Critical Control: How Different Forms of Vocabulary Control Aid and Hinder Novice Indexers Aiming to Support Racial and Social Justice

Chris Holstrom

University of Washington, Department of English and The Information School,
A101 Padelford Hall, Box 354330, Seattle, WA 98185-4330, USA
cholstro@uw.edu

Chris Holstrom is an assistant teaching professor in the Department of English at the University of Washington. He holds a PhD in Information Science and an MS in Technical Communication – both from the University of Washington. His research interests include subject indexing, indexing languages, and social tagging. He focuses on understanding the human role in these technologies and processes, and on critically analyzing their social and cultural impact. Chris first became interested in knowledge organization when he designed and built tagging systems while working as a technical writer at Google. As an academic, Chris expanded his research interests to include more formal indexing languages and developed a particular affection for the *Sears List of Subject Headings*, which he makes a point of teaching in his cataloging and knowledge organization courses. In addition to his work in knowledge organization, Chris teaches courses and conducts research in technical communication, rhetoric, and data science.

Holstrom, Chris. 2024. "Critical Control: How Different Forms of Vocabulary Control Aid and Hinder Novice Indexers Aiming to Support Racial and Social Justice". *Knowledge Organization* 51, no. 7: 521-542. 56 references. DOI:10.5771/0943-7444-2024-7-521.

Abstract: Controlled vocabularies are widely used in Knowledge Organization Systems (KOS); however; they are criticized for perpetuating biases, being slow to change, and not reflecting the language of many groups and cultures. This paper examines and challenges these criticisms by identifying five forms of vocabulary control in use today (minimal control, passive control, post hoc control, flexible control, and rigid control), and studying their effect on the subject indexing and subject tagging processes from the perspective of novice indexers and taggers. The study uses a mixed methods approach, including a survey, a think-aloud protocol that was employed while participants indexed and tagged documents, and a retrospective interview, to better understand participants' actions, thoughts, and reactions during the indexing and tagging process. The study explores how controlled vocabularies make indexers and taggers feel controlled, in control, or out of control while they index and tag; measures how different forms of vocabulary control affect coextensiveness between aboutness statements and indexing terms; and analyzes how these experiences inform the criticisms of controlled vocabularies, particularly in relation to how vocabulary control aids and hinders racial and social justice in indexes. The study finds that the form of vocabulary control has a significant effect on the subject indexing and subject tagging processes, including an effect on participants' feelings of control during the processes, an effect on how participants transformed concepts during subject representation, and an effect on how participants navigate complex issues related to racial and social justice.

Received: 1st July 2024; Accepted: 11 July 2024.

Keywords: controlled vocabulary; subject indexing process; subject tagging process; racial and social justice.

1.0 Introduction

Controlled vocabularies – collections of authorized indexing terms and the rules that govern their use – minimize ambiguity and promote consistency (ISO 1985), improve precision and recall in information retrieval (Lancaster 1972; Soergel 1985, 123-126), and help to collocate related resources (Svenonius 2000 p. 21-22). However, controlled vocabularies have been criticized for reinforcing systemic biases (Olson 2001; Furner 2007; Fox 2016), can be slow to

adapt to changes in language and culture (Bowker and Star 1996; Tennis 2012), and often do not reflect the vocabulary choices of many information seekers (Howard and Knowlton 2018) and many groups and cultures (Duarte and Belarde-Lewis 2015, Littletree and Metoyer 2015). These criticisms of controlled vocabularies have been voiced for decades by indexers aiming to promote racial and social justice while working with rigid controlled vocabularies like the Library of Congress Subject Headings (LCSH) (Berman 1971). For example, the LCSH authorized subject heading "Indians of North America" has long been criticized for perpetuating colonialist attitudes and not representing Native American peoples as they want to be represented (Duarte and Belarde-Lewis 2015) and has also been criticized by authors whose works have been assigned the controversial heading (Bullard et al. 2022). The dramatic case of the "Illegal aliens" headings in the LCSH again raised criticisms around representation and social justice in controlled vocabularies (Baron and Gross 2021) and illustrated how controlled vocabularies can control the actions of indexers and can frustrate their efforts to promote racial and social justice (Holstrom 2022).

