Jacobmeyer, W. (1998). Das Schulgeschichtsbuch – Gedächtnis der  
Gesellschaft oder Autobiographie der Nation? *Geschichte, Politik  
und ihre Didaktik*, 26, 26–35.
- Jalal, A. (1995). Conjuring Pakistan: History as official imagining.  
*International Journal of Middle East Studies*, 27(1), 73-89.
- Joshi, S. (2010). Contesting histories and nationalist geographies: a  
comparison of school textbooks in India and Pakistan. *South  
Asian History and Culture*, 1(3), 357–377.  
doi:10.1080/19472498.2010.485379.

- Kumar, J. (2022). History, National Identity and Politics of History Textbooks in India. *J Pol Sci Pub Aff*, 10:417.
- Kumar, K. (1988). Origins of India's "Textbook Culture". *Comparative Education Review*, 32(4), 452–464.
- Kumar, K. (2001). *Prejudice and Pride*. India: Penguin Books.
- Loewen, J. W. (1995). *Lies My Teacher Told Me: Everything Your American History Textbook Got Wrong*. New York.
- Martin, J. R. (2002). Meaning beyond the clause: SFL perspectives. *Annual Review of Applied Linguistics*, 22, 52–74. Cambridge University Press.
- Meyrowitz, J. (1985). *No Sense of Place: The Impact of Electronic Media on Social Behavior*. Oxford: Oxford University Press.
- Mickes, L., Darby, R. S., Hwe, V., Bajic, D., Warker, J. A., Harris, C. R., et al. (2013). Major memory for microblogs. *Memory & Cognition*, 41(4), 481–489.
- Mittal, D. (2022). Narrativizing Partition and producing stigmatized identities: an analysis of the representation of Muslims in two Indian history textbooks. *India Review*, 21(3), 329-348. DOI: 10.1080/14736489.2022.2086406
- Mozafari, M., Farahbakhsh, R., & Crespi, N. (2020). Hate speech detection and racial bias mitigation in social media based on Bert Model. *PLOS ONE*, 15(8). <https://doi.org/10.1371/journal.pone.0237861>.
- Myskow, Gordon. (2017). Surveying the Historical Landscape: The Evaluative Voice of the Textbook. *Functional Linguistics*, 4(7), 1–15.
- Nathan Rambukkana. (2016). *Hashtag Publics: The Power and Politics of Discursive Networks*. New York: Peter Lang, 30.
- National Council of Educational Research and Training. (2020). *Textbooks from Grade VI to XII*. New Delhi, India: NCERT.
- Nayyar, Abdul Hameed, & Salim, Ahmed. (2003). *The subtle subversion*.
- Pathak, Avijit. (2002). *Social Implications of Schooling. Knowledge, Pedagogy and Consciousness*. Noida, Delhi, and Ahmedabad: Rainbow Publishers.
- Phillips, M. S. (2013). *On Historical Distance*. Yale University Press.
- Polkinghorne, D. E. (2005). Narrative Psychology and Historical Consciousness: Relationships and Perspectives. In J. Straub

- (Ed.), *Narration, Identity and Historical Consciousness* (pp. 3–22). New York, NY: Berghahn.
- Powell, Avril. (1996). Perceptions of the South Asian past: Ideology, nationalism and school history textbooks. In *The Transmission of Knowledge in South Asia* (pp. 190-228).
- Prakash Tandon. (1968). *Punjabi Century, 1857-1947*. Berkeley: University of California Press.
- Prashanth Bhat & Kalyani Chadha. (2023). The Mob, the State and Harassment of Journalists via Twitter in India. *Digital Journalism*, 11(10), 1788-1808.
- Punjab Curriculum & Textbook Board. (2020). *Pakistan Studies textbooks for Grade IX to Grade X*[Textbooks]. Lahore, Pakistan
- Punjab Curriculum & Textbook Board. (2020). *Social Studies textbooks for Grade IV to Grade VIII* [Textbooks]. Lahore, Pakistan
- Report of the Education Commission 1964-66. (n.d.). D. S. Kothari report.
- Ricoeur, P. (1984). *Time and Narrative* (Vol. 1). Chicago, IL: University of Chicago Press.
- Rosser, Yvette Claire. (2004). Contesting Historiographies in South Asia: The Islamization of Pakistani Social Studies Textbooks. In *Religious Fundamentalism in the Contemporary World* (pp. 265-307).
- Rüsen, J. (2004). Historical Consciousness: Narrative Structure, Moral Function, and Ontogenetic Development. In P. Seixas (Ed.), *Theorizing Historical Consciousness* (pp. 63–85). Toronto, Ontario, Canada: University of Toronto Press.
- Said, Edward W. (2000). Invention, Memory, and Place. *Critical Inquiry*, 26(2), 175–192.
- Sam Wineburg. (2001). *Historical Thinking and Other Unnatural Acts: Charting the Future of Teaching the Past*. Philadelphia: Temple University Press.
- Savage, Neil. (2011). Twitter as Medium and Message. *Communications of the ACM*, 54(3), 18–20.
- Smit, R., Heinrich, A., & Broersma, M. (2018). Activating the Past in the Ferguson Protests: Memory Work, Digital Activism and the Politics of Platforms. *New Media & Society*, 20(9), 3119–3139.

- Susan Leigh Star and James Griesemer. (1989). Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3), 387–420.
- Tankard, E., Flowers, C., Li, J., & Rawat, D. B. (2021). Toward Bias Analysis Using Tweets and Natural Language Processing. In 2021 IEEE 18th Annual Consumer Communications & Networking Conference (CCNC) (pp. 1-3). <https://doi.org/10.1109/CCNC49032.2021.9369461>
- Teague, M. (2014, November 30). Long history of racial tension set the stage for Ferguson protests. *Los Angeles Times*. <http://www.latimes.com/nation/la-na-ferguson-tensions-20141130-story.html>
- Timo Spinde, Elisabeth Richter, Martin Wessel, Juhi Kulshrestha, Karsten Donnay. (2023). What do Twitter Comments Tell about News Article Bias? Assessing the Impact of News Article Bias on its Perception on Twitter. *Online Social Networks and Media*, 37–38, 100264.
- Tripathi, D. (2018). Manufacturing Enemy: The Presentation of India in Pakistani Textbooks. In *Cultural and Educational Exchanges Between Rival Societies* (pp. 99–114).
- Turner, E., & Rommetveit, R. (1968). Experimental Manipulation of the Production of Active and Passive Voice in Children. *Journal of Language and Speech*, 10(3), 169-180.
- van Dijck, J. (2007). *Mediated Memories in the Digital Age*. Stanford University Press.
- Villa-Nicholas, M. (2019). Latinx Digital Memory: Identity Making in Real Time. *Social Media+ Society*, 5(4).
- Waller, J. (2002). *Becoming Evil: How Ordinary People Commit Genocide and Mass Killing*. Oxford University Press.
- Williams, Raymond. (1961). *The Long Revolution*. London: Chatto and Windus.
- Xiao, Z., Zhu, J., Wang, Y. et al. (2023). Detecting Political Biases of Named Entities and Hashtags on Twitter. *EPJ Data Sci*, 12, 20.

## Chapter 5

### AI Biases: A Digital Humanities Response

*“Maybe the only significant difference between a really smart simulation and a human being was the noise they made when you punched them.”*

— Terry Pratchett, *‘The Long Earth’*

“Digital Humanities is the methodological community of Data-driven research of humanities, which is based on new technologies” (Liu & Ye, 2017). The "Methodological Commons" of digital humanities encompasses a comprehensive set of practices aimed at discovering, accumulating, comparing, disseminating, and collaborating, all facilitated by a diverse array of technologies. These technologies include digitization, data management, visualization, and machine learning, among others (McCarty & Short, 2017). In essence, the field of digital humanities is built upon the foundational principles of uncovering knowledge, accumulating resources, comparing information, providing access, and fostering collaboration among scholars and practitioners. Moreover, the establishment and development of digital humanities can be delineated across three key dimensions. Firstly, there is the dimension of arrangement and depiction, which involves structuring and visually representing data and information in meaningful ways. The second dimension is of restoration and dissemination, which pertains to the efforts to reconstruct historical or cultural artifacts digitally and make them widely accessible to diverse audiences. Finally, there is the dimension of computer-facilitated advancement and expansion, which encompasses the utilization of computational tools and methodologies to further enhance and extend the scope of research and scholarship within the field (Luo, 2015). Through these dimensions, digital humanities endeavors to advance knowledge creation, preservation, and dissemination in the digital age, fostering interdisciplinary collaboration and innovation along the way.



The fundamental distinction between "humanities computing" and "digital humanities" is rooted in the evolution from a paradigm where humans were solely responsible for "hypothesis testing and material verification" to a collaborative approach involving both humans and machines. In this contemporary context, the process encompasses not only the traditional activities of formulating hypotheses, observing, uncovering, analyzing, and testing but also emphasizes iterative cycles of hypothesis refinement, facilitated by the integration of computational tools and methodologies (Dai et al., 2016).

In recent years, digital humanities have witnessed significant growth driven by two key trends.

- The shift from small-scale manual databases to big automated data projects
- Lowering the technical threshold to use computers for textual analysis

These trends foster the integration of Digital Humanities and Artificial Intelligence by lowering the barrier to entry, as they now demand only intermediate computational skills from humanities researchers to effectively leverage automated tools in their research endeavors (Dai et al., 2016).

Artificial Intelligence (AI) has emerged as a transformative force in the field of digital humanities, revolutionizing the ways in which scholars explore, interpret, and understand humanities data. By leveraging machine learning algorithms, natural language processing techniques, and other AI-driven tools, researchers are able to analyze vast amounts of textual, visual, and auditory data with unprecedented speed and accuracy. AI enables humanities scholars to uncover hidden patterns, extract meaningful insights, and generate new perspectives on texts, artworks, and cultural artifacts. Moreover, AI facilitates the creation of innovative digital archives, virtual exhibitions, and interactive learning platforms that enhance public engagement with cultural heritage and historical narratives. As AI continues to advance,

its integration into digital humanities promises to further enrich scholarly inquiry, and foster interdisciplinary collaborations. With its looming importance in Digital Humanities AI has become both a research methodology and an area of research for a humanist with the underlying questions of ethics and responsibility. While the involvement of AI in complex decision-making tasks in humanities research was to reduce human subjectivity, it can rather cause fallible and discriminatory predictions due to embedded biases. Thus, it is important to address these biases not just to harness the transformative potential of AI, but also to decrease the antithetical repercussions it can cause on the underrepresented communities.

In recent years, the discourse surrounding AI biases and ethical considerations within computer science has expanded to encompass broader solutions and perspectives from the humanities and social sciences. Scholars in humanities cultural analytics have contributed significantly to this dialogue, examining the algorithmic and software aspects of AI within cultural contexts. Works by authors such as Berry (2021), Bratton (2016), Chun (2008, 2017), Dix et al. (2003), Fuller (2003), Galloway (2006), Hayles (2012), Manovich (2001, 2008), Montfort (2016), Marino (2006, 2020), and Wardrip-Fruin (2009) delve into the intersection of AI, culture, and society. Additionally, there's a growing body of research focusing on the intersection of new media, machine learning, and ethics, with studies by Amoores (2020), Benjamin (2019), Chun (2011), Floridi (2013), Floridi and Taddeo (2016), Gray (2014), Han (2017), and Noble (2018) shedding light on ethical considerations and biases inherent in AI systems. Moreover, humanities scholars have directed attention towards issues of data governance, privacy breaches, and surveillance, as evidenced by works by Browne (2015), D'Ignazio and Klein (2019), and Zuboff (2019). Works like Safiya Umoja Noble's "Algorithms of Oppression" provides a critical examination of how search engine algorithms can perpetuate and reinforce racial biases, highlighting the ways in which digital technologies intersect with systems of power and inequality. Ruha

Benjamin's "Race After Technology" further expands on this discourse, exploring the concept of the "new Jim code" and proposing abolitionist tools for dismantling racialized technological systems. Cathy O'Neil's "Weapons of Math Destruction" delves into the impact of opaque algorithms on society, illustrating how these "weapons" can amplify disparities and perpetuate injustices, particularly for marginalized communities. Similarly, scholars like Virginia Eubanks, in "Automating Inequality," and Meredith Broussard, in "Artificial Unintelligence," offer critical perspectives on the ways in which AI technologies often fail to address the needs of marginalized populations, exacerbating socio-economic disparities instead of alleviating them. Eubanks investigates how high-tech tools can profile, police, and punish the poor, exposing the hidden consequences of automated decision-making systems in welfare, housing, and criminal justice contexts. Broussard, on the other hand, examines the limitations and biases of AI technologies, arguing that computers frequently misinterpret the complexities of the world, leading to unintended consequences and reinforcing existing power structures.

