Examining Notions of Racism in Science, Technology, Engineering, and Mathematics: A Quantitative Historical Analysis

# Abstract

In discussions of racism across the sciences, medicine, and technology are important critiques of dominant ideas and practices. Dominant ideas are artifacts of the historical foundations of a discipline, notable in references to “founding fathers” and a discipline’s foundational theories and beliefs. Dominant practices, in the more contemporary sense, are modern representations of the professional priorities resulting from a discipline’s dominant ideas – which often serve to maintain a status quo. In this article, the authors will employ mixed methodologies to examine how critically oriented scholars have challenged dominant ideas and practices in relation to racism in STEM (science, technology, engineering, and mathematics.s). In doing so, the authors develop a framework and critical data scientific method to highlight varying levels and typologies of racism, as well as historical patterns of white racial violence in both theory and practice.

*Keywords*: racism, STEM, citation analysis, synthesis, race and ethnicity

# Introduction

A global network of academics have identified how racism deeply impacts the well-being of communities, and the families and individuals that make up those communities. As these scholars increasingly cross disciplinary boundaries to make sense of how racism is both studied and understood in different contexts and across various ecosystems, overwhelming evidence is generated to show racism’s widespread impact. This interdisciplinary body of scholarship not only identifies how racism holds consequences that are unique to each of the contexts in which racists attitudes and actions manifest, they also frame racism’s similarities across social and political boundaries. As a working example, scholars in sociology have structured the concept and meaning of racism through individual perspectives (a single person’s experience), what Bonilla-Silva refers to as “psychological phenomena” (Bonilla-Silva, 1997) and Shiao & Woody as “individual attitudes” (Shiao & Woody, 2021), which are differentiated from two other main perspectives: cultural perspectives and structural perspectives. Other scholars consider intersectional approaches in their studies of racism. For example, Nzinga (2020) situates “an intersectional feminist [lens] that recognizes the multiplicity of not only people’s identities and experiences but also institutionalized hierarchies of power such as race and gender, as well as stratified and segmented labor structure” (p. 11). These diverse perspectives are representative of the varied meanings that scholars attribute to racism and the intersectional perspectives that may be taken on in a particular domain of study.

However, the meanings that scholars attribute to racism across disciplinary contexts may be informed by a constantly evolving set of dynamic indicators which change as new laws, policies, and systems are developed. As scholars develop scholarship around recent policies, such as those regarding the development of ethnically-focused curricula in the United States, such as African American and Asian American studies, examine the structures of racism differently than, say, those scholars studying racism in the medical context. Despite any disciplinary differences, many of the studies in this broader body of scholarship extend the general understanding of racism while also providing acute insights into different settings which make up the various set of indicators which point researchers to an improved analysis of racism, and more generally contribute to ongoing efforts to combat racism.

As this body of scholarship on racism grows in both the STEM and STEM education domains, the extensions to the work of scholars from other disciplines, and those who have identified, or at least attributed a conceptual map on the different meanings that are attributed to racism within a field, can offer insight to both the similarities and differences across disciplines *within* a domain. Much of the insights through which these large-scale citation analyses develop are based, however, are from research within a particular journal subject or thematic index in the bibliographic content of an academic article’s citation. For example, in many of the well-known academic resource databases (such as EBSCO or PubMed), results for various logic commands can be filtered by academic discipline. In the case of STEM, however, it is not clear through which channels emerging scholarships on racism in STEM is centered.

To make better sense of the general scholarship on racism in STEM and STEM education content, we analyze two large-scale databases that provide insights for understanding how the evolving body of scholarship on racism in STEM is distributed. These distributions come in a variety of forms, some regarding the geographical distribution, as well as the distribution of various keywords in context. Questions also remain regarding the additional considerations about the general social structures of an extending academic area of study. Given these needed considerations, this study aims to map and describe the state of the intellectual structure and emerging notions in research on racism across the “science, technology, engineering, and mathematics” areas of study.

# Theoretical Framework

## Dimensions of complexity in the social sciences

Modern studies in quantitative history engage in diverse computational analyses using different types of data (Furet, 1971; Haskins & Jeffrey, 2011). In the original disciplinary markers tied to quantitative historical study, economists and historians sought to make sense of various historical trends in larger scaled data sets. These analyses included both the collection, comparison, and modeling of data sets that presented a nuanced understanding that, due to its potentially reductive nature, has only recently gained traction as a viable alternative for more critical engagement within research scholarship.

In their study on quantitative historical analysis, Turchin et al. (2017) examine how characteristics of social scale across various geographical boundaries and in different countries exhibit strong evolutionary relationships, in one aspect, in terms of how different ideas and concepts may be studied by various human social groups and societies. The authors’ argue that as societies develop, so does the strength of how we might measure different “complexity characteristics.” Using a global historical data bank, the authors found key aspects of human social organization tend to co-evolve in predictable ways,” pointing to “huge commonalities” in the way that human societies evolve (p. E147). In the discussion on the expanding discourses on research on racism in STEM globally, Turchin et al. (2017)’s study presents the component of the theoretical framework which applies to the dimensions of complexity, a term which they note has a “that has many colloqiual meanings, and there are many valied ways in which it could be applied to human social organization” (p. 147). In our analysis, this social organization is rested up the expansion of a research domain across a set of disciplines from a global perspective, which was the primary focus of Turchin et al. (2017)’s study. Our study integrated the global component of the quantitative historical analysis to inform the analysis and production of scholarship on racism in the global context.

## Specificity across the social sciences

We develop methods from the theoretical framework based on the approaches identified by scholars regarding conceptions of race and notions of a needed specificity in the study of racism. One primary example in sociology regards Du Bois’ intellectual inquiries (Besek et al., 2021), a majority of which focus on racism, to examine the various dimensions of complexity of the social conditions that surround racism. This “Du Boisian” framing (Hughey, 2023; Monteiro, 2000; K. Ojeh, 2024; Shuford, 2017) informs both critiques of and methodological opportunities for a dynamic investigation of racism in a particular disciplinary context, and across disciplines and domains. These scholars notice how conceptions of race and racism contribute to both connected but niche conceptualizations that balance scholars’ use of broader ideas, terms, and keywords and their analysis of these concepts in specific contexts. Du Bois is noted as one of the early critical scholars of race who contributed to this duality in the methodological real of sociological study and is now regarded as one of the founding fathers of sociology (K. S. Ojeh & Wright, 2024), an interdisciplinary scholar who also integrated computational methods into the sociological lexicon during its early development (Besek et al., 2021; Monroe-White & Lecy, 2023).