While these criticisms, especially criticisms of specific problematic terms and specific controlled vocabularies, are both important and widespread, critical cataloging scholarship offers somewhat limited insight into indexers' experiences of working with controlled vocabularies. Chu and O'Brien (1993), Bertrand and Cellier (1995), Šauperl (1999), and Joudrey (2005) all observed indexers at work; however, only Bertrand and Cellier and Šauperl studied how indexers work with controlled vocabularies, and they did not address the criticisms of controlled vocabularies related to racial and social justice. Baron and Gross (2021) detail librarians' experiences as they tried to work around and update the "Illegal aliens" headings. However, their account details a specific case, a specific controlled vocabulary, and specific problematic terms – and the case begins after library resources have already been assigned subject headings. What, then, is the whole experience of an indexer aiming to index a library resource? What is their experience working with a controlled vocabulary? How does that experience aid or hinder racial and social justice in a catalog or index? This study aimed to gain insight into these questions by observing indexers, specifically novice indexers committed to racial and social justice, as they work with different forms of vocabulary control. The study asked participants (n=18) to analyze and assign subject headings and tags to four books about race and racism while verbalizing their thought processes. Next, participants were interviewed about their experiences, with questions that probed: 1) how different forms of vocabulary control affected their processes and the subject headings and tags that they assigned to each book, 2) whether they felt in control, controlled, or out of control as they worked with different forms of vocabulary control, and 3) how the effect of vocabulary control on their subject headings and tags and on their feelings of control aided or hindered their efforts to promote racial and social justice through indexing and tagging.

While vocabulary control is often viewed as binary (controlled or uncontrolled), this study asked participants to use, compare, and discuss the racial and social justice implications of many forms of vocabulary control that are currently used today:

1) Minimal control: Also known as "free tagging" or an "uncontrolled" vocabulary. Allows complete freedom of expression but also provides indexers with no guidance (Mathes 2004; Kipp and Campbell 2006; Adler 2009).
2) Passive control: Associated with tagging. A user interface or other mechanism suggests tags but allows taggers to assign any tags that they like (Holstrom 2020).
3) Post hoc control: Involves active editorial control by another agent to manage synonyms, homographs, and typos after the indexer or tagger assigns initial indexing terms (Bullard 2018; Jackson et al. 2018).
4) Flexible control: Encourages indexers to add subject headings to the controlled vocabulary, when necessary, by following specified guidelines. Associated with the Sears List of Subject Headings.
5) Rigid control: Requires indexers to use only the approved subject headings in the controlled vocabulary. Associated with the LCSH in the United States.

Quantitative and qualitative analysis found that these different forms of vocabulary control affected the indexing terms that participants assigned and their experiences with the subject indexing and tagging processes. Participants felt most controlled by rigid, post hoc, and flexible vocabulary control, citing 1) indexing terms that they disagreed with, found insufficiently specific, considered inaccurate, or perceived as missing; 2) the outside influence on their indexing and tagging choices; 3) and the complexity and size of indexing languages. They also transformed their initial aboutness statements more frequently and more significantly when translating into rigid and flexible forms of vocabulary control. However, many participants also felt "out of control" when using passive and minimal control, which they felt provided too much autonomy and insufficient guidance.

These experiences of control related directly to participants' feelings about how different forms of vocabulary control aided or hindered their efforts to promote racial and social justice through subject indexing and tagging. For example, one participant expressed feeling controlled when they said, "Oh, man, with the LCSH you gotta use some terms that hurt your soul sometimes." Another participant criticized specific vocabulary that made them feel con-

trolled: "I feel like 'Blacks' (plural) has a pejorative connotation...that's just not how we really talk about groups. We'd say, 'White people' or 'Black people'...so it is outdated language." This participant used "White people" as a tag while working with minimal vocabulary control, explaining "because I can do that." However, when participants indexed and tagged documents that described lived experiences that did not match their own racial identity or documents with controversial views on race, they, in many cases, felt supported (not controlled) by the outside guidance provided by rigid and flexible control because they felt they lacked the expertise or authority to assign accurate and authentic indexing terms on their own.

This paper explores the balance of these effects of vocabulary control on racial and social justice in subject indexing and tagging, considers how the racial perspective of resources and the racial identity of indexers and taggers informs the subject indexing and tagging process, and makes recommendations for promoting racial and social justice through intentionally designed approaches to vocabulary control. These recommendations focus on the goals of not doing further harm with colonialist and racist language, using vocabulary that groups and cultures use for themselves, and promoting representation and visibility for underrepresented groups and cultures.

2.0 Literature Review

Vocabulary control is employed near the end of multi-step processes in which indexers assign subject headings or tags to resources. This section discusses different forms of vocabulary control in that larger context and explores the critical relationship between vocabulary control and racial and social justice. First, this section describes the steps in the subject indexing and subject tagging processes, emphasizing how transformations that occur during translation from subject analysis to subject representation aid or hinder social and racial justice. Next, this section explores criticisms of controlled vocabularies and criticisms of and alternatives to controlled vocabularies, like tagging and folksonomies. Third, this section presents a taxonomy of the five forms of vocabulary control investigated in this study: minimal, passive, post hoc, flexible, and rigid. Finally, this section explores the concept of control in the relationship between vocabulary control and indexers and taggers, and establishes two axes of control: 1) from in control to controlled, and 2) from in control to out of control.