Adding to this emerging scholarship this chapter aims to understand

- How does digital humanities currently employ Artificial Intelligence?
- What obstacles do digital humanists encounter when working with or incorporating AI into their research?
- What are the ethical guidelines to be followed by a researcher who embarks on AI-integrated DH research?

Despite the recent focus on ethical concerns of AI, there is no tangible evidence-based study that practically assesses algorithms. This chapter attempts to address this gap by 'use cases' of AI algorithmic modeling, precisely the AI model for offensive speech detection for Twitter, to articulate with evidence how bias enters algorithms and how it can cause discrimination and fallible outcomes. The 'complete debias' goal of AI systems is unmitigable as it isn't only technical but it has innate societal,

cultural, and ethical implications. The objective here is to acknowledge why and how bias is embedded in AI systems, its long-term impacts, and effective ways to address it.

This chapter is divided into 8 sections where section 1 gives a brief introduction to Digital Humanities, Artificial Intelligence, and Big Data. Section 2 delineates ways in which digital humanities scholars employ AI and section 3 addresses the transformation in the process of research and the roles of the researcher with the integration of AI in DH research. While section 4 briefs the ethical and social implications of AI, sections 5 and 6 study AI biases in detail with a focus on offensive speech detection in Tweets. Section 7 deals with the challenges of a digital humanities scholar in employing and studying AI and in response to that section 8 proposes a set of guidelines for a DH researcher working on/with AI for humanities research.

### **5.1. Digital Humanities, Artificial Intelligence, and Big Data**

There are three distinguishable phases in the evolution of automated practices in Digital Humanities

1. First phase, where target-oriented software was created to address specific humanities research questions. In this phase, the software was mostly programmed by a self-trained humanist coder. This phase focused on tools that could help sort and classify large amounts of data for humanities research.
2. Second phase, where ready-to-use packages like Voyant tools, ATLAS.ti, and Text encoding software were made available. These tools addressed several research problems all in one place and were created with a beginner-friendly user interface. This phase also saw a rising number of collaborations between humanists and computer scientists. Tools during this phase fall under two categories based on their objectives
  - a. Digital tools for textual data
  - b. Digital archives which were mostly image and transcription-based

3. Third and the current phase is where humanist coders attempt to customize the existing machine-learning packages to suit their needs.

While the initial integration of AI into digital humanities research commenced during the second phase, its widespread adoption has become increasingly conspicuous in the current phase. This broad uptake can be ascribed to various factors. Primarily, advancements in AI technology have rendered it more accessible and user-friendly, thus diminishing the barriers to entry for researchers lacking extensive technical expertise. Additionally, the proliferation of large-scale datasets and computational resources has empowered digital humanists to employ AI algorithms across a diverse spectrum of research endeavors, spanning text analysis, data mining, image recognition, and natural language processing. Moreover, collaborative efforts among computer scientists, data scientists, and humanities scholars have yielded innovative methodologies that harness the capabilities of AI to tackle intricate research inquiries within the humanities domain.

## **5.2. How does the Digital Humanities use AI?**

Artificial intelligence displays great potential in addressing large-scale data-driven questions in digital humanities. It has become an effective tool aiding in a sophisticated understanding of even complex concepts like social behaviors, human cognition, cultural dynamics, and human relationships. “AI ushers in novel views for analysis and poses significant problems and opportunities for the humanities and social sciences” (Gefen, 2021). AI is a lot more than a “research program”, “It builds particularly original bridges between biology and mathematics, rethinking the unity of scientific knowledge through a new tree of sciences and imposing a strong epistemological reflection to historians and philosophers” (Feyerabend and Lakatos, 2014). Meanwhile, humanists are raising questions about the foundations, history, and social, cultural, and ethical implications of AI. They also propose effective work benches and strategies to overcome the humanistic

impediments faced by AI researchers. Thus, it is a research tool and area of research for digital humanists.

AI was first used by digital humanists for image recognition tasks and later it was extended to analyze other forms of data including text, speech, and historical artifacts. It aids in data analysis, data classification, decision-making support, machine translation, financial predictions, and various other tasks. It is used by present-day digital humanists as a research tool and methodology mostly in semi-automated research. They either use the open-accessed datasets and tools or modify and adapt from the existing models according to their research objectives. The prominent AI-assisted research tasks in humanities are as follows:

1. Natural Language Processing and Text Analytics

Text analytics is an interdisciplinary field that works at the intersections of data mining, computational linguistics, machine learning, information extraction, and statistics. Also known as “Text Data Mining or Knowledge Discovery in Text (KDT)”, Text analytics is defined as an extension of data mining that identifies patterns from huge, complex, and unstructured textual sources, while data mining tools handle structured data. Digital Humanists employ text analytics to study unstructured and/or semi-structured data sets including, literary texts (novels, short stories and, poetry), social media writings, interpersonal communications, newspaper reports, and so on. A basic text analysis process consists of

1. Collection and retrieval of document
2. Preprocessing
3. Text analysis

(A detailed description of the process can be found in Chapter 2). This process generates structured or sometimes semi-structured data that can be used for further tasks. Digital humanists

extensively use opinion classification and classification approaches while conducting a text mining task.

AI in text mining tasks addresses the challenges in classification, annotation, and information retrieval. It instigates new quantitative and qualitative methodologies by facilitating complex data analysis tasks such as semantic analysis, named entity recognition, topic modeling, data and document classification using machine learning, sentiment analysis, network analysis, etc. It reveals insights that wouldn't have been visible for a human researcher through distant and or close reading.

Employing AI in text analytics poses significant challenges. Biases inherent in AI models can perpetuate existing biases in textual data. The interpretability of AI models is often limited, making it challenging for researchers to understand decision-making processes. Scalability issues arise when dealing with large volumes of unstructured data, and ensuring accuracy across diverse datasets remains a challenge. Employing AI in text analytics poses significant challenges. Biases inherent in AI models can perpetuate existing biases in textual data. The interpretability of AI models is often limited, making it challenging for researchers to understand decision-making processes. Scalability issues arise when dealing with large volumes of unstructured data, and ensuring accuracy across diverse datasets remains a challenge. The challenges pertinent to text mining in the case of literary texts are pertinent to Natural Language Processing (NLP). The ambiguities of language because of regional dialects, jargon, acronyms, and abbreviations raise a significant challenge in the analysis of varied content like social media narratives, newspaper reports, and blogs. Underlying meanings, trolls, incoherent sentences, and poor grammar can affect the interpretation of texts.

## 2. Archiving

The vast amount of data management required for producing digital archives has coerced digital humanities researchers into using Artificial intelligence, understanding its potential to make archives user-friendly and accessible. AI aids in the

1. recognition of sensitive information,
2. keyword searches and information retrieval
3. organization of metadata,
4. Automatic content extraction using Optical Character Recognition software (OCR) and its indexation
5. discoverability of Archives.

Using AI for digital archiving poses a significant set of challenges associated with sensitive data classification due to Insufficient relevant context variables or the use of wrong variables. These challenges further raise the question of privacy and preserving of categorised information There are also significant challenges in terms of duplication of information and infiltration by encoding biases.

### 3. Image processing

AI is employed for the analysis of images, especially by researchers in Art History. AI is employed in comparing artwork, which is theoretically called ‘iconology’ by Aby Warburg. With the assistance of scanning systems, a searchable compilation of visual data can be created and one example of such a corpus is the Google Art Project. This big visual data can be compared, analyzed, and grouped according to the formation of patterns and diffusion of styles and also considering the metadata using AI image processing systems. Similar to other AI-assisted analyses, image processing also poses challenges in terms of biases, data privacy, and security.

Though Digital humanists were keen on harnessing the benefits of AI, a few scholars found that it is necessary to defend the humanities from AI. Lauren M. E. Goodlad & Samuel Baker addressed AI as a



‘moth without a brain’. There were arguments that humanities, especially Arts became undervalued with the advent of AI. Dana Kim in her essay ‘Ai Poses a Threat to the Humanities’ creates a narrative suggesting that the humanities are on the verge of extinction with AI and she concludes by saying ‘there is still a way to explore our humanity before it’s too late’. The shift to AI in humanities according to Bachimont, operates a triple rupture: “a rupture of data from its source and nature ... a rupture of the processing from the data ... and finally, a rupture of what is shown from what is calculated” (Bachimont, 2015). But as Teresa Heffernan rightly said “It is only by keeping alive the tensions between artificial intelligence and the humanities that we can hope to have an informed debate about the limits and possibilities of this technology”

### **5.3. Changing roles of DH research and researcher**

Artificial intelligence has not only transformed curation and analysis of data but it has transformed research approaches in humanities. It reveals details that are otherwise not attended by a researcher. The fundamental roles and frameworks are altered and it provides scholars with a renewed vantage point to data-driven research.

“AI is changing the way Social and Human Sciences model and analyze their data, understand their objects, and visualize their results... First emerged in the field of image recognition, applications of machine learning have soon extended to other types of data and tasks including printed characters, handwriting, and speech recognition, but also financial prediction in the financial, decision support in law or medicine as well as machine translation” (Gefen et.al, 2021)

AI has challenged the existing notions of methods and disciplinary boundaries in DH. AI in DH is about exploring new ideas and generating evidence to justify the interpretations of those ideas to a certain extent. It has helped DH in achieving its previously defined values of

collaboration and experimentation. Exposure to big data allows a DH researcher to

approach interesting questions from multiple and interdisciplinary points of view, in the way that citations to textual sources do not. Again, we are arguing not for wholly replacing close readings and textual analysis in ... research but, rather, for complementing them with our explorations of data” (Kitchin, 2014).

In humanities where data is mostly noisy and fragmented, a combination of qualitative and quantitative methodology for research is significant in a holistic understanding with methodological rigor and data determinism.

A DH researcher with the advent of AI is a digital archivist and curator of their work. Here the humanist has to deeply engage with the AI systems objects to create, substantiate, and archive their research. But scholars here are not just users of the AI systems, the critical digital humanist speculatively perceives it with a critical humanistic perspective, thereby questioning technological determinism. “Politically committed academics with humanities skill sets must engage technology and its production not simply as an object of our scorn, critique, or fascination but as a productive and generative space that is always emergent and never fully determined” (Mcpherson, 2012).