Besek et al. (2021) note that “DuBois directly reflected upon this interdisciplinary approach in an essay that remained unpublished until the first year of the twenty-first century ((Du Bois, 2000); also see (Go, 2020)) entitled ‘Sociology Hesitant’, which provides a window into how he fit his cautiously realist picture of science within his own sociological practice. Its title refers to a contemporary identity crisis in sociology, one derived from open questions regarding what sociology is per se, what sociologists should actually be doing, and how sociology should relate to other, more established, disciplines,” citing Du Bois, that “[sociologists] had been heretofore ‘hesitant’ to acknowledge the limits of sociology – that it will never be an exact science and that ‘this is a world of Chance as well as Law’ (p. 7–8).

These extensions relate to the ongoing critiques within the discipline of sociology regarding the conceptual and methodological approach to study inequality and social systems and structures. Despite our analysis of and reliance on the ways that sociologist and the body of work in sociology focused on racism, we take into account these dynamic shifts in our development of the theoretical framework for the study. Green & Wortham (2015) provide a related discourse around the scholars’ ongoing neglect of Du Bois’ body of scholarship, especially within the domain of sociology, which focused increasingly on the various systems through which the conceptions of race and racism developed by Du Bois would service the broader field of sociological research.

Ojeh (2024), following Wright (2020), refines a subset of contemporary studies of sociology within the frame of a “Black sociology,” where the analysis of systems and social power are rooted more deeply in the foundations of the Atlanta Sociological Laboratory (ASL) at Atlanta University. In this analysis, Ojeh (2024) provides an extension to the field’s understanding of Black sociology as a distinct area of study, a specificity within sociology, using seven principles: “(1) Defines Black sociology as the scientific analysis of systems of oppression and social power; (2) Avoid race deficit approaches derived from mainstream sociology; (3) The research centers on the experiences of Black Americans and people of African descent throughout the Diaspora; (4) The research implements an interdisciplinary and community-based approach; (5) Research about Black people is understood as contextually valid and valuable to scientific research; (6) Scholars use their scholarship to promote social change through social policy and practitioner engagement; (7) The research is conducted by scholars who implement a Black sociological standpoint” (p. 39).

While Ojeh (2024) presents a focused frame on sociological research in the context of “Black sociology,” the conceptual approach taken in the study contributes to our understanding of complexity in the general study of racism as examined in social structures, and the analysis of racism *within* a racial group. This dual conception, that at the intersection of race and racism, in extensions of sociological thought are articulated in other studies related to different groups, such as anti-Asian racism (Yu et al., 2024), research on racial equity research and Indigenous communities (Gordon & Around Him, 2024), and related studies on the dynamic shifts in studies of racism (Neville et al., 2024).

## Variation in the study of racism by discipline and within a domain

STEM is a domain of study under which many disciplines have been situated. There have also been ongoing calls and the use of extended acronyms by national bodies to help foster a more interdisciplinary focus across the various STEM disciplines; one example is the movement to include The Arts in STEM using the STEAM (science, technology, engineering, arts, and mathematics) acronym. There have been a host of other acronyms situated and used by scholars across the globe. As the STEM domain expands in both its reach and its ability to offer more than valuable insights into the hidden aspects of racism throughout society, there is a need to consider the ways that traditional insights into a specific discipline of study may shift a broader domain of interest. Some questions that arise from these differences may be, for example: which scholars discuss racism in the STEM domain? How many of the scholars who publish on racism in STEM come from a particular discipline or sub-area of study within the domain?

The interdisciplinary framing of STEM and the term’s use across both geographical and theoretical boundaries establishes a need to explore how discussions of racism have been codified in STEM" and in the broader collection of critixal\footnote{An `x' is used here primarily as a literary tool -- and as a variable indicator -- to note that various meanings and conceptions ofcritical” can be identified in scholarly research. It is beyond the scope of this study to define the term `critixal’ but we point readers to (2010) for a sample of possible discourses around the general treatment of the term critical in education research.} studies on science, technology, engineering, and mathematics. Here, we notice that studies of racism have also been situated in broader engagements with the sciences, technology, medicine, engineering, mathematics, psychology, and the many STEM related disciplines. We use the term to identify the context-rich evidence identifiable in the syntactical structures and keywords used by researchers across disciplines, and their diverse representations of racism (, 2021). These representations also relate to conceptual and methodological inquiries which serve to inform how different meanings may be constructed. These questions, given the increasingly computationally-dependent nature in studies of bibliometric data and citations, systematics reviews, and meta-analytic studies, calls for further engagement with different analytic priorities across disciplinary areas (, 2021).

The work in traditional cliometrics provides a large portion of the frames through which these questions may be explored. However, there is still a need to not only make sense of but to implement new ways of thinking about deriving the information about the various disciplines within a domain, and making sense of how different disciplines and their understanding of a particular topic of study (in this case racism) relate to the ways that scholars across other disciplines make sense of the concept of racism and its various meanings across different situations and contexts.

For example, Crisp et al. (2024) conduct a systemic review of the varied theoretical perspectives guiding research on racism in higher education. In their study, the authors conduct a content analysis that integrated interdisciplinary perspectives and methods on the various theories used to study race and racism between 2010 and 2019. In their work, they found the use of critical race theory (CRT) perspectives were particularly situated as central to study in this area. The findings by Crisp et al. (2024) relate to the use of those perspectives situated across CRT frameworks and the systematic methods in understanding how particular terms are studied. Shiao & Woody (2021) make this intersection in their study on the meaning of racism in STEM, which lends to the core framing on the conceptualization and meanings of racism within the complex ways that studies of racism may be situated.

Shiao & Woody (2021)’s “three meanings of racism expand on Bonilla-Silva (1997)’s distinction between conceptualizations of racism as (1) an “ultimately psychological phenomenon,” (2) the “cultural processes [of an] all powerful ideology,” and (3) “the foundation or structure of the social system” by observing that sociologists typically conceptualize each meaning in less restrictive terms. To examine racism across disciplines and within a domain, we build on Shiao & Woody (2021)’s systematic study of racism in sociology. In this work, Shiao & Woody (2021) make note that the the “decision to connect concepts across the divide between comparative ethnicity and critical race theory may be controversial, but [they] find the general terminology of comparativists to be helpful for differentiating among uses of racism as structure and integrating the historical and group-specific analyses of U.S. critical race scholars” (p. ??). With these divides in mind, the two frameworks of racism presented by Bonilla-Silva (1997) and Shiao & Woody (2021) were used to guide the conceptual replication and data analysis for the study.