2.1 Steps in the subject indexing and subject tagging processes

To understand the effect of vocabulary control on subject indexing and subject tagging, one must consider the steps involved in these processes and where vocabulary control fits in those steps. While the literature generally agrees that subject analysis is the first step in subject indexing, Knowledge Organization (KO) scholars have presented many different models for the subject indexing process. These models differ primarily in the number of steps that they list in the process. Frohman (1990), Farrow (1991), Smiraglia (1991), and Lancaster (2003 p. 9) all propose a two-step model, which Lancaster summarizes as: "1) Conceptual analysis, and 2) Translation." Miksa (1983), Langridge (1989), and Mai (1999 p. 277) posit a three-step process. Mai summarizes his version of the three-step process:

> The first step, the document analysis process, is the analysis of the document for its subject. The second step, the subject description process, is the formulation of an indexing phrase or a subject description. The third step, the subject analysis process, is the translation of the subject description into an indexing language.

Whether the process comprises two steps or three, vocabulary control is employed during the translation of an initial conceptual analysis into a formal subject description or subject representation.

This process of translation, and how different forms of vocabulary control affect that translation in ways that aid or hinder racial and social justice, are the focus of this study. Therefore, isolating the effect of vocabulary control and analyzing the transformations of subjects during the translation process is critical. To do so, this study intentionally separates its procedure into sequential and discrete steps of subject analysis and subject representation. This experimental design aligns with Langridge's arguments for keeping the subject analysis step independent (1989). Lancaster (2003 p. 26) concurs:

> The "conceptual analysis" stage of indexing should not be influenced by the characteristics of the vocabulary control to be used at the translation stage. That is, indexers must first decide what topics need to be represented; only later (momentarily perhaps) should they consider whether or not the vocabulary can represent these topics adequately. Put somewhat differently, indexers should not ignore a topic because they know or suspect that it cannot be expressed adequately.

While Langridge and Lancaster agree that subject analysis and subject representation should remain separate, Šauperl (1999) observes that experienced indexers regularly combine these steps, often because their knowledge of a controlled vocabulary allows them to anticipate the available vo-

cabulary for subject representation. Šauperl's finding is the primary reason that the current study uses novice indexers, who are less likely to combine subject analysis and subject representation, thus obscuring the effect of vocabulary control on translation.

2.2 Criticisms of and alternatives to controlled vocabularies

Because vocabulary control is so central to subject indexing, the KO community has often reflected on and critiqued controlled vocabularies, studying their value in modern KOS (Fidel 1992; Gross et al. 2015); contrasting controlled vocabularies with natural language search (Rowley 1994) and with uncontrolled vocabularies (Tennis 2006; Adler 2009); and criticizing controlled vocabularies for reinforcing systemic biases (Berman 1971; Olson 2001; Furner 2007; Fox 2016), for being too slow to adapt to change (Tennis 2012; Bowker and Star 2000; Berman and Gross 2017), and for not reflecting the language of information seekers and many groups and cultures (Howard and Knowlton 2018; Duarte and Belarde-Lewis 2015). These criticisms are closely intertwined with each other. For example, controlled vocabularies are often slow to change because of systemic biases and systemic biases are often revealed in the problematic or non-representative language in controlled vocabularies. The following subsections explore these criticisms in more detail, describe how leaving these criticisms unaddressed can cause continued harm to historically marginalized groups, and discuss the emergence of folksonomies as a potentially more democratic and responsive alternative to controlled vocabularies.

2.2.1 Controlled vocabularies reinforce systemic bias

A defining characteristic of controlled vocabularies is that they have authorized terms and unauthorized terms. Authorized terms are used to index documents, and unauthorized terms are not allowed by the controlled vocabulary. This system of authority allows for standardized vocabulary, which helps to promote key objectives of KOS such as collocation of resources. However, vocabulary control can also reflect the implicit or explicit biases of the editors of the controlled vocabulary, reinforce those biases, and mark them as authoritative for all the systems, people, and documents that the KOS touches. Sanford Berman (1971) explains why these biases are persistent and harmful, and notes that deeply embedded indexing languages like the LCSH are unlikely to be superseded despite catalogers' frustrations with their shortcomings. While the work of critical catalogers has improved many of the problematic headings that Berman identified in *Prejudices and Antipathies*, Berman notes in a more recent interview (2017) that systemic bias and colonial attitudes persist in the LCSH and that some of the headings that he critiqued decades earlier remain in place.

Other scholars note that persistent bias is not tied only to terminology but also to the structure of indexing languages (Olson 2001) and can even be exacerbated by attempts to remove bias (Furner 2007). Importantly, these persistent and systemic biases affect how information seekers find and perceive documents, and how they perceive themselves. For example, Howard and Knowlton (2018) show that researchers looking for information about Black Studies and LGBTQ+ studies (particularly scholars looking for information at the intersection of these two fields) struggle to find information that reflects their identities because of systemic bias in the indexing languages and tools that they use to find information, even though these tools are often presented and perceived as neutral. Melissa Adler (2017 p. 4, 27) analyzes a history of racism in classification schemes and calls for "taxonomic reparations" and for "multiple, local, community-based, and personal ways to organize knowledge and ideas."