The researcher here becomes a ‘creative technologist’ who combines professional software expertise with a humanistic understanding of research. While advancements in data organization and initial processing have proliferated, the intrinsic complexity of humanities data necessitates that researchers allocate substantial time and effort to the coding phase of research rather than focusing primarily on analysis stages. With the traditional monographic pattern of scholarly communication, there is a lack of data and methodological transparency in AI-enabled humanities research challenging the reproducibility of the results. Scholarly writing in DH should critically engage more with the hermeneutics of data queries, tool workflows and failures, and

generation and analysis of data visualizations, and thus similar to Gibbs and Owens's contention about historical writing “our methodologies might not be as deliberate or as linear as they have been in the past. This means we need more explicit and careful (if not playful) ways of writing about them” (Gibbs & Owens, 2018).

#### **5.4.Challenges of a Digital Humanist working with AI**

Along with the questions of biases and fallibilities, the major challenge for a digital humanities scholar employing AI systems is that of access, transparency, interpretability, and explanatory deficit. In most cases, scholars may not have access to the algorithm and thus will have a blind eye toward how AI analyses and classifies the underlying research data. In other cases, even when they have access to a humanities scholar, lacking the background and infrastructure will be unaware of the complex ethical decisions entangled with the code. To take the case of Voyant tools, an online text analysis environment that performs tasks including word frequency lists, frequency distribution plots, and KWIC displays, though the project is open-sourced and the codes are available on GitHub it is comprehensible by only a few digital humanists. Thus, the ways in which the code operates on the data and its strategies of classification remain unknown to researchers. Thus, the humanities researcher assumes the role of a user of AI tools than that of a mediator in their research. This way the researcher is also unaware of the biases embedded in the code. “Computational systems tend towards the generation of quantifiable hierarchies of entities, particularly when classifying, sorting, and filtering, and even the notion of such a hierarchy is, of course, an imposed framework of understanding” (Johnson 2018). “Similarly, because of the underlying programming logics of a particular software tool, certain design assumptions, such as objects, networks, collections, or even the presumed suitability for particular properties of entities to be valued over others, may have effects on the way in which those digital objects are subsequently manipulated by the software” (Berry 2011a; Emerson 2014; Marino 2020; Chun 2021).

The second set of challenges arises from the legibility of code especially in a machine-learning system. “due to the way in which machine-learning systems work, they tend to encode their classifying processes within a table of numbers which represent relationships between entities” (Berry, 2023). In a big data system, these tables will be huge increasingly so according to the parameters of classification and will be beyond the reasoning of a humanities researcher. “One can see how these automating processes might cover ethical issues by transferring them into the hashtables of the machine-learning system. Machine learning also has the power to automate larger parts of the humanities, replacing research work that was previously thought could only be the preserve of the human researcher such as identifying key concepts” (Berry, 2023).

Thus interpretability and explainability become a key concept for employing AI systems in humanities and also for a humanities researcher to probe the ethical issues of biases of AI. Explainability and interpretability are proposed to generate sufficient context and reasoning for the classificatory structure and decision-making of an AI system.

While interpretable AI provides clear and concise insights about the decision-making process that can be easily understood by human experts who aren't experts in complex mathematical algorithms. Whereas explanatory AI reasons and interacts with the user to provide comprehensive explanations about the decision-making process by providing sufficient information about the context, features, and underlying reasoning of classification “thereby giving us the power to change our environment in order to meet our own ends and thereby raise interesting ethical questions” (Berry 2021).

Google was quick to work on the criticisms of Safiya Umoja Noble in her work ‘Algorithms of Oppression: How Search Engines Reinforce Racism’ (Noble, 2018). She writes against data discrimination and how racism and stereotypes are reinforced through the digital. Though there are changes it is quite unclear how these clarifications

were made. It was previously noted that in most cases the responses to these issues are faulty and result in further biases. As Hern identifies Google's response to the outrage against how images of black men were labeled as “gorillas” in Google Photos was to censor the tags, so that no images were ever labeled gorilla, chimpanzee, or monkey, even if they were pictures of the primates themselves. Similarly in another instance on Flickr when a picture of a black man was labeled as an “ape”, the term was removed from the tagging lexicon of Flickr (Hern 2018). The changes here are dubious but concealed. This raises the issue of transparency – where a researcher communicates the presence of a bias or an ethical issue but isn’t informed about the adopted solutions to mitigate the bias.

### **5.5.Ethical and Social Implications of AI**

Despite all its advances, AI has profound implications for democratic values, privacy, and order in civil society, raising many ethical, cultural, legal, and social challenges. These challenges are primarily rooted in humanities and social sciences and are of immense research significance for digital humanities. Figure 5.5.1. shows a word cloud of AI now reports from 2017 -2023 to show the key discussions in the area for the past years. This research in the further sections focuses mainly on AI biases. The illustration indicates a lack of emphasis on the issues surrounding algorithmic bias, the transparency and explainability of "black box" systems, and concerns regarding data privacy.



*Figure 5.5.1: Word Cloud of AI Now Reports*

These omissions underscore the pressing need for heightened attention and comprehensive examination of these vital aspects in the ongoing discourse surrounding AI ethics and governance. Addressing these issues is essential for fostering responsible AI development and deployment and ensuring that AI technologies align with democratic principles, uphold individual privacy rights, and promote societal well-being.

## 5.6. AI Bias

Biases in AI systems are primarily from human creators mirroring the dominant culture. Perspectives of the developer and values of the created organization are often reflected on an AI systems biased integration. Based on the Computing Research Association (CRA) survey, 2022, and AI index report the percentage of women in AI computing is 26% which is lower than women in computing during the 1960s. 18.3% of all Ph.D. graduates in AI Ph.D. programs are female and 16.1% of tenure track faculty from the area are women. According to the Artificial Intelligence Index Report 2021 across the majority of countries global relative AI skills penetration rate for women is lower than that of men. Considering the demography of Ph.D. in AI the largest

percentage (45.6%) are non-Hispanic white, 22.4% are Asian, 2.4% are Black or African American, 3.2% are Hispanic and 1.6% are multi-racial. Similarly, in the case of tenure track faculty in the field in top universities, 67% are white while 14.3% are Asian and 0.6, 0.8, and 2.7 are Black, Hispanic, Middle Eastern, or North African respectively. In the case of tech firms, among Microsoft employees only 4.5% are black and 6.3% are Hispanic which is not different from other companies. Based on the data from Queer in AI (QAI) survey half of the respondents observe a lack of inclusiveness in the field and 40% have experienced discrimination or harassment. This lack of gender, sexual, demographic, and racial diversity is one major reason for the bias.

Bias can enter an AI system during the data generation and data management stages and also during the composition and evaluation of the algorithm. In the creation of a dataset, bias enters because of multiple reasons such as

1. Non-existence of data points
2. Lack of disaggregation of data points
3. Existing data points are biased
4. Biased data generated by users
5. Historical (ex: race, gender, caste, ethnicity, geography, etc) and current inequities (ex: digital divide, etc) that introduce the bias or limit the data collection
6. Biased selection of data points into a dataset
7. Subjective and discriminatory data labels
8. Constraints of AI model around binaries

This results in a dataset that is unrepresentative or under-representative of the society or it will create a dataset that is representative of the existing unjust society. Similarly, algorithms become biased as

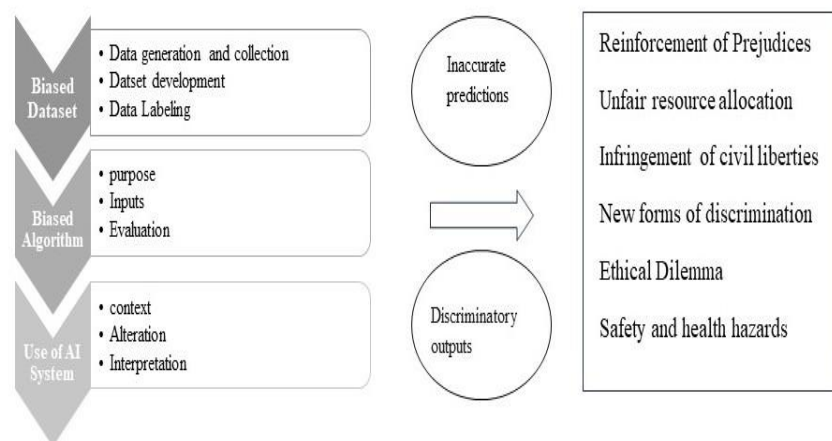
1. The objectives of composition are narrow and technical lacking equity considerations
2. Usage of biased dataset
3. Usage of a dataset that is not created for the target population

4. Variables or proxies that inadvertently (or not) discriminating communities
5. Lack of consideration for social group variables
6. Algorithms are either not audited or are audited with limited metrics.

These results in algorithms that are inaccurate and discriminatory especially for certain subsets of the population. This issue is further solidified by the lack of accountability and transparency in the algorithm. Often this happens quite unintentionally, but the resultant inaccurate and discriminatory outputs can have adverse results. AI systems can result in biased output also at the user stage when

1. They are used in the context of different populations
2. They are used in a way they are not initially operationalized for
3. The algorithm is discriminatorily altered
4. Outcomes are interpreted with bias

Figure 5.6.1. shows how bias enters an AI system and the hazards paved by biased algorithms and datasets.



*Figure 5.6.1.: Bias in AI Systems*

The pervasive biases inherent in AI systems are intricately linked with the demographic composition and perspectives of their human creators. The notable underrepresentation of women, racial and ethnic



minorities, and LGBTQA+ individuals within the AI domain significantly contributes to the perpetuation of bias. This dearth of diversity not only influences the formulation and assessment of AI algorithms but also extends to the stages of data generation and management, where biases may be inadvertently introduced or magnified. From the creation of biased datasets to the delineation of narrow algorithmic objectives and the inadequacies of auditing mechanisms, various factors converge to foster the proliferation of inaccurate and discriminatory AI systems. Furthermore, the lack of accountability and transparency in these systems exacerbates these issues. Addressing these challenges necessitates concerted efforts to promote diversity and inclusivity within the AI field, enhance accountability and transparency in algorithmic processes, and prioritize equity considerations throughout all phases of AI development and implementation.

### **5.7. Offensive Speech Detection in Tweets**

This section analyses biases in AI-integrated offensive speech detection in Tweets to understand how bias is embedded in AI language tools as a ‘user case’. This analysis looks into the intricate nature of algorithmic biases and how they result in biases towards certain groups. This analysis is not an attempt to analyze the performance of speech metrics or its efficiency or inefficiency in monitoring social media content rather the study delineates how this algorithmic classification can lead to a bias.

Twitter’s developer platform has constituted ‘hate lab’ as part of Cardiff University’s ‘Social Data Science Lab’, which is a partnership between the university’s School of Social Sciences and School of Computer Science and Informatics to “understand the dynamics of the toxic conversations on social media - the enablers, the drivers, the inhibitors - to help inform how to foster healthy conversations online” (Hate Lab, 2023). Twitter says that at ‘Hate Lab’ they use “data science methods, including ethical forms of AI, to measure and counter the

problem of hate both online and offline” (Hate Lab, 2023). In 2022 over 5 million tweets were removed by Twitter, out of which 1.3 million were flagged as hate and another 1.3 were flagged for abuse and harassment. But relying completely on AI to detect and combat offensive speech wouldn’t be effective, rather it may even work contrarily as

1. AI cannot take context into account
2. There is a lack of adequate datasets to develop algorithms
3. AI isn’t cognizant of the differences in speech patterns and changing patterns while tagging a sentence as offensive
4. There are no metrics for the assessment of offensiveness as it is mostly contextual

Thus, this may further increase the opacity of content moderation and cue in unaccustomed disputes in terms of fairness and justice towards certain sects of people.

Previous studies on bias in offensive speech corpuses primarily concentrated on detecting and addressing bias against explicit identity references (Park and Fung, 2017; Dixon et al., 2018). Sap et al. investigated how "insensitivity to dialects can result in discrimination against minorities". Meanwhile, Wiegand, Ruppenhofer, and Kleinbauer (2019) examined biased datasets containing implicit abusive content, and Nozza, Volpetti, and Fersini (2019) analyzed bias in identifying misogyny.