### A Complexity, Specificity, and Variation Framework

We extend a *complexity, specificity, and variation* (CSV) framework to inform the development of the study data sets and to conduct our analysis of selected data from studies of racism in STEM. This framework allows us to systematically dissect and understand the multifaceted nature of racism research, offering a structured approach to examine the intricate details embedded within the data. By leveraging the CSV framework, we aim to provide a comprehensive and nuanced analysis that captures the depth and breadth of scholarly work in this critical area. This framework enhances the analysis of racism studies by providing a structured and multifaceted approach that captures the complexity, specificity, and variation inherent in this field of research.

The CSV framework is situated primarily around the intersection of perspective and methods. Other frameworks that have been advanced in studies of racism through a structural approach, such as the Narrative Policy Framework (NPF; (Bertrand et al., 2024; Dool & Schlaufer, 2024; Stauffer et al., 2024)). Broadly, NPF offers a methodical way to study how policy actors use storytelling to influence debates about race and racism, allowing researchers to identify key structural elements of policy stories, such as settings, characters, plots, and morals. NPF also helps researchers analyze the strategies used by policy actors to influence public opinion on race-related issues by operating through analyses at “micro, meso, and macro” levels, providing for a more comprehensive examination of how narratives about race function at individual (micro), group (meso), and societal (macro) levels. NPF can help trace how narratives about race influence policy decisions and public support for racial equity initiatives. While there are multiple connections.

### Complexity

Complexity refers to the constant development and evolving intricacies of social science research. In the context of racism studies, complexity encompasses the multifarious dimensions and layers that characterize the phenomena under investigation. This includes the historical evolution of racism, the diverse theoretical perspectives employed by researchers, and the intricate interplay of social, economic, and political factors that shape racial dynamics. By acknowledging and addressing this complexity, our analysis aims to unravel the sophisticated web of factors that contribute to the understanding of racism.

The CSV framework addresses the evolving and intricate nature of racism studies by acknowledging the multifaceted dimensions of racism. This includes understanding how racism operates at various levels, from individual to structural, and how it intersects with other social determinants. For example, structural racism is a complex construct that operates through multiple domains and institutions, shaping morbidity and mortality across racial and ethnic groups over time. By incorporating complexity, the CSV framework allows researchers to unravel the sophisticated web of factors that contribute to racial disparities.

### Specificity

Specificity within the CSV framework builds on the detailed work of scholars who frame studies within both historical and contemporary contexts. This focus allows for precise and meaningful conclusions about the state of racism research. For instance, Critical Race Theory (CRT) emphasizes the importance of presenting stories about discrimination from the viewpoint of people of color and addressing the intersections of race, sex, and class. By honing in on specific aspects of racism, such as particular populations, methodological approaches, and outcomes, the CSV framework ensures that the analysis is contextually grounded and detailed.

Our use of specificity builds on the work of scholars in sociology who frame studies in a discipline across both the historical. This scholarship aims to help develop our analysis of the *notions* of racism attributed not only in a discipline but across the various sub-disciplinary areas of study in sociology, as noted by (K. Ojeh, 2024) and others (Besek et al., 2021; Clair & Denis, 2015; Wright, 2020; Zuberi, 2004). In this way, specificity involves honing in on particular aspects of racism research, such as the methodological approaches used, and the particular perspectives that are undertaken and conceptually developed. This focus allows us to draw notions and meaningful conclusions about the state of racism research, identifying trends, gaps, and areas for further inquiry. By integrating specificity into our framework, we ensure that our analysis is detailed and contextually grounded.

### Variation

Our analysis of variation is rooted in traditional methods employed in studies of bibliometric records (Airyalat et al., 2019), enriched with insights developed from the complexity and specificity arms of the study. We develop these insights from the complexity and specificity arms of the study, and use prior research on the systematic analysis of research in STEM fields (Gil-Doménech et al., 2020). Variation examines the differences and similarities across studies, such as variations in research findings, methodological approaches, and theoretical frameworks. This aspect of the CSV framework helps us to understand the diversity of perspectives and approaches within the field of racism research, highlighting the dynamic and evolving nature of scholarly inquiry. By analyzing variation, we can identify patterns and divergences that inform a more holistic understanding of the research landscape.

Interestingly enough, the CSV framework situates the ways that large-scale data, often formatted in comma-separated values (CSV), results in a dynamic analysis of the content of studies across disciplines and within a particular domain. This technological aspect of the framework underscores the importance of data management and analysis tools in contemporary research. By leveraging large-scale data, we can conduct robust and comprehensive analyses that transcend disciplinary boundaries and provide deeper insights into the study of racism. This integration of technology and methodology exemplifies the innovative potential of the CSV framework in advancing our understanding of complex social phenomena.

# Data and Methods

In this article, we explore the insights of large-scale computational critiques in fostering an understanding of the intellectual structure of racism in STEM education. By taking on a historical lens, we examine how racism might be explored differently based on the varied development of a domain. More plainly, there may be a noticeable difference regarding scholars in different domains and disciplines and the uptake of various social issues, especially withing those STEM disciplines of study deemed as the “hard” sciences (e.g., biology, chemistry, physics). In making sense of the diversity of this scholarship, we examine the contributions of various related disciplines and the intellectual structure and keywords used by sets of highly cited disciplinary scholars within a domain.

This section describes the procedures undertaken in the conceptual replication of Shiao and Woody’s (2021) study, The Meaning of ‘Racism’, which further examines three broad constructs that frame sociologists’ use of the term ``racism’’ that they extend in their analysis of the literature in sociology. We use a quantitative historical analytic lens to examine the framework’s mapping on to racism-related research across the various science, technology, engineering, and mathematics (STEM) sub-disciplines with a focus on the global distribution, disciplines and domains, and variations in the key framings used across these disciplines. We map the intellectual structure of racism-focused research in STEM and STEM education journal articles and model similarities across database samples. Results indicate that research on racism in the STEM disciplines varies by disciplinary network but not necessarily by the associated content. Despite the charge to examine more systemic forms of racism, the samples are primarily inclusive of research on two of the three constructs: individual attitudes and cultural schema. Three distinct citation-based communities were identified across samples. There is a change in the variance of the network citation metrics when themes are varied. We discuss the need for more empirical studies on the structural conceptions of racism and present a model of a subgroup of highly-networked authors.