2.2.2 Controlled vocabularies are slow to change

Changes to indexing languages (whether structural or terminological) typically require some type of warrant or evidence that a new term or a new usage is well enough established that the editors should update their indexing language. The expenses associated with updates are one reason that indexing languages tend to be conservative about change and updates tend to lag common usage (Tennis 2012, Bowker and Star 2000). This lag can be annoying when new technologies like personal computers and smartphones are slow to appear in controlled vocabularies, but the lag can be harmful when hard-won civil rights are not reflected, when people's identities cannot be accurately portrayed, and when offensive terminology persists in controlled vocabularies. This type of problem is evident in the "Indians of North America" and "Illegal aliens" headings in the LCSH. It is also evident in the Sears List's slow adoption of modern vocabulary for people of color, indigenous people, and members of the LGBTQ+ community (Holstrom 2021).

2.2.3 Controlled vocabularies do not reflect the language of many groups and cultures

Controlled vocabularies can cause harm to historically marginalized and underrepresented groups. One form that this harm can take is making information (especially information tied to one's self-identity) difficult to find because the controlled vocabulary does not use the language that people from minority and historically marginalized groups use to describe themselves (Howard and Knowlton 2018).

This misrepresentation can also cause symbolic and spiritual harm by echoing colonial misnaming and limiting the visibility and legitimacy of marginalized groups, including indigenous tribes (Duarte and Belarde-Lewis 2015).

One illustrative example of these criticisms and how they are intertwined is the previously discussed case of the "Illegal aliens" headings in the LCSH (Baron and Gross 2021, Lo 2019). This dated and prejudiced term appeared while an undocumented student was conducting research at the Dartmouth University Library, causing that student emotional harm. Despite intense efforts by students, librarians, university administrators, the American Library Association, and even United States congresspeople, it took years for the headings with "Illegal aliens" in them to be partially updated to a slightly less offensive phrasing. This case shows that controlled vocabularies can cause harm (see Adler and Tennis, 2013, for a taxonomy of harm in KOS) and that controlled vocabularies can control more than just indexing terms. Controlled vocabularies can also control the self-identities of information seekers and the ability of indexers to represent documents in a socially just and equitable way.

2.2.4 Folksonomies and tagging as an alternative to controlled vocabularies

To this point, this literature review has focused on controlled vocabularies associated with subject indexing, like the LCSH and the Sears List. This subsection introduces another type of KOS at the heart of this study: tagging. Tagging (or social tagging) is similar to subject indexing because tagging involves associating keywords or phrases with documents and because these tags, taken in aggregate can be used to form a KOS. Beyond these high-level similarities, however, lie key differences. First, tagging is typically part of a social internet experience and open to all web users, not just indexing professionals. Second, tagging is often performed by these individuals on the web for their own information organization needs and then aggregated with tags from other users (Feinberg 2006) to establish a consensus set of tags for a resource (Kipp and Campbell 2006; Halpin et al. 2007). The biggest difference, however, is the role that vocabulary control has historically played in social tagging: no role at all. In fact, tagging was often referred to as working with an "uncontrolled" vocabulary, and criticisms of tagging centered around how poorly tagging handled polysemy and synonymy (Golder and Huberman 2005), two issues that the uniform and unique headings in controlled vocabularies are designed to address.

Despite these criticisms, folksonomies were considered revolutionary alternatives to traditional KOS when they emerged during Web 2.0 because of three big advantages over traditional subject indexing: inclusiveness, rapid updates, and low cost. Tagging in early folksonomies was more inclusive than traditional indexing because taggers did not need to learn a controlled vocabulary and were free to use tags of their choice. As a result, early folksonomies were more likely to represent diverse user perspectives (Bates and Rowley 2011), to "directly reflect the vocabulary of users" (Mathes 2004), and to better "reflect the population's conceptual model of information" (Quintarelli 2005). Adler frames this difference in terms of power and control: "Perhaps the greatest power of folksonomies, especially when set against controlled vocabularies like the Library of Congress Subject Headings, lies in their capacity to empower user communities to name their own resources on their own terms" (Adler 2009 p. 309). Adler contrasts how the LibraryThing folksonomy represents transgender materials with how the LCSH does and finds that LibraryThing is able to adopt inclusive terminology more readily as language evolves, is able to better and more directly represent the voices of the transgender community, and can more exhaustively represent diverse voices. In other words, tagging addresses (or promises to address) some of the criticism of traditional controlled vocabularies, such as not reflecting the language or users, being slow to adapt, and perpetuating biases.

2.3 Five forms of vocabulary control

The previous section contrasted vocabulary control in subject indexing with the lack of vocabulary control in tagging and subject tagging. While this dichotomy is instructive, it is incomplete because tagging technologies and communities have evolved since early folksonomies. The following subsections expand upon the dichotomy by describing five forms of vocabulary control that are in use today and that are investigated in this study: 1) minimal control, 2) passive control, 3) post hoc control, 4) flexible control, and 5) rigid control.