This research studies the results from TILT an online Hate Speech Detection Solution for Twitter, that uses a machine learning algorithm to filter hate, toxicity, and threat. The platform helps to

1. Set up a customized investigation of offensive content posted
2. Filter and rank hateful content
3. Visualise network patterns and relationships

This section reused the Twitter data used for analysis in Chapter 3, for performing the analysis. The results from the analysis using TILT are then compared with the manual hate offensive speech tagging by 10

researchers. The research team replicated the algorithmic process of categorization by deeming a tweet as offensive if the vast majority of its content contains hate. The model is tested by accounting for both the ‘false positive rate’ (FPR) and the ‘false negative rate’ (FNR), where FNR is the percentage of tweets that are classified as offensive by the researchers, whereas not offensive by TILT. Similarly, FPR is the percentage of tweets that are classified as offensive by the researchers but not TILT. Both FNRs and FPRs are significant in social media platforms. FPRs can subject the tweet through objectionable censorship while FNRs can fail the objective of curbing hate and vile targeting on Twitter.

### *Key Findings*

- Previous studies (European Union Agency for Fundamental Rights, 2022 & Nascimento et.al, 2022) conclude the word ‘Muslim’ and its associated terms can lead to the prediction of an offensive speech even when there aren’t any. This was termed as an ‘overreaction’ of the algorithm by previous researchers. But on the contrary, this study has identified that 54% of the Islamophobic tweets weren’t flagged for hate partly because the words like ‘ulcers that bleed’, ‘They’, ‘League mobs’, ‘2nd largest community’, and ‘illegal residents’ were used. Tweeps also use special characters in place of words as in M@sl!ms to escape the algorithm.
- TILT shows a very low error rate on non-offensive comments
- The highest rate of false positives is reported when identity terms are used
- AI algorithms easily react to single words and negative words such as ‘hate’, ‘violence’, ‘internal war’ ‘fight’, and so on and predict that the sentence is offensive
- Using positive words for ex: ‘love’, ‘unity’, and ‘peace’ reduce the chances of content being categorized as offensive
- The algorithm is not designed to identify the dialectical changes, especially in short strings of texts like tweets, thus Indian

English variations and sentences with words from regional languages were not rated precisely.

It is important to note that these results are dependent on the applied methodology and thus are not necessarily generalizable for other algorithmic models for offensive speech detection or a similar algorithm in a different context. However, there are suggestive patterns of biases that can reappear in other AI models. Thus, these results should caution developers and policymakers to inspect their models for biases that can contribute to discriminatory outputs. Despite efforts such as Twitter's 'Hate Lab' initiative, overreliance on AI for content moderation presents significant challenges, including the inability to contextualize content and the lack of standardized metrics for assessing offensiveness. The findings underscore the importance of scrutinizing AI models for biases that may contribute to discriminatory outputs, emphasizing the need for greater accountability and transparency in content moderation processes.

### **5.8.Ethical Framework for Digital Humanists Engaging with AI**

The solutions identified by previous researchers to the ethical problems with AI are mostly based on deontology, consequentialism, or virtue ethics (van den Hoven, 2010). Berry in his chapter 'AI, Ethics and Digital Humanities' suggests Van den Hoven's approach to ethics to address the ethical dilemma of AI. Deontological principles advise "compliance with the highest ethical principle: the Categorical Imperative," while making choices, irrespective of the results. In contrast to this consequentialists approach ethics in terms of consequences of action and thus choices should be made based on the "course of action which maximizes utility compared to the alternatives" (van den Hoven, 2010). Virtue ethics is based on the idea of identifying "the morally relevant features in every situation and determining the right course of action" (van den Hoven, 2010)

There is a lack of ethical guidelines when one embarks on a digital humanities project that employs AI. Instead, most ways of operation around ethical questions in “digital humanities tend towards a form of virtue ethics in terms of the kinds of character that digital humanists are thought to have” (Berry, 2023) such as being “‘nice’: collegial, egalitarian, non-hierarchical, and able to work well with others” (Koh 2014, 95). “Among core humanistic values are inquiry, critical thinking, debate, pluralism, balancing innovation and tradition, and exploration and critique” (Spiro 2012). In addition, digital humanities fully embrace the Internet values reflected in the “DH community’s focus on openness, iterative development, and transparency, as well as in its adoption of open-source approaches to code development and community education” (Scheinfeldt quoted in Spiro 2012).

This study proposes a set of guidelines a digital humanities researcher can adopt while creating or working with an AI system for research. Along with the propositions of the ‘Berne DHSS Declaration on Research Ethics in Digital Humanities’ a Humanities researcher who works with and/or on AI should work with the following social and ethical responsibilities.

1. **Critical lens:** It is important for a DH researcher to be not pre-empted by either techno-utopianism or technopositivism while approaching AI. The researcher should have a definite and stark critical lens toward the creation, processing, and results of AI.
2. **Data:** It is important to understand that AI is not just the algorithm but the underlying data plays a significant role. While curating huge datasets the nature of humanities data should be taken into account. Below are the key cautions while creating a humanities data set:
  - A. datasets should be fairness-unaware
  - B. researcher should refrain from repurposing datasets

- C. curator should draft a purpose statement focusing on motivation rather than cause before the process
- D. informed consent is crucial, ensuring privacy and security of the data as well as the participant
- E. there should be a well-laid policy on who owns the data

F. When seeking consent for data collection, it is imperative to consider marginalized and vulnerable populations who may lack the capacity or authority to provide informed consent. Special attention should be given to underrepresented communities who may face systemic barriers to participation. Furthermore, it is crucial to meticulously document the source of the data, ensuring transparency and accountability in the research process.

G. Ensuring the representation of underrepresented groups and actively mitigating encoding biases that perpetuate historical prejudices are critical aspects of ethical data practices.

H. the digital divide and accessibility should be a key concern.

3. **Transparency:** A theoretical understanding of the process of the system should be a focus.
4. **Explainability and interpretability:** generate sufficient context and reasoning for the classificatory structure and decision-making of an AI system

5. **Adaptability and Access:** The model should be adaptable and should be accessible for further research whenever possible following the base principles of Digital Humanities
6. **Linguistic diversity** of DH should be realized.
7. **Labour-** digital or manual should be acknowledged
8. **Societal, and ethical consequences** of the model should be the focal point.

## References

- Amoore, L. (2020). *Cloud Ethics: Algorithms and the Attributes of Ourselves and Others*. Duke University Press.
- Bachimont, B. (2015). Le numérique comme milieu: enjeux épistémologiques et phénoménologiques. *Interfaces numériques*, 4(3/2015), 385–402.
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity Press.
- Berry, D. M. (2011). The Computational Turn: Thinking about the Digital Humanities. *Culture Machine*, 12.
- Berry, D. M. (2021). Explanatory Publics: Explainability and Democratic Thought. In B. Balaskas & C. Rito (Eds.), *Fabricating Publics: The Dissemination of Culture in the Post-truth Era*. Open Humanities Press.
- Bratton, B. (2016). *The Stack: On Software and Sovereignty*. MIT Press.
- Broussard, M. (2018). *Artificial unintelligence: How computers misunderstand the world*. MIT Press.
- Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness*. Duke University Press.
- Chun, W. H. K. (2008). On ‘Sourcery,’ or Code as Fetish. *Configurations*, 16, 299–324.
- Chun, W. H. K. (2011). *Programmed Visions: Software and Memory*. MIT Press.
- Chun, W. H. K. (2017). *Updating to Remain the Same: Habitual New Media*. MIT Press.
- Chun, W. H. K. (2021). *Discriminating Data: Correlation, Neighborhoods, and the New Politics of Recognition*. MIT Press.
- D’Ignazio, C., & Klein, L. F. (2019). *Data Feminism*. MIT Press.
- Dai, A. D., Jiang, W. T., & Zhao, W. B. (2016). Digital humanities as a method: Current status and prospects of Western studies. *Shandong Social Sciences*, (11), 26–33. doi:10.14112/j.cnki.37-1053/c.2016.11.00
- Dix, A., Finlay, J. E., Abowd, G. D., & Beale, R. (2003). *Human-Computer Interaction*. Prentice Hall.
- Emerson, L. (2014). *Reading Writing Interfaces: From the Digital to the Bookbound*. University of Minnesota Press.



- Eubanks, V. (2018). Automating inequality: How high-tech tools profile, police, and punish the poor. St. Martin's Press.
- European Union Agency for Fundamental Rights. (2022). Bias in algorithms - Artificial intelligence and discrimination. Luxembourg: Publications Office of the European Union. doi:10.2811/25847.
- Floridi, L. (2013). The Ethics of Information. Oxford University Press.
- Floridi, L., & Taddeo, M. (2016). What is Data Ethics? Philosophical Transactions of the Royal Society A, 374, 20160360.
- Fuller, M. (2003). Behind the Blip: Essays on the Culture of Software. Autonomedia.
- Galloway, A. (2006). Protocol: How Control Exists after Decentralization. MIT Press.
- Gefen, A., Saint-Raymond, L., & Venturini, T. (2021). AI for Digital Humanities and Computational Social Sciences. In B. Braunschweig & M. Ghallab (Eds.), Reflections on Artificial Intelligence for Humanity (pp. 191–202). Springer.
- Gibbs, F., & Owens, T. (n.d.). The hermeneutics of data and historical writing. In Writing History in the Digital Age. doi:10.2307/j.ctv65sx57.18.
- Goodlad, L. M. E., & Baker, S. (2023, February 21). Now the Humanities Can Disrupt 'AI'. Public Books. Retrieved from <https://www.publicbooks.org/now-the-humanities-can-disrupt-ai/>
- Gray, K. L. (2014). Race, Gender, and Deviance in Xbox Live: Theoretical perspectives from the Virtual Margins. Routledge.
- Han, B.-C. (2017). Psychopolitics: Neoliberalism and New Technologies of Power. Verso.
- HATELAB. (n.d.). HateLab. Retrieved from <https://hatelab.net/>
- Hayles, N. K. (2012). How We Think: Digital Media and Contemporary Technogenesis. University of Chicago Press.
- Heffernan, T. (n.d.). The Ethical Imagination: Humanities Versus Artificial Intelligence (Ethics of AI in Context). Centre for Ethics, University of Toronto. Retrieved from <https://ethics.utoronto.ca/events/teresa-heffernan-ethics-of-ai-in-context/>
- Ji Ho Park & Pascale Fung (2017). One-step and twostep classification for abusive language detection on Twitter. Proceedings of the Workshop on Abusive Language Online.