## Methods

This study is motivated, in part, by the work of (Shiao & Woody, 2021) who discuss the various meanings of racism in sociological journal content. For the purpose of this study, our team sought to make sense of the nature of this intellectual structure in STEM and the various concepts and themes situated within it. We integrate systematic review techniques with computational methods to generate the methods for the study. The systematic review process builds on qualitative meta-synthesis techniques, as well as content and thematic analysis. The computational methods follow the general scientific mapping workflow for working with large sets of citation data.

Systematic review techniques differ from traditional literature reviews in several key ways, offering a more rigorous and structured approach to synthesizing research evidence. Unlike traditional reviews, systematic reviews begin with a specific, focused research question and a pre-defined protocol outlining the methodology. They employ a comprehensive, detailed search strategy across multiple databases and sources to find all relevant studies, in contrast to the less rigorous search often used in traditional reviews. Systematic reviews also utilize specific, pre-defined inclusion and exclusion criteria for studies, whereas traditional reviews may rely on subjective or unclear selection criteria. A critical aspect of systematic reviews is the rigorous assessment of study quality and risk of bias, which may be limited or variable in traditional reviews. In terms of synthesis and analysis, systematic reviews often employ quantitative methods (e.g., meta-analysis) or structured qualitative synthesis, while traditional reviews typically provide a narrative summary of findings.

Reproducibility is a key feature of systematic reviews, with detailed documentation of methods allowing for replication, unlike traditional reviews which often contain subjective elements and lack methodological transparency. Systematic reviews are designed to minimize bias through rigorous methods, including dual independent screening and data extraction, making them less prone to various biases compared to traditional reviews. Additionally, systematic reviews can be periodically updated to incorporate new evidence, whereas traditional reviews are typically one-time efforts. These differences make systematic reviews more reliable for answering specific research questions, particularly in healthcare and policy-making contexts. However, traditional literature reviews, while less structured, can still be valuable for providing broad overviews of topics or contextual information. The choice between these two approaches depends on the research goals, available resources, and the level of evidence synthesis required for the particular study or field of inquiry.

Despite the value of systematic analysis, there are few replication studies conducted in certain disciplines, one being research in education (Perry et al., 2022). In studies on replication methods, the focus on the value of certain methodological insights is presented and challenged. For example, Bonett (2021) focus on the various criteria needed in replication studies, confirming the research conducted by Anderson & Maxwell (2016) and Perry et al. (2022) that direct replications are rare. While focusing on the various quantitative and computational facets of replication studies, Bonett (2021)’s analysis provides insights into the dynamic nature of replication studies and the increasing requirements of computational methods for replication. Perry et al. (2022) review replication studies in education between 2011 and 2020. In their analysis, they examine the number of replications and types of replication studies conducted. The authors note that many studies in education research are conceptual in nature, and that direct replication studies are rare in the discipline of education. Each of these studies contributes to our broader engagement with large-scale bibliometric content in education and the development of a more diverse approach to examining the evolution and development of ideas in a field.

### Conceptual framework for the replication

We outlined questions early in our discussion regarding engagement with the conceptual and methodology tasks associated with making sense of racism within and across disciplines: Through what lenses might scholars navigate the expanding research literature and intersecting areas of study that cross disciplinary boundaries? How can research scholars address the complex dynamics of increasingly niche disciplinary terminologies in light of new computational tools? And how can individual researchers and teams make sense of intersections in terminology that crosses several academic disciplines?

While aspects of this study provide insight to these questions, we see each question as an ongoing inquiry that should be used to develop visions for a set of priorities and clinical practices (e.g., procedural considerations) that take into account disciplinary priorities and increasingly niche research content. In the review of research on racism in STEM and across other disciplinary areas, like STS, we identified of racism in published academic research through the year 2023. We used a sample to study associations, and we explored variation in other samples. In the study samples, there were differences in the number of results returned and multiple lenses through which scholars situated racism.

Scholars’ framings were representative of a set of key terms that have been increasingly studied in recent decades. Some primary frameworks used include a host of critixal theories, and specific mentions of intersectionality, critical race theory, and postcolonial perspectives. Applying these frameworks has allowed scholars to understand the varying ways that racism interacts with other forms of oppression, such as sexism, ableism, and classism; and further analyze how racism operates within structures and the standard, everyday practices in STEM settings. These studies encourage the analysis of power dynamics, institutional policies, and cultural norms that contribute to the reproduction of racial attitudes, behaviors, and systems of inequality. From a methodological perspective, researchers noted that these frameworks present opportunities for others to continue to expand the study of the historical impact of colonialism and imperialism, racial injustice, and the construction of racial systems on scientific knowledge production. Researchers also describe the importance of identifying communities and examples that represent equitable scientific practice in the development of more inclusive societies.

Frameworks on the conceptualization and meaning of racism

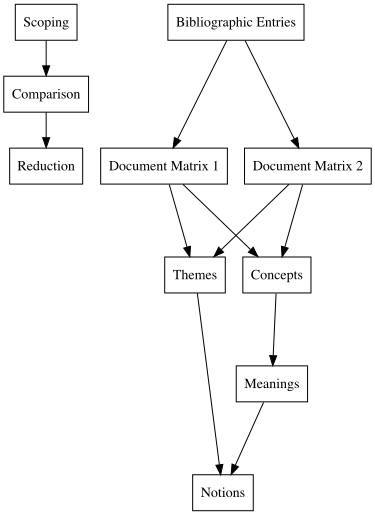
| **Bonilla-Silva (1997)** | **Shiao and Woody (2021)** |
| --- | --- |
| 1. Psychological phenomena | 1. Individual attitudes |
| 1. Cultural processes | 1. Cultural schema |
| 1. Social structure | 1. Structure |
|  | (3) Structure: Pre-existing conditions |
|  | (3) Structure: Create or maintain inequity |
|  | (3) Structure: Cultural mobilization |
|  | (3) Structure: Racial dominance |