2.3.1 Minimal control

Minimal control, or free tagging, entails tagging with an "uncontrolled" vocabulary and complete freedom of expression. The tagger experiences essentially no controls over what tags they can enter and receives no suggestions for tags. With minimal control, the vocabulary is an amorphous collection of tags that is constantly updated as each new tag is added, including any typos and synonyms. Minimal control is the cheapest form of vocabulary control because there is no centralized effort required to maintain the vocabulary. Of course, that low maintenance cost means that KOS that employ minimal control are likely to experience many of the problems that KO researchers note in early folksonomies (Mathes 2004; Quintarelli 2005; Kipp and Campbell 2006) and, while taggers are free to promote racial and social jus-

tice through any tags that they like, other taggers can use tags that actively undermine those goals.

While web moderation standards and anti-spam measures have made completely free tagging somewhat rare, many websites employ systems that can be classified as minimal control. For example, the community weblog MetaFilter has a free tagging user interface that asks blog posters to enter tags simply by entering space-separated words (e.g., "politics") or concatenated phrase (e.g., "uspolitics") to describe their blog post. Similarly, the "Fandoms," "Relationships," "Characters," and "Other Tags" fields on the fanfiction website Archive of our Own are examples of minimal vocabulary control because authors can assign any tags (and as many tags) as they want in these fields.

2.3.2 Passive control

Passive vocabulary control is similar to minimal control in that taggers are free to enter any tags that they like. However, the user interface for the tagging experience in passive control is different in an important way. With passive control, taggers see suggested tags from a controlled vocabulary that they can select (or not select) when they tag documents. Passive control requires slightly more effort and expense than minimal control to implement because the feature must be coded and enabled, an initial vocabulary must be established, and, in some cases, the vocabulary must be updated with new tags. Many popular social media sites, such as Instagram and Twitter, employ passive vocabulary control. For example, typing "university of washington" into an Instagram tagging field prompts suggested tags like "#universityofwashington," "#universityofwashingtonseattle," "#universityofwashingtonhuskies." Users can choose to use or ignore these algorithmically generated suggestions, but the ease of selecting a popular tag is considered a key feature and a defining characteristic of passive vocabulary control.

2.3.3 Post Hoc control

While minimal control and passive control function with little editorial intervention, post hoc vocabulary control involves active editorial control of tags to manage synonyms, homographs, and typos. Julia Bullard (2019) studies three websites that employ post hoc vocabulary control (Stack Overflow, LibraryThing, and Archive of Our Own) and terms them "curated folksonomies." Bullard (p. 9-10) lists the tenets of curated folksonomies as follows:

1) Users create tags
2) Some intentional agent combines synonymous tags and/or differentiates homographic tags
3) Recall and precision are improved

Jackson et al. (2018) adopt the metaphor of "gardening" from Peters and Weller (2008) to describe how they use this approach of harnessing user-generated tags with human judgment to build a structured folksonomy in a citizen science project.

Because human or algorithmic judgment is required for post hoc control, it is more labor-intensive and potentially more expensive than minimal or passive control. However, websites like the programming question-and-answer website Stack Overflow have encouraged qualified community members to do this editorial work voluntarily by providing incentives. For example, a trusted Stack Overflow member marked the user-provided tag "linux-development" as a synonym of the canonical tag "linux," which improves collocation, precision, and recall on Stack Overflow (n.d.). While humans currently perform most of this editorial work, algorithms are likely to supplant humans. Bullard (2019 p. 8) explores the role of algorithmic judgment by noting that "computational approaches [to post hoc control] may exacerbate [social] harms, especially where the method is blackboxed or inherently inaccessible to auditing."

2.3.4 Flexible control

While minimal control, passive control, and post hoc control are most associated with tagging, flexible control (like rigid control, which will be discussed next) is most associated with subject indexing. The most popular indexing language that employs flexible vocabulary control is the Sears List of Subject Headings. The Sears List features a much shorter list of subject headings than the LCSH and instructs indexers that more specific headings are "to be added as needed," following established rules and patterns (Wilson 2022). By making the creation of new subject headings an explicit and encouraged feature of the indexing language, the Sears List balances the benefits of standardized indexing terms with local agency for indexers. For example, this author's used copy of the 21st edition of the Sears List includes handwritten subject headings that represent the LGBTQ+ community. These headings were absent from the Sears List until the U.S. Supreme Court's Obergefell v. Hodges decision provided what Sears editors considered sufficient warrant to add more inclusive subject headings.

However, flexible vocabulary control comes with restrictions that are not present in minimal vocabulary control, passive vocabulary control, or, to some extent, post hoc vocabulary control. For example, the Sears List requires that indexers use an authorized heading if it is available and requires that indexers not use unauthorized entry terms, such as "Clothing designers" when an authorized heading such as "Fashion designers" is available. Additionally, cost can be a downside of flexible vocabulary control. Even though vocabularies like the Sears List are designed to be relatively easy

to use, collections that use flexible control tend to be more expensive to maintain than collections that use any of the forms of control associated with tagging because collections that use flexible control tend to employ professional indexers and catalogers and use cataloging services that charge fees.