- Johnson, J. M. (2018). Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads. *Social Text*, 36(4), 57–79.
- Kim, D. (2023, October 16). AI Poses a Threat to the Humanities. Youth Journalism International. Retrieved from <https://youthjournalism.org/ai-poses-a-threat-to-the-humanities/>
- Lakatos, I., & Feyerabend, P. K. (2014). Lakatos and Feyerabend: research programmes and anarchism. In *Historical and Philosophical Foundations of Psychology* (pp. 46-68). Cambridge University Press.
- Liu, A. (2011). The state of the digital humanities: A report and a critique. *Arts and Humanities in Higher Education*, 11(1-2), 8–41. doi:10.1177/1474022211427364
- Lucas Dixon, John Li, Jeffrey Sorensen, Nithum Thain, & Lucy Vasserman. (2018). Measuring and mitigating unintended bias in text classification. *Proceedings of Conference on AI, Ethics, and Society*.
- Luo, F. Y. (2015). Digitization of intangible cultural heritage protection from the perspective of technology. *Sichuan Opera*, (2), 125–128.
- Maarten Sap, Dallas Card, Saadia Gabriel, Yejin Choi, & Noah A. Smith. (2019). The Risk of Racial Bias in Hate Speech Detection. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 1668–1678.
- Manovich, L. (2001). *The Language of New Media*. MIT Press.
- Manovich, L. (2008). *Software Takes Command*. Retrieved from <http://lab.softwarestudies.com/2008/11/softbook.html>
- Marino, M. C. (2006). *Critical Code Studies*. Retrieved from <http://www.electronicbookreview.com/thread/electropoetics/codology>
- Marino, M. C. (2020). *Critical Code Studies*. MIT Press.
- McCarty, W., & Short, H. (2017). Mapping the field. Retrieved from <http://www.org/node/188>
- Mcpherson, T. (2012). Why Are the Digital Humanities So White? Or Thinking the Histories of Race and Computation. Retrieved from <http://dhdebates.gc.cuny.edu/debates/text/29>
- Montfort, N. (2016). *Exploratory Programming for the Arts and Humanities*. MIT Press

- Nascimento, F. R. S., et al. (2022). Unintended bias evaluation: An analysis of hate speech detection and gender bias mitigation on social media using Ensemble Learning. *Expert Systems with Applications*, 201, 117032. doi:10.1016/j.eswa.2022.117032
- Nestor Maslej, L., et al. (2023). The AI Index 2023 Annual Report. AI Index Steering Committee, Institute for Human-Centered AI, Stanford University.
- Noble, S. U. (2018). *Algorithms of Oppression*. New York University Press.
- Nozza, D., Volpetti, C., & Fersini, E. (2019). Unintended Bias in Misogyny Detection. In *IEEE/WIC/ACM International Conference on Web Intelligence* (pp. 149–155).
- O'Neil, C. (2016). *Weapons of math destruction: How big data increases inequality and threatens democracy*. Crown.
- Pratchett, T., & Baxter, S. (2013). *The Long Earth*. Harper.
- Recht, R. (2013). *L'iconologie avant Warburg. Images Re-vues*, Hors-série 4. Retrieved from <http://journals.openedition.org/imagesrevues/2898>
- Spiro, L. (2012). “This is why we fight”: Defining the Value of the Digital Humanities. In M. Gold (Ed.), *Debates in the Digital Humanities*. University of Minnesota Press. Retrieved from <http://dhdebates.gc.cuny.edu/debates/text/13>
- The Computing Research Association. (2023, May 19). CRA Taulbee Survey. Retrieved from <https://cra.org/resources/taulbee-survey/>
- van den Hoven, J. (2010). The Use of Normative Theories in Computer Ethics. In L. Floridi (Ed.), *The Cambridge Handbook of Information and Computer Ethics* (pp. 59–76). Cambridge University Press.
- Wardrip-Fruin, N. (2009). *Expressive Processing: Digital Fictions, Computer Games, and Software Studies*. MIT Press.
- Wiegand, M., Ruppenhofer, J., & Kleinbauer, T. (2019). Detection of Abusive Language: The Problem of Biased Datasets. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)* (pp. 602–608).
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Profile Book.

## Chapter 6

### Conclusion

*"Our inherent cognitive biases make us ripe for manipulation and exploitation by those who have an agenda to push, especially if they can discredit all other sources of information"*

- Lee McIntyre

This thesis broadens the horizons of scholarly investigations into the semantics and symbolism in Partition literature by leveraging computational tools for text analysis to explore the 'language of bias'. This research advocates for the integration of both humanities and computational methodologies to factor in the complexities involved in analyzing bias patterns within intricate narratives. Through the integration of insights from the humanities with computational techniques, it seeks to develop advanced tools capable of more accurately identifying and understanding linguistic biases embedded within complex texts. This approach not only acknowledges the inherent limitations in conventional methods of bias analysis but also aims to overcome them by harnessing computational capabilities to provide deeper insights and more nuanced interpretations of biased language and discourse. Through this interdisciplinary approach, the thesis enhances the understanding of bias in narratives and contributes to the development of more efficient and effective analytical tools in this field.

Through the evaluation and comparison of four metrics for media bias detection and removal, the thesis identifies a gap in current research and introduces a specialized tool for analyzing and mitigating biases in texts related to the 1947 Partition. This tool is designed to address biases at different levels of granularity ranging from words, sentences, and paragraphs to the entire text. It identifies, categorizes, and suggests the manifestation of linguistic biases across ten categories within different genres of texts pertaining to the 1947 Partition, thereby enhancing the overall understanding and handling of biases in historical narratives. The research comprehends the diverse functions of historical

texts, personal testimonies, educational materials such as school textbooks, and social media narratives in shaping and perpetuating enduring mnemonic hegemonies related to the Partition. These mnemonic hegemonies refer to dominant collective memories or interpretations of historical events that influence societal perceptions and understandings. Furthermore, the study acknowledges the significance of varied sources in disseminating knowledge about the Partition, which contributes to the formation and reinforcement of shared narratives and cultural memory surrounding Partition history. The thesis also examines the linguistic patterns used in constructing narratives and manifesting biases within Partition historiography. It specifically concentrates on understanding the significant social repercussions stemming from these biases. This analysis aims to uncover how language choices and biased portrayals in historical accounts influence societal perceptions and attitudes toward the Partition event and its aftermath. Through a methodological probe, the research also investigates the applications, challenges, and ethical concerns associated with utilizing AI-integrated models for digital humanities research. This investigation is particularly focused on the interpretation of individual texts rather than identifying patterns within a larger corpus. Thus, the thesis comprehends the potential of AI technologies to aid in the analysis and interpretation of individual texts, all the while acknowledging and confronting the ethical implications and challenges that accompany such pursuits.

Using close reading with computers as a methodology serves to connect digital method users with scholars involved in traditional literary-critical pursuits, thus narrowing the divide between the digital humanities and broader communities of literature, book history, and textual criticism scholars. This approach facilitates a seamless integration of computational tools and techniques into the study of literature, enabling scholars to analyze texts at a deeper level and uncover insights that may not be immediately apparent through traditional close reading methods alone. The implicit project of Close

Reading with Computers in literary studies positions scholars even “closer” to the text as Eve says “through computational reading [...] this book goes ‘back to the text’” (11). This thesis demonstrates with an analytical outcome how revisiting the text necessitates a reinterpretation of the text, challenging conventional assumptions about the text within computational literary studies. While computational literary studies have typically regarded the text as isolated units divorced from their broader context within the book, this approach emphasizes the need to consider texts within their larger framework.

### **1.1. Summary and Key Findings of the Chapters**

Following the introductory chapter, the second chapter of this thesis elaborates on the methodology employed for linguistic bias analysis, drawing from both social sciences and computer sciences. Through a comparative analysis of four recent bias analysis models, the chapter highlights persistent limitations stemming from issues related to dataset size, diversity, and generalizability. Additionally, it underscores the tendency of existing approaches to overlook insights from the humanities, resulting in deficiencies within bias detection and removal modules. The chapter further outlines the architecture and workflow of a semi-automated target-oriented bias detection and classification tool designed to analyze texts at various levels of granularity, leveraging a diverse dataset sourced from Wikipedia articles, newspaper editorials, and published works on the 1947 Partition of India. Annotations for the dataset were collected through crowd-sourcing micro-tasks, ensuring a wide range of perspectives. The tool's detection module utilizes methods adapted from existing metrics, employing recurrent neural network (RNN) classifiers with a pre-trained word embedding model. It also employs reverse feature analysis to estimate bias probabilities at both word and sentence levels, while incorporating contextual information to enrich the analysis. Finally, the chapter elucidates how the methodology integrates the developed computational tool with close reading techniques to uncover linguistic biases in Partition texts, drawing from frameworks proposed by Martin Paul Eve and Alan Liu.

The first section of the third chapter examines Nisid Hajari's "Midnight's Furies: The Deadly Legacy of India's Partition (2015)," highlighting biases in narrative patterns and archival practices. The linguistic bias analysis underscores susceptibility to misinterpretation, omissions, citation failures, and misleading explanations. A notable deficiency in conveying the profound sentiments driving the call for Pakistan is observed, as the absence of Muhammad Iqbal's influence leaves the historical impetus solely attributed to Jinnah. The narrative is marred by inaccuracies, such as the misattribution of quotes and excessive reliance on secondary sources. Hajari's bias is primarily rooted in slanted objectivity and framing effect, particularly against Jinnah. His religious and political biases are revealed through a lack or an overloading of evidence. Moreover, his reliance on selective evidence and references to present-day incidents introduces biases like misinformation effect, attribution, confirmation, and selection biases. The section also argues that bias is manifested through the chain of causality of events, metaphors employed by the author, and pitfalls in Hajari's approach to history as a perennial conflict between two men who are central to the events. Discussing the subtle religious and gender biases in Hajari's historiography, the section argues that the complex history of 'nationalist and patriarchal fetish on women's sexuality' is reinforced without criticism in Hajar's narrative. Hajri's work is clear evidence of insufficiencies in adequately portraying the narratives of the under-represented within the social and cultural historiography of the 1947 Partition irrespective of f the Oral history wave.

The second section computationally analyzes Urvashi Butalia's "The Other Side of Silence: Voices from the Partition of India," focusing on biases inherent in testimonial narratives and challenges in historical interpretation. The inherent biases within Butalia's work are notably influenced by memory biases and entrenched patriarchal perceptions, particularly in the treatment of women's roles during Partition. The tendency to portray women primarily as victims or symbols of sacrifice, coupled with the perpetuation of gendered stereotypes, reflects the

pervasive societal perception of male heroism and female victimhood. Moreover, the narratives tend to externalize the violence of Partition, attributing it to individuals outside the affected communities, thus perpetuating a communal 'other.' Beyond these biases, the text underscores the impact of childhood amnesia on the reliability of memories recounted by adult witnesses, exposing the fallibility inherent in testimonial narratives. Butalia's narrative strategy, characterized by a thematic organization that transcends individual stories and acknowledges the subjective nature of historical interpretation, offers a nuanced exploration of Partition themes. However, it also raises questions about the author's exclusive agency in narrative representation and the potential loss of dynamics inherent in oral testimonies during their transition to text. Her attempt to challenge the notion of formal history as sacrosanct on the one hand and personal stories or oral histories as less worthy of formal documentation on the other is a thread that runs throughout the book. The chapter also explores the language of describing violence and the use of Holocaust language to capture the political volatility of Partition and its aftermath, drawing parallels with other genocides.

The first section of this chapter studies the Social Studies textbooks from Grade IV to Grade VIII and Pakistan Studies from Grade IX to Grade X compiled by the Punjab Textbook Board and NCERT textbooks from Grade VI to XII using the Semi-Automated tool. It delves into the historical context of school textbooks in India and Pakistan and analyzes their portrayal of the Partition. Through bias analysis, a significant prevalence of positive sentences is identified, indicating a linguistic positivity bias, which reflects writers' attempts to align with societal perceptions by emphasizing objective circumstances and emotional states. Moreover, the victim-centered narrative in both sets of textbooks, coupled with the use of affect-inducing language to depict hardships, suggests a negative attribution bias. Specifically, the higher utilization of passive voice construction in NCERT textbooks is highlighted, potentially impacting emotional association and content



recall for younger readers. The bias-inducing pattern is attributed to the concept of 'historical distance,' where writers strategically manipulate emotional responses to subtly convey their ideological values. Additionally, the section underscores the multifaceted roles of textbooks as tools for memory practice, boundary objects, conveyors of national ideology, and enforcers of homogeneity in understanding Partition. It also discusses various themes such as the Language of Causality and Othering, Selective Amnesia of Conflicts and Violence, and the geopolitical implications of imagined cartographies in Partition textbooks. The subsequent section, titled 'Recollection and Transmedial Reiteration: Twitter and the New Epistemology of 'Subjugated knowledges' about Partition,' investigates the nature of discourses and bias manifestation in tweets related to the 1947 Partition on August 14 and 15, 2021, coinciding with Partition Horrors Remembrance Day. The analysis reveals a potential correlation between the level of hostile comments and the inherent bias in the original tweets, suggesting the use of comment hostility as an indicator of bias. Political bias in tweets often emerges through entities and hashtags, with nouns and noun phrases playing a pivotal role in introducing bias. Bias is further facilitated by sentiment-charged language, primarily employing negative emotion words and emphasizing group identity. Twitter is depicted as a mediated site of contention, a mnemonic medium, a site of false memory formation, and a creator and manipulator of collective consciousness in recalling and reiterating Partition. Moreover, the study emphasizes the transmedial structure of Twitter, which aids in recalling the past within a social network's collective memory.