### Analytic framework for the replication

The analytic framework for this study was an iterative process that included elements from cliometric analytic methods (Aria & Cuccurullo, 2017) and qualitative meta synthesis (Thunder & Berry, 2016). Given that the research on racism in STEM draws from a more diverse set of intersecting disciplinary sources, our study departs from the methods outlined in (2021), and adopts some of the methods engaged in the systematic processes in related studies in education on critical content. In our study, we focus on the distribution of research across domains and our process on extracting information from bibliographic entries and citations to examine the . As we sought to identify the meaning of racism in STEM, we understood that a diverse collection of disciplines and foundations can be found in STEM (e.g., physics, psychology, zoology). We notice that scholarship at this intersection is parsed across multiple disciplinary boundaries and domains of study described by category and research area. As a result, we develop a conceptual replication of ’s (2021) study to examine how various of racism in STEM relate to the major constructions of racism identified in the discipline of sociology. While study results present a summary of the research literature on racism and STEM, here we build an analytic process around the conceptual framework through which we identify studies on racism and STEM.

Bibliometric analysis and quantitative historical methods share common ground in their utilization of quantitative data to derive insights and draw conclusions. Both approaches employ systematic methodologies that examine information and uncover patterns. The two frameworks diverge in their primary objectives and by the nature of their respective disciplinary foundations. As noted by (2021), these analytic choices hold benefits and limitations. Traditional bibliometric analysis primarily focuses on the quantitative assessment of scholarly publications, including citation patterns, authorship networks, and journal impact factors; it aims to evaluate research productivity, influence, and trends within specific disciplines or research communities. Quantitative historical methods aim to uncover and analyze patterns in historical data, in this case publications at the intersection of racism and STEM.

By employing integrated techniques in our analysis, the quantitative historical component allowed us to examine any historical phenomena and identify related concepts in the study data. Our use of bibliometric analytic techniques center on the study of scholarly texts and their diverse , so that we may delve further into the exploration of historical phenomena using quantitative measures. While further research is called for in the general area of database research (, 2019; , 2006), there is an ongoing need to consider extensions to tradition in bibliometric analyses. These new considerations would look for various limitations in the method and provide grounding contexts that can help frame scholars’ decisions with the use of certain terms and lexical patterns that prescribe meaning.



**Figure 1**. Analytic framework for the study

### Limitations

In their discussion, Turchin et al. (2017) also comment that “others have criticized [the quantitative] approach on the grounds that [the] proposed measures focus too narrowly on size and hierarchy or that there are multiple dimensions of variable manifestations of complexity” (p. E145).

## Data

### Scoping and comparison

An iterative scoping process was followed to make sense of the various databases and document structures. The comparisons that followed were used to differentiate the contents of each potential data set. Four databases were used to identify the possible collections of sources, and to compare the various results obtained. Using multiple databases provides a more complete comparative picture of the available literature, as each database has different coverage and indexing practices. This approach also presents an opportunity to examine differences across variations in terminology.

*EBSCO*. As one of the leading providers of research materials, EBSCO (Elton B. Stephens Company) initially sold magazines and military supplies but has now become one o the largest providers of research databases for libraries and academic institutions. For this study, a subset of databases were selected that followed the work of other scholars (Thunder & Berry, 2016). The databases were: Academic Search Premier, Social Sciences Full Text (H. W. Wilson), Education Research Complete, ERIC, and APA PsychArticles and APA PsychInfo Collection). For example, ERIC (Education Resources Information Center) is a specialized database for education research, providing access to over 1.6 million texts related to all aspects of education. It is sponsored by the U.S. Department of Education and is considered the largest database for education literature.

*Scopus.* Scopus is one of the leading abstract and citation database that covers multiple academic fields and is considered as a comprehensive selection for academic researchers. It covers scientific journals, books, conference proceedings across various areas of study. It provides journal metrics and advanced search capabilities that can be used to identify specific areas of study. Scopus is also integrated with other research tools and software to incorporate large-scale analysis.

*Google Scholar*. As a part of Google’s broader universe, Google Scholar is one of the most widely used free academic search engines. The search engine indexes a range of scholarly literature across multiple disciplines and does not include a peer review process for inclusion, unlike the other search engines. As a result, it is considered one of the broader academic search engines that contains one of the largest collections of scholarly materials.

*Web of Science (WoS)* is a comprehensive and highly respected citation database used for scholarly research. It is one of the main databases used to identify and compare collections of citations and bibliographic sources. Given the conceptual replication method selected for this study, we continued by using the Web of Science Core Collection of cliometric data. This decision provided us with an opportunity to understand differences across disciplinary boundaries despite the increasingly interdisciplinary nature of the research on racism in STEM.

Table 2 provides results for the keywords across their related keyword patterns. In some engines, the use of racis\* produces results that begin with “racis” such as racist and racism.

Table: Results of title keyword search[[1]](#footnote-37)

| **Keyword** | **EBSCO** | **Google Scholar** | **Scopus** | **Web of Science** |
| --- | --- | --- | --- | --- |
| anti-racism | 381 | 3,420 | 685 | 592 |
| anti-racist | 427 | 3,620 | 826 | 685 |
| race | 44,057 | ~313,000 | 65,203 | 69,537 |
| racial | 24,467 | ~212,000 | 39,930 | 41,527 |
| racialization | 519 | 2,740 | 1,108 | 1,077 |
| racialized | 958 | ~5,600 | 1,863 | 2,120 |
| racism | 8,669 | ~99,500 | 11,418 | 10,674 |
| racist | 1,965 | ~15,100 | 2,170 | 1,906 |

**PRISMA Flow Diagram**

The protocol outlined in the PRISMA (**prisma2009?**) flow diagram was modified to examine the reduction of records across the two sets. This reduction method integrated a critical analysis of data loss and considered the various ways that contemporary methods may not include specific articles. Each of these steps is outlined in three section. The first of these sections focuses on the reduction methods used to generate the final two data sets for the study. Then a discussion of the inclusion criteria (IC) and exclusion criteria (EC) is discussed. The research categories and counts of articles in each of these categories across the two data sets is then outlined.