2.3.5 Rigid control

With rigid vocabulary control, indexers can use only authorized indexing terms. While rigid control provides indexers with little flexibility, it is used widely to ensure uniform and unique indexing terms and is considered a gold standard for supporting precision and recall in information retrieval. The most widely used indexing language that employs rigid control is the LCSH, which is freely available and closely integrated with cataloging tools and services like MARC, WorldCat, and OCLC Contract Cataloging. This type of integration enforces rigid control by making unauthorized headings difficult and expensive to create and maintain. For example, an indexer using the LCSH might disagree with a problematic subject heading. They could use MARC field 650, control subfield $2 to assign a subject heading that is outside of the LCSH, but that process is labor-intensive and requires local cataloging expertise. Furthermore, rigid control requires centralized approval of new and updated subject headings. Unlike the flexible control of the Sears List, the LCSH requires an extensive review process through the Subject Authority Cooperative Program (SACO) and that committee's approval before a new or updated heading can be used.

Finally, it is important to note that even with rigid vocabulary control, not all authorized subject headings are fully enumerated. The LCSH and the Sears List maintain a list of approved subdivisions that can be used to create synthetic headings. For example, "African Americans" is an authorized LCSH subject heading and "Blogs"' is an authorized LCSH subdivision. If an indexer wanted to index a document about blogs by African Americans, they could assign the synthetic subject heading "African Americans – Blogs" within the rules of rigid vocabulary control even though "African Americans – Blogs" is not enumerated in the LCSH.

2.4 Control in subject indexing and subject tagging

This section reviews literature about the concept of control in KOS, emphasizing how indexers and taggers experience vocabulary control. It then discusses vocabulary control in relation to two axes of control in KOS: 1) feeling in control to feeling controlled, and 2) feeling in control to feeling out of control, both of which are examined extensively in this study and are critical to indexers' ability to promote racial and social justice in their indexing and tagging work.

2.4.1 Controlled vocabularies, systemic bias, and control of indexers

In *Two Kinds of Power*, Wilson (1968) describes two types of bibliographical control: 1) exploitative control, and 2) descriptive control. Exploitative control, an ideal that Wilson argues we cannot reach, would provide each information seeker with the best text or texts available for their specific needs. While Wilson says that exploitative control is unattainable, he argues that descriptive control, which is widely implemented in libraries and other KOS, is an effective and attainable approximation for exploitative control. Olson (2001) problematizes that conclusion, stating that "universal language marginalizes and excludes the other," and argues that effective bibliographical control (even imperfect descriptive control) is unattainable because of systemic biases in controlled vocabularies. These systemic biases have historically meant that members of advantaged groups experience effective descriptive control while members of less advantaged groups do not. Olson argues that actors from historically marginalized groups in controlled roles (including indexers, taggers, and information seekers) suffer a lack of control, agency, and self-representation when their reality is not reflected in a controlled vocabulary.

Like Olson, Furner (2007) finds that indexing languages cannot "provide a value-neutral snapshot of an objective reality." Furner uses a critical race theory (CRT) lens to critique the decision by DDC 22 editors to "decracialize" Table 5, "Racial, Ethnic, and National Groups" in an effort to remove dated terminology like "Mongoloid" and "Negroid" and to make the DDC race-neutral. While sympathetic to the challenges that editors and designers of indexing languages face, Furner criticizes these changes on multiple grounds, most importantly stating "a scheme should support its users in the retrieval of documents about topics relating to the populations with which they self-identify. So, for instance, a person self-identifying as black ought to be able to use the scheme in order to retrieve – easily, effectively, and efficiently – documents about topics relating to black people" (p. 21). Furner's criticisms reveal many aspects at the intersection of race, representation, institutions, and information science. Importantly for this study, Furner shows 1) that indexing languages can perpetuate bias, even when editors have good intentions, 2) that these biases are communicated through what is included in and what is omitted from controlled vocabularies, and 3) that the bias in controlled vocabularies cascades downward from the editors of indexing languages to information seekers.

One promising way for indexers to work around control in rigid controlled vocabularies is to develop their own indexing languages to complement or supplement broad subject headings lists like the LCSH. While this strategy does not guarantee that large indexing languages will adopt in-

dexing terms from these alternative controlled vocabularies, the languages have value as independent entities and can serve as the warrant for new and updated terms in large indexing languages. For example, the Homosaurus linked data vocabulary is "is a robust and cutting-edge vocabulary of LGBTQ-specific terminology that enhances the discoverability of LGBTQ resources" and is "designed to serve as a companion to broad subject term vocabularies" (Cifor and Rawson 2022). That cutting-edge perspective comes from people within the LGBTQ+ community who have lived experiences and domain expertise that inform their vocabulary decisions. The Mashantucket Pequot Thesaurus of American Indian Terminology is another controlled vocabulary built from a historically marginalized perspective. The creators, Sandy Littletree and Cheryl Metoyer, state that while the Thesaurus has value as an independent indexing language, "the primary goal of the Thesaurus is to inform LCSH" (2015). This approach recognizes the top-down structure of control relationships in KOS and smartly works around and within that structure by demonstrating an effective controlled vocabulary for Indigenous materials and establishing compelling warrant for the LCSH to adopt all or part of their controlled vocabulary, which represents a meaningful path forward for indexers who feel controlled by rigid vocabulary control.