The fifth chapter traces the discernible phases in the advancement of automated methodologies within Digital Humanities (DH), highlighting AI's role in natural language processing, archiving, image processing, and text analytics. The chapter argues that AI has profoundly transformed data curation and analysis in the humanities, providing researchers with novel insights beyond conventional methods. While AI presents opportunities for collaboration and experimentation

in DH, there are concerns regarding transparency, reproducibility, and the imperative for critical engagement with AI systems. DH researchers are positioned as creative technologists, digital archivists, and critical humanists navigating the intersection of technology and humanistic inquiry. The chapter contends that bias can permeate AI systems during various stages, including data generation, dataset creation, algorithm design, and evaluation. Factors contributing to dataset bias encompass non-representative or skewed data, historical inequities, and subjective or discriminatory labeling, potentially leading to datasets that perpetuate societal injustices. Algorithmic bias may arise due to narrow objectives, biased datasets, lack of consideration for diverse populations, and inadequate auditing, resulting in inaccurate and discriminatory outcomes. At the user level, biased AI outputs may occur due to contextual variations, discriminatory alterations, or biased interpretations. The chapter analyzes biases in AI-integrated offensive speech detection through a case study examining the performance of TILT, an online Hate Speech Detection Solution for Twitter. Findings reveal challenges in accurately identifying offensive content, particularly in the context of dialectical variations and regional language usage in short text strings like tweets. The chapter concludes by proposing ethical guidelines for humanities researchers working with or on AI models.

## **1.2. Limitations**

- The dataset used to create the tool for this study, though it is extendable is target-oriented in its present form which affects its adaptability
- The dataset cannot be termed as fully diverse as user-generated sites like Wikipedia have their structural factors including moderation practices that make them less accessible for marginalized communities.
- Data is not static. The dataset for the tool was curated between the period of 2020-2021. The model has to be retrained

continuously and the training data should also be updated, which is unachievable considering the nature of this project.

- Although the thesis draws its argument from school textbooks, it does not claim to offer a comprehensive survey of textbooks across India and Pakistan. It is acknowledged that even widely used textbooks such as those published by NCERT and the Punjab Textbook Board reach only a fraction of students in Indian and Pakistani schools.
- As Twitter's API has been under a paywall since February 2023, we couldn't extend the study further or couldn't access the history of the tweets from 2021.

### **1.3. Scope for further research**

- The study could be expanded to incorporate memory and cognitive biases, as well as other linguistic and narrative biases. This broader approach would offer a more comprehensive understanding of the various factors influencing human cognition and communication.
- The dataset and tool could be expanded to examine events beyond Partition and include various genres like newspaper reports. This expansion would enhance efforts to combat fake news and deepen understanding of biased content across different contexts.
- The tool can also serve as a valuable resource to understand the intricate patterns of bias and its manifestations evident in an individual author's body of work, facilitating a more profound exploration of their writing.
- Furthermore, by applying the tool to specific genres, particularly social media narratives, researchers can conduct qualitative investigations into the manifestation of bias within these contexts deviating from the 'close reading with computers' approach. A qualitative approach can enhance the understanding

of bias and its manifestation within varied literary and online landscapes, contributing to a richer comprehension of its implications in contemporary discourse.

## References

- Butalia, U. (2018). *The other side of silence: Voices from the Partition of India*. Duke University Press.
- Eve, Martin Paul. *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*. Stanford University Press, 2019.
- Hajari, N. (2017). *Midnight's furies: The deadly legacy of India's Partition*. Amberley.
- McIntyre, L. (2021, July 22). *How cognitive bias can explain post-truth?* The MIT Press Reader.  
<https://thereader.mitpress.mit.edu/how-cognitive-bias-can-explain-post-truth/#:~:text=Our%20inherent%20cognitive%20biases%20make,no%20defense%20against%20post%2Dtruth.>

## APPENDIX – A

### Dataset

No	Work	Author	Category
1	Wikipedia Articles		Wikipedia
2	India Wins Freedom	Maulana Abdul Kalam Azad	Book
3	Partition: The Story of Indian Independence and the Creation of Pakistan	Barney White Spinner	Book
4	Guilty Men of India's Partition	Ram Manohar Lohia	Book
5	The Other Side of Silence	Urvashi Butalia	Book
6	The Great Partition: The Making of India and Pakistan	Yasmin Khan	Book
7	Remnants of Separation: A ghara and a gaz: From Lahore to Amritsar to Delhi	Aanchal Malhotra	Book
8	Midnight's Furies: The deadly legacy of India's partition	Nisid Hajari	Book
9	Pity of Partition	Ayesha Jalal	Book
10	Violent Belongings: Partition, gender, and national culture in postcolonial India	Kavita Daiya	Journal article
11	The Shadow of the Great Game: The Untold Stories of India's Partition	Sarila	Journal article
12	Partition of India: Why 1947	Harshan Kumarasingham	Journal article
13	The Master and the Maharajas: The Sikh Princes and the East Punjab Massacres of 1947	Ian Copland	Journal article
14	Partition, Pakistan, and South Asian History: In Search of a Narrative	David Gilmartin	Journal article
15	Partition of India: The Gurdaspur Dispute	Shivam Sharma	Journal article
16	The Box, the Fish, and Lost Homes	Suranjana Choudhury	Journal article
17	Understanding The Trauma of 1947 India-Pakistan Partition – An Account of Toba Tek Singh	Sejal Choudhary	Journal article
18	Partition: The “Pulsing Heart That Grieved	Gerald James Larson	Journal article
19	Was the partition of India in 1947 inevitable?	Cyril Philips	Journal article
20	Fingernails Torn from Flesh: Intizār Husain, Rām La‘l, and	Daniel Majchrowicz	Journal article

	Travel Writing across the India-Pakistan Border		
21	Unfinished stories of the Partition: Across 75 years	Kamran Asdar Ali and Tabish Khair	Journal article
22	The Origins of the Partition of India, 1936-1947	Judith M. Brown and Anita Inder Singh	Journal article
23	Legacies of Partition for India and Pakistan	Ian Talbot	Journal article
24	"Bound hand and foot and handed over to the caste Hindus': Ambedkar, untouchability and the politics of Partition	Jesús Francisco Cháirez-Garza	Journal article
25	The dark forest of exile: A Dandakaranya memoir and the Partition's Dalit refugees	Debjani Sengupta	Journal article
26	Hidden Histories	Urvashi Butalia	Journal article
27	Memories of a Fragmented Nation: Rewriting the Histories of India's Partition	M. Hasan	Journal article
28	Partition Narratives	M. Hasan	Journal article
29	Muslim Women and the Partition of India: A Historiographical Silence	R. U. Ali	Journal article
30	Remembering partition: women, oral histories and the Partition of 1947	P. Virdee	Journal article
31	Partition Studies: Prospects and Pitfalls	J. Chatterji	Journal article
32	Women during the Partition: Victim and agent	R. Roy	Journal article
33	Forms of Memory: Partition as a Literary Paradigm	A. Bernard	Journal article
34	Partition: The Holocaust	R. Roy	Journal article
35	Memories of Partition: Revisiting Saadat Hasan Manto	S. Tiwari	Journal article
36	The High Politics of India's Partition: The Revisionist Perspective	A. Roy	Journal article
37	I Am Not a Refugee': Rethinking Partition Migration	M. Rahman & W. van Schendel	Journal article

38	The Princely States, the Muslim League, and the Partition of India in 1947	I. Copland	Journal article
39	Redefining the Partition Refugee	P. C. Ghosal	Journal article
40	Mission with Mountbatten	Alan Campbell-Johnson	Book
41	India: The Transfer of Power 1942-1947	N. Mansergh, E. W. R. Lumby, and Penderel Moon	Book
42	Divide and Quit	Penderel Moon	Book
43	The Partition of India: Policies and Perspectives 1935-47	C. H. Philips and M. D. Wainwright	Book
44	Disturbances in the Punjab 1947: A Compilation of Official Documents	Government of Pakistan	Book
45	The Partition of Punjab: A Compilation of Official Documents	Government of Punjab	Book
46	Foundations of Pakistan: All-India Muslim League Documents, 1906-1947	Syed Sharifuddin Pirzada	Book
47	Tryst with destiny' — free but divided	Aijaz Ahmad	Journal Article
48	The 1947 Partition of India: A Paradigm for Pathological Politics in India and Pakistan	Ishtiaq Ahmed	Journal Article
49	Countdown to Partition: The Final Days	Ajit Bhattacharjea	Book
50	The Transfer of Power in South Asia: An Historiographical Review	H. V. Brasted and Carl Bridge	Journal Article
51	Liberty or Death: India's Journey to Independence and Division	Patrick French	Book
52	Empire and Islam: Punjab and the Making of Pakistan	David Gilmartin	Book
53	The Sole Spokesman: Jinnah, the Muslim League and the Demand for Pakistan	Ayesha Jalal	Book
54	The Break-up of British India	B. N. Pandey	Book
55 56	The Prose of Otherness	Gyanendra Pandey	Book
57	Experiment with Freedom, India and Pakistan 1947	Hugh Tinker	Book



58	The Aftermath of Partition in South Asia	Tan Tai Yong and Gyanesh Kudaisya	Book
59	The Partition of India and Mountbatten	L. A. Sherwani	Book
60	The Origins of the Partition of India, 1936-1947	Anita Inder Singh	Book
61	Freedom's Cry: The Popular Dimension in the Pakistan Movement and Partition Experience in Northwest India	Ian Talbot	Book
62	Iqbal, Jinnah and Pakistan: The Vision and the Reality	C. M. Naim (editor)	Book
63	Hindu-Moslem Conflict in India	David G. Mandelbaum	Journal Article

## BIBLIOGRAPHY

### Primary Sources

- #Partition1947 [Tweets]. (2021, August 14-15).
- Butalia, U. (2017). *The other side of silence: Voices from the Partition of India*. Penguin UK.
- Hajari, N. (2016). *Midnight's Furies: The Deadly Legacy of India's Partition*. Penguin UK.
- National Council of Educational Research and Training. (2020). *Textbooks from Grade VI to XII*. New Delhi, India: NCERT.
- Punjab Curriculum & Textbook Board. (2020). *Pakistan Studies textbooks for Grade IX to Grade X*[Textbooks]. Lahore, Pakistan
- Punjab Curriculum & Textbook Board. (2020). *Social Studies textbooks for Grade IV to Grade VIII* [Textbooks]. Lahore, Pakistan

### Secondary Sources

- 1947 archive. (2020). Retrieved from <https://www.1947Partitionarchive.org/>
- Ahmad, A. (1997). " 'Tryst with destiny' — free but divided." *India* (Special Number), 21-28.
- Ahmed, I. (2002). "The 1947 Partition of India: A Paradigm for Pathological Politics in India and Pakistan." *Asian Ethnicity*, 3(1).
- Alan Liu. (2012). The state of the digital humanities: A report and a critique. *Arts and Humanities in Higher Education*, 11(1–2), 8–41. <https://doi.org/10.1177/1474022211427364>
- Aleksandrova, D., Lareau, F., & Ménard, P. A. (2019). Multilingual sentence-level bias detection in Wikipedia. In *Proceedings of the International Conference on Recent Advances in Natural Language Processing* (pp. 42-51).
- Alessio, D., & Allen, M. (Year). Media bias in presidential elections: A meta-analysis. *Journal of Communication*, 50(4), 133–156.
- Ali, K. A., & Khair, T. (2022). Unfinished stories of the Partition: Across 75 years. *Journal of Commonwealth Literature*, 57(3), 562–573. <http://dx.doi.org/10.1177/00219894221115911>.