### Reduction

Data reduction took place in three steps. The first reduction of data occurred in the scoping process where articles were reviewed for containing complete cliometric records. In this step, JSON files of the citation records were entered into the Quanteda (**quanteda?**) R/Posit software package which summarized the sources of any incomplete records. The second reduction of data occurred in the comparison step where any results that were returned which did not fit the requirements outlined in the inclusion and exclusion criteria (see (**table?**)) following traditional cliometric analysis methods were excluded. The third and final reduction of data occurred in the comparison step where articles were reviewed for their fit within the goals of the study. In this step, if any articles were not found during the initial two steps and any ancillary results that were returned were excluded. This resulted in a set of cliometric records across the two initial samples for the study. A third sample was generated by integrating the two samples obtained from the analysis of records from STEM education journals and those records from STEM journals; the difference between these two cases is noted in the (**table?**) below.

### Inclusion and exclusion criteria

Inclusion and exclusion criteria for the study

| Criteria | Code | Criteria |
| --- | --- | --- |
| **Inclusion criteria** | IC1 | Article contains keyword pattern in either the title (TI), abstract (AB), or author keywords (DE) |
|  | IC2 | Article published before 2024 |
|  | IC3 | Article originally written in English |
|  | IC4 | Article is a peer-reviewed journal article |
|  | IC5 | Article problem, purpose, or core question(s) center on the topical subjects of the analysis as opposed to only mentioning author keywords (DE) |
| **Exclusion criteria** | EC1 | Article does constitute a complete citation record |
|  | EC2 | Article contained ancillary content that is not focused on a clear and related research question |

**Table 3**. : WoS results by research category for the study samples

| Research category | M1 | M2 |
| --- | --- | --- |
|  | n() | n() |
| Education/Educational Research |  |  |
| Education Scientific Discipline |  |  |
| Public Health |  |  |
| Ethnic Studies |  |  |
| Sociology |  |  |
| Women’s Studies |  |  |
| Social Science Other Topics |  |  |
| Government Law |  |  |
| Social Psychology |  |  |
| Environmental Sciences Ecology |  |  |
| Science Technology Other Topics |  |  |
| Educational Psychology |  |  |
| Interdisciplinary Social Sciences |  |  |
| Political Science |  |  |
| Health Care Sciences Services |  |  |
| Multidisciplinary Psychology |  |  |
| Multidisciplinary Sciences |  |  |
| Urban Studies |  |  |
| Chemistry Multidisciplinary |  |  |
| Engineering Multidisciplinary |  |  |
| Geography |  |  |
| History |  |  |
| Psychology Developmental |  |  |
| Other |  |  |
| Total |  |  |

# Findings

We identify the distribution of research on racism in STEM across academic disciplines to help navigate the expanding research literature and intersecting areas of study. We examined various database options to search and retrieve published scholarly studies. Among these databases, the Web of Science (WoS) was situated very early in our search as one of the most widely-expanding citation databases used by researchers. WoS is often compared to other well-known citation databases and networks such as the Google Scholar, Pubmed, and Scopus databases.

## Database summaries

Each database, , included a different number of observations but a total of 61 variables. A matching method was used to identify and remove all duplicate records. In our coding schema for the reduction, when duplicate records were noted, a new data frame was generated and used as the main object.

Publication count and rank by country

| Country | M1 |  |  | M2 |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | *Rank* | *Count* | *% of total*  (n = 474) | *Rank* | *Count* | *% of total*  (n = ???) |
| USA | 1 | 352 |  |  |  |  |
| England | 2 | 35 |  |  |  |  |
| Canada | 3 | 31 |  |  |  |  |
| Australia | 4 | 19 |  |  |  |  |
| Brazil | 5 | 7 |  |  |  |  |
| Germany | 6 | 6 |  |  |  |  |
| Israel | 7 | 5 |  |  |  |  |
| Ireland | 8 | 4 |  |  |  |  |
| Spain |  | 4 |  |  |  |  |
| India | 9 | 3 |  |  |  |  |
| Mexico |  | 3 |  |  |  |  |
| Norway |  | 3 |  |  |  |  |
| China |  | 3 |  |  |  |  |
| South Africa |  | 3 |  |  |  |  |

### Document Matrix 1

Document matrix 1 includes the records from education journals.

### Document Matrix 2

Document matrix 2 includes those records from STEM and non-education journals.

For example, in the STEM education research indices, Kendi (2016) is an outlier in the co-citation network. However, in the STEM indices, the same author is a central node. This difference points to the varied intellectual histories within the index work. Still, one limitation of this conclusion is that the disciplinary tags (indices) and outlets (journals) are not random but include some amount of random variation.

# Results and Analysis

Descriptive results for each database is provided in Table 3. M is the merged data from M and M.

**Table 3**. Database record counts and labels.

| **Database** | **Complete Records** | **Columns** | **Label** |
| --- | --- | --- | --- |
| M | 278 | 61 | edrecords |
| M | 132 | 61 | stemrecords |
| M | 351 | 61 | combined |

# Discussion

Studies of racism in contemporary sociology offer a host of frameworks to describe the potential of racism in the sociology of race and ethnicity. As research on racism in science, technology, engineering, and mathematics (STEM) continues to gain traction, the dynamic developments and interdisciplinary perspectives leveraged by scholars will require the use of dynamic frameworks that incorporate diverse voices. In large scale studies of citation networks, dynamic representations of research literature help map the increase of cross-disciplinary research and intersecting themes (Cite).

These systematic studies and the analysis of corpora of bibliometric data offer research scientists ways to understand the study of the various phenomena. However, the diverse topics across what is conceived as `STEM’, while useful for framing a set of highly related disciplines, fails to provide a common framework for the analysis of scientific literature. As research on racism expands, the framing used around various concepts in relation to STEM will offer important insights. These insights can be understood through an analysis of the various notations scholars have situated around the study of racism.

Some examples and widely cited studies that focus on racism and STEM include research on the experiences and perceptions of groups racialized as non-white (, 2019; , 2007; , 2018; , 2019; , 2013; , 2011; , 2012; , 2021; , 2017); studies on the various cultural schemas, conceptions of race and racialization, and intersecting identities (, 2019; , 2020; , 2022; , 2023; , 2016; , 2017; , 2017; , 2020; , 2020); and studies that analyze various structural dimensions, often examining interlocking systems and power relations (, 2017; , 2020; , 2022; , 2019; , 2018). Importantly, the discussions across this extensive body of work often reside in multiple thematic areas; and many studies on racism and STEM provide insight into the ideologies, habits and traditions, social and cultural processes, and core structural components of racialization. These studies also highlight the dynamic and context-dependent nature of racism across social contexts, physical settings, and complex ecosystems (, 2018). Yet, questions remain about how scholars situate ``racism” across different disciplines of study (, 2021; , 2021).