2.4.2 Control in less rigid forms of vocabulary control

While criticisms of more rigidly controlled vocabularies show that indexers often lack control or have minimal control, other forms of vocabulary control can afford indexers more immediate control. For example, speakers of Sámi, an indigenous and endangered language from northern Europe, used Twitter hashtags to "increase language use and contribute to the visibility of languages that are not present in mainstream and traditional media" (Cocq 2015 p. 283). Because Twitter has passive vocabulary control for tags, Sámi speakers do not need to propose vocabulary changes to editors or designers and wait for a decision – they just post their tags freely. Of course, being in control to freely use any tag means that these Sámi speakers have little control over the "uncontrolled vocabulary" that comprises the amorphous collection of hashtags on Twitter and have no control over how others tag documents with similar content. Still, taggers like Sámi speakers, can feel empowered to express themselves in the vocabulary that they choose.

Other less rigid forms of vocabulary control can offer similar dynamics of control. For example, the canonical example of flexible control, the Sears List of Subject Headings, encourages indexers to add subject headings as required by their local collection, within certain constraints but without any type of approval from editors of the language. These additions are often made by writing in the margins of a print copy of the Sears List at a local library. While these handwritten annotations afford indexers little control over the centralized Sears List vocabulary, the annotations can give indexers a feeling of control over their local index. And in some cases, local additions are eventually incorporated into the official Sears vocabulary. Perhaps the most visible example of this type of migration is how the subject headings in the Sears List of Subject Headings Canadian Companion were eventually included in the full Sears List.

These differences in feelings of control while working with more rigid and less rigid forms of vocabulary control are central to the questions of this study. Do indexers and taggers experience feelings of control as described here? Does vocabulary control affect their processes and the indexing terms that they assign? How do indexers' and taggers' experiences and choices of indexing terms relate to criticisms of controlled vocabularies and racial and social justice?

2.4.3 Controlled, in control, and out of control

So far, this paper has contrasted feeling in control (i.e., Sámi language use on Twitter and adding local Sears List subject headings) with feeling controlled (i.e., Dartmouth librarians struggling to change the "Illegal aliens" subject headings). While these examples suggest that indexer and tagger control decreases as vocabulary control becomes more rigid, these examples consider only one axis of control, an axis with "in control" at one end and "controlled" at the other end. However, this study proposes that there are at least two axes of control for subject indexers and subject taggers. The other axis also has "in control" at one end but has "out of control" at the other end.

The second axis, between feeling in control and feeling out of control, reveals an important dynamic for subject indexers and taggers. For example, one might expect that more rigid forms of vocabulary control might make some indexers and taggers feel more controlled and less in control because they must follow rules and conform to an established vocabulary. Conversely, one could expect that the lack of guidance provided by minimal vocabulary control might make some indexers and taggers feel out of control as they attempt to determine useful indexing terms without the constraints (and guidance) provided by authorized indexing terms.

Understanding the dynamics of these two axes is particularly relevant for the "middle" forms of vocabulary control in this study (passive vocabulary control, post hoc vocabulary control, and flexible vocabulary control) because they have not been studied as extensively. For example, do taggers feel less out of control with post hoc control because they know that they can tag freely and have editors "fix" any tags that do not conform to the controlled vocabulary? Or do

they feel more controlled because they know that the tags that they choose are subject to editing by someone else? Do indexers and taggers using passive vocabulary control feel controlled by tags suggested by autocomplete? Or do they feel out of control because they still have complete freedom to choose their own tags?

Finally, this discussion of different forms of vocabulary control, and feelings of control is incomplete without addressing the relationship between control and ease of use, particularly for inexperienced indexers and taggers. This study asks inexperienced indexers and taggers to index and tag books using different forms of vocabulary control, including forms of vocabulary control represented by complex and possibly intimidating indexing languages like the Sears List and the LCSH. Indexers and taggers with limited exposure to these indexing languages might find them difficult to use and might encounter confusion and roadblocks unrelated to the focus of this study. For example, some participants might be confused by how to construct synthetic subject headings by using subdivisions and other indexers might be intimidated by the size of the print copy of the Sears List or the user interface for searching Library of Congress Authorities. This study used a series of questions and ranking activities to understand how feelings of control relate to ease of use, with the aim of homing in on the effect of vocabulary control on feelings of control across two axes in the subject indexing and subject tagging processes.