- Ali, R. U. (2009). Muslim Women and the Partition of India: A Historiographical Silence. *Islamic Studies*, 48(3), 425–436.
- Azad, A. K. (2021). *India wins freedom*. Prabhat Prakashan.
- Azad, M. A. K. (1988). *India Wins Freedom*. New Delhi: Orient Longman.
- Baly, R., Karadzhov, G., Alexandrov, D., et al. (2018). Predicting factuality of reporting and bias of news media sources. arXiv preprint arXiv:1810.01765
- Benjamin, R. (2019). *Race After Technology: Abolitionist Tools for the New Jim Code*. Polity Press.
- Bernard, A., & (2010) برنارد, أ. Forms of Memory: Partition as a Literary Paradigm / أشكال الذاكرة: التقسيم بوصفه نموذجاً أدبياً. *Alif: Journal of Comparative Poetics*, 30, 9–33.
- Bharathi, S. V., & Geetha Angelina. (2019). Determination of news biasedness using content sentiment analysis algorithm. *Indonesian Journal of Electrical Engineering and Computer Science*, 882-889.
- Bhattacharjea, A. (1998). *Countdown to Partition: The Final Days*. New Delhi: HarperCollins.
- Bolukbasi, T., Chang, K. W., Zou, J. Y., Saligrama, V., & Kalai, A. (2016). Man is to Computer Programmer as Woman is to Homemaker? Debiasing Word Embeddings. In *Advances in Neural Information Processing Systems* 29.
- Bolukbasi, T., Chang, K., Zou, J. Y., et al. (2016). Man is to computer programmer as woman is to homemaker? debiasing word embeddings. *Advances in Neural Information Processing Systems*, 29.
- Brasted, H. V., & Bridge, C. (1994). "The Transfer of Power in South Asia: An Historiographical Review." *South Asia*, 17(1), 93-114.
- Brown, J. M., & Singh, A. I. (1989). The Origins of the Partition of India, 1936-1947. *American Historical Review*, 94(3), 832. <http://dx.doi.org/10.2307/1873913>.
- Bruce, R., Bell, M., & Martin, M. (2002). Learning subjective language. Computer Science Technical Report, University of Pittsburgh.
- Butalia, U. (1995). Hidden histories. *Index on Censorship*, 24(4), 81–89. <http://dx.doi.org/10.1177/030642209502400423>.
- Butalia, U. (2017). *The other side of silence: Voices from the Partition of India*. Penguin UK.

- Campbell-Johnson, A. (1982). *Mission with Mountbatten*. London: Hale; reprinted, New York: Atheneum, 1986.
- Carder, B., & Rooney, J. (2018). Thinking, fast and slow. *The Journal for Quality and Participation*, 41(2), 31–32.
- Cháirez-Garza, J. F. (2018). 'Bound hand and foot and handed over to the caste Hindus': Ambedkar, untouchability and the politics of Partition. *Indian Economic & Social History Review*, 55(1), 1–28. <http://dx.doi.org/10.1177/0019464617745925>.
- Chatterji, J. (2014). Partition Studies: Prospects and Pitfalls. *The Journal of Asian Studies*, 73(2), 309–312.
- Chen, W., Khatib, K., Sten, B., et al. (2020). Detecting media bias in news articles using Gaussian bias distributions. arXiv preprint.
- Chen, W., Khatib, K., Wachsmuth, H., et al. (2020). Analyzing political bias and unfairness in news articles at different levels of granularity. arXiv preprint.
- Chen, W., Wachsmuth, H., Khatib, K., et al. (2018). Learning to flip the bias of news headlines. In *Proceedings of 11th Conference of Natural Language Generation* (pp. 79-88).
- Choudhary, S. (2022). Understanding The Trauma of 1947 India-Pakistan Partition – An Account of Toba Tek Singh. *International Journal of English Literature and Social Sciences*, 7(5), 111–113. <http://dx.doi.org/10.22161/ijels.75.18>.
- Choudhury, S. (2020). The Box, the Fish, and Lost Homes. *Migration and Society*, 3(1), 259–261. <http://dx.doi.org/10.3167/arms.2020.030124>.
- Copland, I. (1991). The Princely States, the Muslim League, and the Partition of India in 1947. *The International History Review*, 13(1), 38–69.
- Copland, I. (1998). "The Further Shores of Partition: Ethnic Cleansing in Rajasthan, 1947." *Past and Present*, 160.
- Copland, I. (2002). The Master and the Maharajas: The Sikh Princes and the East Punjab Massacres of 1947. *Modern Asian Studies*, 36(3), 657–704. <http://dx.doi.org/10.1017/s0026749x02003050>.
- Culler, J. (2011). The closeness of Close reading. *Adfl*, 41(3), 8–13. <https://doi.org/10.1632/adfl.41.3.8>.
- Daiya, K. (2011). *Belongings: Partition, gender, and national culture in post-colonial India*. Temple University Press.
- Das, A., Dantcheva, A., & Bremond, F. (2018). Mitigating bias in gender, age, and ethnicity classification: A multi-task

- convolution neural network approach. In European Conference of Computer Vision Workshops.
- D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press.
- Entman, R. (1993). Framing: Towards clarification of a fractured paradigm. In McQuail, D. (Ed.), *McQuail's Reader in Mass Communication Theory* (pp. 390-397).
- Ethayarajh, K. (2019). Attenuating Bias in Word Representations: An Evaluation of Euphemisms. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*.
- Ethayarajh, K., Duvenaud, D., & Hirst, G. (Year). Understanding undesirable word embedding associations. *arXiv preprint*.
- Eve, M. P. (2019). *Close Reading with Computers: Textual Scholarship, Computational Formalism, and David Mitchell's Cloud Atlas*. Stanford University Press.
- Fan, L., White, M., Sharma, E., et al. (2019). In plain sight: Media bias through the lens of factual reporting. *arXiv preprint*.
- Frampton, G., Whaley, P., Bennett, M., et al. (2022). Principles and framework for assessing the risk of bias for studies included in comparative quantitative environmental systematic reviews. *Environmental Evidence*, 11, 12. <https://doi.org/10.1186/s13750-022-00264-0>
- Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, 345(6203), 1502–1505.
- French, P. (1997). *Liberty or Death: India's Journey to Independence and Division*. London.
- Gangula, R., & Mamidi (2018). Resource creation towards automated sentiment analysis in Telugu (a low resource language) and integrating multiple domain sources to enhance sentiment prediction. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation*.
- Gentzkow, M., & Shapiro, J. M. (Year). What drives media slant? Evidence from US daily newspapers. *Econometrica*, 78(1), 35–71.
- Ghosal, P. C. (2014). Redefining the Partition Refugee. *Proceedings of the Indian History Congress*, 75, 546–558.

- Gilmartin, D. (1988). *Empire and Islam: Punjab and the Making of Pakistan*. London: Oxford University Press.
- Gilmartin, D. (1998). Partition, Pakistan, and South Asian History: In Search of a Narrative. *Journal of Asian Studies*, 57(4), 1068–1095. <http://dx.doi.org/10.2307/2659304>.
- Gonen, H., & Goldberg, Y. (2019). Analyzing Gender Bias in Harry Potter with Word Embeddings. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing*.
- Gonen, H., & Goldberg, Y. (2019). Lipstick on a pig: Debiasing methods cover up systematic gender biases in word embeddings but do not remove them. *arXiv preprint*.
- Gordon, E. W., Miller, F., & Rollock, D. (1990). Coping with communication centric bias in knowledge production in the social sciences. *Educational Researcher*, 19(3), 14–19.
- Government of Pakistan. (1995). *Disturbances in the Punjab 1947: A Compilation of Official Documents*. Islamabad.
- Government of Punjab. (1993). *The Partition of Punjab: A Compilation of Official Documents*. Lahore: Sang-e-Meel.
- Greene, S., & Resnik, P. (2009). More than words: Syntactic packaging and implicit sentiment. In *Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics* (pp. 503-511).
- Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, 120(4), 1191–1237.
- Hajari, N. (2015). *Midnight's furies: The deadly legacy of India's Partition*. Houghton Mifflin Harcourt.
- Hamborg, F. (2020). Media bias, the social sciences, and NLP: Automating frame analyses to identify bias by word choice and labeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: Student Research Workshop* (pp. 79-87).
- Hamborg, F., Donnay, K., & Bela, G. (2019). Automated identification of media bias in news articles: An interdisciplinary literature review. *International Journal on Digital Libraries*, 20(4), 391–415.

- Hamborg, F., Meuschke, N., & Gipp, B. (2020). Bias-aware news analysis using matrix-based news aggregation. *International Journal on Digital Libraries*, 21(2), 129–147.
- Hasan, M. (1993). *India's Partition: Process, Strategy and Mobilization*. New Delhi: Oxford University Press.
- Hasan, M. (1997). *Legacy of a Divided Nation: India's Muslims Since Independence*. New Delhi: Oxford University Press.
- Hasan, M. (1998). "Memories of a Fragmented Nation: Rewriting the Histories of India's Partition." *Economic and Political Weekly*, 33(41), 2662-2668.
- Hasan, M. (1998). *Memories of a Fragmented Nation: Rewriting the Histories of India's Partition*. *Economic and Political Weekly*, 33(41), 2662–2668. <http://www.jstor.org/stable/4407270>.
- Hasan, M. (2000). *Inventing Boundaries: Gender, Politics and the Partition of India*. Delhi: Oxford University Press.
- Hasan, M. (2004). Partition Narratives. *Oriente Moderno*, 23(84)(1), 103–130.
- Hodson, H. V. (1993). *The Great Divide: Britain-India-Pakistan*. Karachi: Oxford University Press.
- Hunter, J., Schmidt, F., & Jackson (1982). Meta-analysis: Cumulating research findings across studies. *Educational Researcher*, 15(8), 20–21.
- Iyyer, M., Enns, P., Graber, J. B., et al. (2014). Political ideology detection using recursive neural networks. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (pp. 1113-1122).
- Jalal, A. (1974). "Conjuring Pakistan: History as Official Imagining." *International Journal of Middle East Studies*, 8(4).
- Jalal, A. (1985). *The Sole Spokesman: Jinnah, the Muslim League and the Demand for Pakistan*. Cambridge: Cambridge University Press.
- Jalal, A. (1996). "Secularists, Subalterns, and the Stigma of 'Communalism': Partition Historiography Revisited." *Indian Economic and Social History Review*, 33(1), 93-103.
- Jalal, A. (1998). "Nation, Reason and Religion: Punjab's Role in the Partition of India." *Economic and Political Weekly*, 33(32), 2183-2190.
- Jalal, A. (2013). *The pity of Partition*. Princeton University Press.