Additionally, as research on racism in STEM extends globally, questions about geographical variation and citation network patterns arise. These patterns are further influenced by discipline-specific cultures and the array of social politics around citation practices in academic research (, 2018; , 2017). In light of these considerations, a set of key questions emerged for our research team: Through what lenses might scholars navigate the expanding research literature and intersecting areas of study that cross disciplinary boundaries? How can research scholars address the complex dynamics of increasingly niche disciplinary terminologies in light of new computational tools? And how can individual researchers and teams make sense of intersections in terminology that crosses several academic disciplines? Through these questions, our analysis of racism in STEM has implications for both theory and practice. Moreover, this analysis contributes to our knowledge about the various features that connect studies of racism in research regarding STEM.

For the purposes of this paper, we view notions as a web of concepts that relate to distinct and overlapping sets of ideas or themes. Notions around a broad topic, like racism, generate a web of concepts and terms. This web produces an interconnected and complex structure of words which imply that notions are not isolated ideas but are interrelated and interdependent. A thematic network map, as opposed to a word cloud, would be one example. In this paper, we model different approaches to generate latent constructs (or themes) as they relate to various measures of association between the metadata of STEM and STEM education journal articles which focus, in some way, on racism-related matters. We discuss the various selection and inclusion criteria, and present models on different associations. Findings reveal a set of ``steady-state” results across all models which structure the set of core notions of racism.

Research by (2021) identified the various ways that scholars structure their discussions of racism, a notable difference between, for example, individual attitudes and broader systems. As studies of racism expand in a field, these scholars have identified how the diverse interpretations of racism" call for examining the various \textit{meanings} ofracism” in a body of research literature (, 2021; , 2021). We examine notions of racism in STEM using a context-based bibliometric approach.

The focus on CRT presented some of the broader discourses developing in the United States which have informed developing studies across nations, contributing to the parallel evolutions noted by Turchin et al. (2017). (Lloyd et al., 2021) note that “this inequality is built into the infrastructure of our country and has formed the foundation for structural racism—a system that privileges White people and results in intentional disadvantage for Black Americans. These inequalities negatively impact the lives of Black people in a number of ways, including where they live; the education they receive; their employment and economic opportunities, access to child care, mental and physical health outcomes, and political standing and power; and the way they are treated in our systems of law and justice” (p. 2).

As scholars better understand the meanings prescribed to racism and the presence of moving syntactical signifiers, there is also a need to understand the history and foundations of popular policies and the use of specific terms, like STEM (, 2020). We examine notions of racism as framed in a body of research on `STEM," and in mathematics, science and technology studies to make sense of how racism is framed and contextualized. For instance, \cite{benjamin2016catching} (2016) uses the termSTS’, instead of STEM, to refer to the specific area of science and technology studies (STS) (, 2018). However, despite focusing on two of the integral dimensions of STEM – science and technology – many traditional approaches in bibliometrics and citation analysis may overlook this contribution due to differences in terminology, which is a common challenge in cross-disciplinary research (CDR, , 2015; , 2020).

Notions, then, are a function of the lexical patterns identified in a set of keywords-in-context. With increases in latent semantic analysis (LSA) and other methods, the development of broad themes in a body of work provides content for further inquiry, and not full results. When themes and primary concepts are situated as the meanings, care must be taken with the broader social and political context of the disciplinary boundaries. For example, Shiao & Woody respond to the development of the journal and a call by Bonilla-Silva (1997). In U.S. mathematics education, the social turn (Lerman, 2000), sociopolitical turn (Gutierrez, 2007), and spatial turn (Cite, XXXX) prompted the inclusion of different methods for the analysis of data. We observe similar patterns across other fields of study. As a result, notions provide a sampling of the ideas that exist across samples with the understanding that different disciplines both uptake and examine developments in the disciplinary area largely from disciplinary perspectives. As interdisciplinary research expands, additional models and frameworks will be required.

# Conclusion

Our study, employing an integrated systematic approach through the complexity, specificity, and variation (CSV) framework, has yielded significant insights into the landscape of racism research in STEM. The application of this framework to our analysis of two distinct datasets has revealed both the intricacies and nuances inherent in this critical field of study. The results of our investigation demonstrate a high degree of complexity within the research on racism in STEM. This complexity is evident in the multifaceted nature of the studies examined, reflecting the intricate interplay of social, institutional, and individual factors that contribute to racial disparities in STEM fields. Our analysis uncovered a rich tapestry of research approaches, theoretical frameworks, and methodological strategies employed by scholars in addressing this pressing issue.

he CSV framework allowed us to identify and analyze specific themes, populations, and contexts within racism research in STEM. This specificity manifested in the detailed examination of particular racial groups, specific STEM disciplines, and various educational and professional settings. Such granular analysis provides a more nuanced understanding of how racism operates and impacts individuals within the STEM ecosystem.

Interestingly, while our study found substantial complexity and specificity across both datasets, the variation in content between the two was not as significant as initially anticipated. Our review of keywords in context revealed a considerable overlap in the core themes, concepts, and approaches used in racism research across different STEM disciplines. This finding suggests a certain level of consistency in how racism is studied and conceptualized within the broader STEM community, despite the diverse nature of individual STEM fields.

The conceptual replication (Anderson & Maxwell, 2016; Bonett, 2021; Perry et al., 2022) of Shiao & Woody (2021)’s study on the various meanings of racism described by scholars in the sociology of race and ethnicity employed a different set of analytic methods that provided a descriptive framework for a critical data scientific method following in the contemporary work of scholars in sociology, such as outlined in Ojeh (2024), in the conceptions of needed specificity within a discipline. Our research explorations included peer-reviewed scholarships across the related but distinct research published in journals focused on STEM education and those published in STEM journals.

The consistency in content across datasets, coupled with the observed complexity and specificity, underscores the importance of collaborative and interdisciplinary approaches in addressing racism in STEM. It highlights the potential for cross-pollination of ideas and methodologies among different STEM disciplines in tackling this pervasive issue. Our application of the CSV framework provides a comprehensive and structured analysis of research on racism in STEM, revealing both the depth and breadth of scholarly work in this area. The findings of complexity and specificity, alongside the consistency in content variation, offer valuable insights for future research directions and policy interventions. As we move forward, these results can inform more targeted and effective strategies to combat racism and promote equity in STEM fields, ultimately contributing to a more inclusive and diverse scientific community who can continue to learn within and across disciplinary boundaries.