- Jha, P. S. (1996). *Kashmir, 1947: Rival Versions of History*. Delhi: Oxford University Press.
- Khan, Y. (2017). *The great Partition: The making of India and Pakistan*. Yale University Press.
- Khosla, G. D. (1989). *Stern Reckoning: A Survey of the Events Leading Up to and following the Partition of India*. New Delhi: Oxford University Press.
- Kiesel, J., Mestre, M., Shukla, R., et al. (2019). Semeval-2019 task 4: Hyperpartisan news detection. In *Proceedings of the 13th International Workshop on Semantic Evaluation* (pp. 829-839).
- Kiritchenko, S., Zhu, X., Mohammad, S. M., & Cherry, C. (2018). Examining Gender and Race Bias in Two Hundred Sentiment Analysis Systems. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*.
- Krane, V. (2019). Bias in the Flesh: A Critical Feminist Technoscience Study of Sex Testing in Sport. *Science, Technology, & Human Values*, 44(6), 1019-1044.
- Kumar, R. (1998). "Partition Historiography: Some Reflections." In R. Mahan (Ed.), *India's Partition, Prelude and Legacies* (pp. xx-xx). Jaipur: Rawat.
- Kumarasingham, H. (2013). Partition of India: Why 1947? *Asian Affairs*, 44(2), 302–303.  
<http://dx.doi.org/10.1080/03068374.2013.795297>.
- Lamb, A. (1994). *Birth of a Tragedy: Kashmir 1947*. Hertingfordbury: Roxford Books.
- Larson, G. J. (2013). Partition: The “Pulsing Heart that Grieved”. *Journal of Asian Studies*, 73(1), 5–8.  
<http://dx.doi.org/10.1017/s0021911813001666>.
- Lash, T. L., Fox, M. P., Cooney, D., Lu, Y., & Forshee, R. A. (2016). Quantitative Bias Analysis in Regulatory Settings. *American Journal of Public Health*, 106(7), 1227-1230.
- Li, C., & Goldwasser, D. (2019). Encoding social information with graph convolutional networks for political perspective detection in news media. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics* (pp. 2594-2604).
- Lin, W., Wilson, T., Wiebe, J., et al. (2006). Which side are you on? Identifying perspectives at the document and sentence levels. In *Proceedings of the Tenth Conference on Computational Natural Language Learning* (pp. 109-116).



- Liu, A., & Svensson, P. (Eds.). (2019). *Decolonizing Digital Humanities: A Reading and Dialogue*. University of Minnesota Press.
- Lohia, R. (2020). *Guilty Men of India's Partition*. B.R Publishing.
- Lohia, R. M. (1960). *Guilty Men of India's Partition*. Allahabad: Navahind.
- Low, D. A., & Brasted, H. (1998). *Freedom, Trauma, Continuities: Northern India and Independence*. New Delhi: Sage Publications.
- Lumby, E. W. R. (1954). *The Transfer of Power*. London: Allen.
- Majchrowicz, D. (2020). Fingernails Torn from Flesh: Intizār Husain, Rām La'l, and Travel Writing across the India-Pakistan Border. *Journal of Urdu Studies*, 1(2), 241–267.  
<http://dx.doi.org/10.1163/26659050-12340012>.
- Malhotra, A. (Year). Remnants of a separation: A ghara and a gaz: From Lahore to Amritsar to Delhi. *South Asian Popular Culture*, 14(1-2), 89–100.
- Malmasi, S., & Dras, M. (2020). An Analysis of the Relationship Between Gender and Authorship in Computer Science Research. arXiv preprint arXiv:2008.06666.
- Mandelbaum, D. G. (1947). "Hindu-Moslem Conflict in India." *The Middle East Journal*, 1(4), 369-385.
- Mansergh, N., Lumby, E. W. R., & Moon, P. (Eds.). (1970-1983). *India: The Transfer of Power 1942-1947* (12 vols.). London: HMSO.
- McCarthy, J., Titarenko, L., McPhail, C., et al. (2008). Assessing stability in the patterns of selection bias in newspaper coverage of protest during the transition from communism in Belarus. *Mobilization: An International Quarterly*, 13(2), 127–146.
- Michel, A. A. (1967). *The Indus Rivers: A Study of the Effects of Partition*. New Haven: Yale University Press.
- Miyuki Yamada, Yuichi Murai, & Ichiro Kumagai. (2013). Story Visualization of Novels with Multi-theme Keyword Density Analysis. *Journal of Visualization*, 16(3), 247–257.  
<https://doi.org/10.1007/s12650-013-0163-4>.
- Moon, P. (1998 [1961]). *Divide and Quit*. Delhi: Oxford UP.
- Moore, R. J. (1982). *Escape from Empire: The Atlee Government and the Indian Problem*. Oxford: Oxford University Press.
- Mosley, L. (1961). *The Last Days of the Raj*. London.

- Munson, S., & Resnick, P. (Year). Presenting diverse political opinions: How and how much. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1457-1466).
- Munson, S., Zhou, D., & Resnick, P. (2009). Sidelines: An algorithm for increasing diversity in news and opinion aggregators. In *Proceedings of the International AAAI Conference on Web and Social Media* (pp. 130-137).
- Naim, C. M. (1979). *Iqbal, Jinnah and Pakistan: The Vision and the Reality*. Syracuse: Syracuse University Press.
- Noble, S. U. (2018). Reckoning with Bias in Algorithmic Culture. In *Data & Society: Points*.
- Owen, R. (Year). Reader bias. *Jama*, 247(18), 2533–2534.
- Pandey, B. N. (1969). *The Break-up of British India*. London: Macmillan.
- Pandey, G. (1994). "The Prose of Otherness." In D. Arnold & D. Hardiman (Eds.), *Subaltern Studies VIII: Essays in Honour of Ranajit Guha* (pp. 188-221). Delhi: Oxford University Press.
- Park, S., Ko, M., Kim, J., et al. (2011). The politics of comments: Predicting political orientation of news stories with commenters' sentiment patterns. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (pp. 113-122).
- Pennebaker, J. W., Boyd, R. L., Jordan, K., et al. (2015). The development and psychometric properties of LIWC.
- Pennington, S., Manning, C. D. (2014). GloVe: Global vectors for word representation. In *Proceedings of the 2014 conference on Empirical Methods in Natural Language Processing* (pp. 1532-1543).
- Philips, C. (1986). Was the partition of India in 1947 inevitable? *Asian Affairs*, 17(3), 243–251.  
<http://dx.doi.org/10.1080/03068378608730233>.
- Philips, C. H., & Wainwright, M. D. (Eds.). (1970). *The Partition of India: Policies and Perspectives 1935-47*. London: Allen and Unwin.
- Pirzada, S. S. (Ed.). (1970). *Foundations of Pakistan: All-India Muslim League Documents, 1906-1947* (Vols. 1-2). Karachi: National Publishing House.

- Posner, M. (2016). Whose Stories Are Heard? Exploring Bias in Digital Humanities Projects. In *Debates in the Digital Humanities 2016*.
- Pottker, H. (Year). News and its communicative quality: The inverted pyramid when and why did it appear? *Journalism Studies*, 4(4), 501–511.
- Preethi, P. G., Uma, V., et al. (2015). Temporal sentiment analysis and causal rules extraction from tweets for event prediction. *Procedia Computer Science*, 48, 84–89.
- Pryzant, R., Martinez, R. D., Dass, N., et al. (2020). Neutralizing subjective bias in text. In *Proceedings of the AAAI Conference on Artificial Intelligence* (pp. 480-489).
- Rahman, M., & van Schendel, W. (2003). “I Am Not a Refugee”: Rethinking Partition Migration. *Modern Asian Studies*, 37(3).
- Rashkin, H., Choi, E., Jang, J., et al. (2017). Truth of varying shades: Analyzing language in fake news and political fact-checking. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing* (pp. 2931-2937).
- Recasens, M., Mizil, C., & Jurafsky, D. (2013). Linguistic models for analyzing and detecting biased language. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics* (pp. 1650-1659).
- Riloff, E., & Wiebe, J. (2003). Learning extraction patterns for subjective expressions. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing* (pp. 105-112).
- Roy, A. (1990). The High Politics of India’s Partition: The Revisionist Perspective. *Modern Asian Studies*, 24(2), 385–408.
- Roy, R. (2010). Partition: The Holocaust. In *South Asian Partition Fiction in English: From Khushwant Singh to Amitav Ghosh* (pp. 33–62). Amsterdam University Press.
- Roy, R. (2010). Women during the Partition: Victim and agent. In *South Asian Partition Fiction in English: From Khushwant Singh to Amitav Ghosh* (pp. 63–88). Amsterdam University Press.
- Samaddar, R. (1997). *Reflections on Partition in the East*. New Delhi: Vikas Publishing House.
- Sarila, N. S. (2017). *The shadow of the great game: The untold story of India’s Partition*. Constable.

- Seaver, N. (2017). Critical Algorithm Studies: A Reading List. Retrieved from <https://www.nickseaver.net/algorithm-studies.html>
- Sengupta, D. (2022). The dark forest of exile: A Dandakaranya memoir and the Partition's Dalit refugees. *Journal of Commonwealth Literature*, 57(3), 520–531. <http://dx.doi.org/10.1177/00219894221115908>.
- Sharma, S. (2021). Partition of India: The Gurdaspur Dispute. *Journal of University of Shanghai for Science and Technology*, 23(07), 1270–1278. <http://dx.doi.org/10.51201/jusst/21/07271>.
- Sherwani, L. A. (1986). *The Partition of India and Mountbatten*. Karachi: Council for Pakistan Studies.
- Singh, A. I. (1987). *The Origins of the Partition of India, 1936-1947*. New Delhi: Oxford University Press.
- Spinner, B. W. (2017). *Partition: The Story of Indian Independence and the Creation of Pakistan*. Simon Schuster UK.
- Stefanie Posavec. (2009). *Writing Without Words*. Retrieved from [www.stefanieposavec.com/writing-without-words](http://www.stefanieposavec.com/writing-without-words).
- Talbot, I. (1996). *Freedom's Cry: The Popular Dimension in the Pakistan Movement and Partition Experience in Northwest India*. Karachi: Oxford University Press.
- Talbot, I. (2019). Legacies of Partition for India and Pakistan. *Politeja*, 16(2), 7–25. <http://dx.doi.org/10.12797/politeja.16.2019.59.01>.
- Talbot, I. (2019). Legacies of the Partition for India and Pakistan. *Politeja*, 59, 7–25.
- Talbot, I., & Singh, G. (1999). *Region and Partition: Bengal, Punjab and the Partition of the Subcontinent*. Karachi: Oxford University Press.
- Tanya E. Clement. (2013). Text Analysis, Data Mining, and Visualizations in Literary Scholarship. In *MLA Commons, Literary Studies in the Digital Age: An Evolving Anthology*. Retrieved from <https://dlsanthology.mla.hcommons.org/text-analysis-data-mining-and-visualizations-in-literary-scholarship>.
- Tinker, H. (1967). *Experiment with Freedom, India and Pakistan 1947*. London: Oxford University Press.
- Tiwari, S. (2013). Memories of Partition: Revisiting Saadat Hasan Manto. *Economic and Political Weekly*, 48(25), 50–58.
- Virdee, P. (2013). Remembering partition: women, oral histories and the Partition of 1947. *Oral History*, 41(2), 49–62.

- West Punjab Government. (1948). Note on the Sikh Plan. Lahore: Superintendent, Government Printing.
- West Punjab Government. (1948). RSS in the Punjab. Lahore: Superintendent, Government Printing.
- West Punjab Government. (1948). The Sikhs in Action. Lahore: Superintendent, Government Printing.
- Yong, T. T., & Kudaisya, G. (2000). The Aftermath of Partition in South Asia. London: Routledge.