# References

Airyalat, S. A. S., Malkawi, L. W., & Momani, S. M. (2019). Comparing bibliometric analysis using PubMed, scopus, and web of science databases. *JoVE (Journal of Visualized Experiments)*, *152*, e58494.

Anderson, S. F., & Maxwell, S. E. (2016). There’s more than one way to conduct a replication study: Beyond statistical significance. *Psychological Methods*, *21*(1), 1.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975.

Bertrand, A. R., Lyon, M. A., & Jacobsen, R. (2024). Narrative spillover: A narrative policy framework analysis of critical race theory discourse at multiple levels. *Policy Studies Journal*, *52*(2), 391–423.

Besek, J. F., Greiner, P. T., & Clark, B. (2021). W.e.b. Du bois and interdisciplinarity: A comprehensive picture of the scholar’s approach to natural science. *Journal of Classical Sociology*, *21*(2), 144–164.

Bonett, D. G. (2021). Design and analysis of replication studies. *Organizational Research Methods*, *24*(3), 513–529.

Bonilla-Silva, E. (1997). Rethinking racism: Toward a structural interpretation. *American Sociological Review*, 465–480.

Clair, M., & Denis, J. S. (2015). Sociology of racism. *The International Encyclopedia of the Social and Behavioral Sciences*, *19*(2015), 857–863.

Crisp, G., Alcázar, L., Sherman, J. R., Schaffer-Enomoto, J., & Rooney, N. (2024). Systematic review of theoretical perspectives guiding the study of race and racism in higher education journals. *Innovative Higher Education*, *49*(2), 247–269.

Dool, A. van den, & Schlaufer, C. (2024). Policy process theories in autocracies: Key observations, explanatory power, and research priorities. *Review of Policy Research*.

Du Bois, W. (2000). Sociology hesitant. *Boundary*, *27*(3), 37–44.

Furet, F. (1971). Quantitative history. *Daedalus*, *100*(1), 151–167.

Gil-Doménech, D., Berbegal-Mirabent, J., & Merigó, J. M. (2020). STEM education: A bibliometric overview. *Modelling and Simulation in Management Sciences: Proceedings of the International Conference on Modelling and Simulation in Management Sciences (MS-18)*, 193–205.

Go, J. (2020). Race, empire, and epistemic exclusion: Or the structures of sociological thought. *Sociological Theory*, *38*(2), 79–100.

Gordon, H. S. J., & Around Him, D. (2024). Understanding racial equity in research with indigenous peoples: Including anti-racism and decolonization approaches. *Identities*, 1–20.

Green, D. S., & Wortham, R. A. (2015). Sociology hesitant: The continuing neglect of WEB du bois. *Sociological Spectrum*, *35*(6), 518–533.

Haskins, L., & Jeffrey, K. (2011). *Understanding quantitative history*. Wipf; Stock Publishers.

Hughey, M. W. (2023). Du boisian sociology and intersectionality. In *Research handbook on intersectionality* (pp. 51–68). Edward Elgar Publishing.

Lloyd, C. M., Alvira-Hammond, M., Carlson, J., & Logan, D. (2021). Family, economic, and geographic characteristics of black families with children. *Racial e Quity: Child t Rends*, 4–5.

Monroe-White, T., & Lecy, J. (2023). The wells-du bois protocol for machine learning bias: Building critical quantitative foundations for third sector scholarship. *Voluntas*, *34*, 170–184.

Monteiro, A. (2000). Being an african in the world: The du boisian epistemology. *The Annals of the American Academy of Political and Social Science*, *568*(1), 220–234.

Neville, H. A., Monette, M., Lewis, J. T., & Safir, S. (2024). Shifting the gaze from racism to healing from racism: A systematic review of selected psychology journals from 1992 to 2022. *American Psychologist*.

Nzinga, S. M. (2020). *Lean semesters: How higher education reproduces inequity*. JHU Press.

Ojeh, K. (2024). Black sociology: Toward a theoretical analysis of systems of oppression and social power. *The American Sociologist*, *55*(1), 38–58.

Ojeh, K. S., & Wright, E. (2024). *An introduction to WEB du bois*. Taylor & Francis.

Perry, T., Morris, R., & Lea, R. (2022). A decade of replication study in education? A mapping review (2011–2020). *Educational Research and Evaluation*, *27*(1-2), 12–34.

Shiao, J., & Woody, A. (2021). The meaning of “racism.” *Sociological Perspectives*, *64*(4), 495–517.

Shuford, J. (2017). Four du boisian contributions to critical race theory. In *WEB du bois* (pp. 193–230). Routledge.

Stauffer, B., Kuenzler, J., & Jones, M. D. (2024). Narrative policy framework and policy implementation. *Handbook of Public Policy Implementation*, 193–203.

Thunder, K., & Berry, R. Q. (2016). Research commentary: The promise of qualitative metasynthesis for mathematics education. *Journal for Research in Mathematics Education*, *47*(4), 318–337.

Turchin, P., Currie, T. E., Whitehouse, H., FranÃ§ois, P., Feeney, K., Mullins, D., Hoyer, D., Collins, C., Grohmann, S., Savage, P., Mendel-Gleason, G., Turner, E., Dupeyron, A., Cioni, E., Reddish, J., Levine, J., Jordan, G., Brandl, E., Williams, A., … Spencer, C. (2017). Quantitative historical analysis uncovers a single dimension of complexity that structures global variation in human social organization. *Proceedings of the National Academy of Sciences*, *115*(2). <https://doi.org/10.1073/pnas.1708800115>

Wright, E. (2020). *Jim crow sociology: The black and southern roots of american sociology*. University of Cincinnati Press.

Yu, M., Coloma, R. S., Sun, W., & Kwon, J. (2024). Dissecting anti-asian racism through a historical and transnational AsianCrit lens. *Sociological Inquiry*, *94*(2), 330–350.

Zuberi, T. (2004). W.e.b. Du bois’s sociology: The philadelphia negro and social science. *The Annals of the American Academy of Political and Social Science*, *595*, 146–156.

1. Reported results obtained on July 16, 2024; search for titles containing keyword patterns. [↑](#footnote-ref-37)