3.0 Method

This empirical study employed a mixed-methods approach, combining both qualitative and quantitative methods and data analysis. The two central research methods in this study were: 1) a think-aloud protocol, as developed by Ericsson and Simon (1984) and described by Charters (2003), and 2) a retrospective interview, which was conducted with each participant after they completed indexing and tagging tasks. The think-aloud protocol was conducted during the subject indexing and subject tagging portion of the study. Participants' verbalized thoughts were recorded, transcribed, and coded, and their aboutness statements and subject headings or tags were collected for each indexing and tagging task. The retrospective interview asked participants to reflect on their experiences with indexing and tagging. Like the think-aloud protocol, the interview session was recorded, transcribed, and coded. In addition to participants' verbal responses, this portion of the study collected participants' rankings of the five forms of vocabulary control on different axes of control.

3.1 Participants

Participants (n=18) were recruited from a large Masters of Library and Information Science (MLIS) program, a student population that fit the desired profile of novice indexers and taggers. Seventeen participants in the study had completed or were currently enrolled in a KO theory course, and four participants had completed a cataloging course. Recruiting materials targeted another aspect of the preferred participant profile, a desire to promote racial and social justice through indexing and tagging, by noting the study's focus on these topics. Participants were surveyed about their familiarity with racial and social justice literature. All participants were at least somewhat familiar with the literature, with 66.7% of participants very familiar with the literature.

The participants recruited for the study somewhat matched the demographics of the American Library Association (ALA) members. For example, the 2017 ALA Demographic Survey (Rosa and Henke 2017) reports that 81% of members identify as "Female," and 77.8% of study participants identified as "Female," "Woman," "Cis Female," or "ciswoman." Participants in the study were slightly more racially diverse than members of the ALA, with 66.7% of participants identifying as "White" or "European American," compared to 81% of ALA members identifying as "White." As might be expected when recruiting novice indexers from an academic program, participants were much younger overall than the average age of ALA members, with 61.1% of participants reporting that their age was between 25 and 29 years old. Only 11.2% of participants were 40 years of age or older. Participants were given $20 gift cards to thank them for their participation.

3.2 Procedure

The study was conducted in four major parts: 1) a survey to collect demographic data and information about participants' education and experience, followed by a welcome to the study and a consent form, 2) a think-aloud protocol to capture participants' thoughts while they analyzed each book and wrote an aboutness statement to summarize each book, 3) a continuation of the think-aloud protocol in which participants translated their aboutness statements into subject headings or tags using different forms of vocabulary control, and 4) a retrospective interview designed to help participants reflect on the subject indexing and tagging tasks that they just completed and the forms of vocabulary control that they used. Research sessions were conducted in a conference room and lasted approximately one-and-a-half to two hours. Sessions were conducted with one participant at a time and used the procedure described in the following subsections.

3.2.1 Demographic Survey, Welcome, and Consent

Before participants arrived at the study, they were emailed a link to a demographic survey. Each participant provided basic demographic data (age, gender, race, and ethnicity), information about their education and experience, and data about their familiarity with books and authors related to racial and social justice. This data is summarized in the Participants section above. When each participant arrived for the study, they were welcomed and briefed on the study, including how long it would take, what they would do, and the goals and background of the study. They also completed a consent form and were told that they could stop the study at any time for any reason.

3.2.2 Subject analysis

Participants were handed books to tag or index, one book at a time in randomized order to avoid order effects. The books were all nonfiction books that address race and society, primarily in the United States, from different perspectives: 1) *Caste: The Origins of Our Discontents* by Isabel Wilkerson, 2) *White Fragility: Why It's So Hard for White People to Talk About Racism* by Robin DiAngelo, 3) *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit It* by Richard Sander and Stuart Taylor Jr., and 4) *The End of Racism: Principles for a Multiracial Society* by Dinesh D'Souza.

For each book, participants were asked to analyze the book to determine its subject or subjects and were instructed to write an aboutness statement summarizing the book. During this portion of the study (and the next portion), participants were asked to "think aloud," and these verbalized thoughts were captured via audio recording. Additionally, the researcher recorded notes to capture non-verbal data like facial expressions and posture and to track which parts of the bibliographic apparatus (e.g. summaries on the cover and inside flap, table of contents, index) participants used. Participants' written aboutness statements, an example of which is shown in Figure 1, were collected at the end of the research session.

3.2.3 Subject representation

After participants indicated that they were satisfied with the aboutness statement for each book, they were introduced to the form of vocabulary control that they would use to tag or index that book and to the materials and apparatus that they would use with that form of vocabulary control. The researcher explained to participants that the subject representation step involved translating their aboutness statement and their understanding of what the book was about into a subject representation in the form of tags or subject headings. The researcher also explained the mechanics and the "rules" of working with the form of vocabulary control presented to the participant. For example, for participants working with rigid vocabulary control, the researcher showed participants how to look up LCSH headings and explained that they could use only authorized headings.

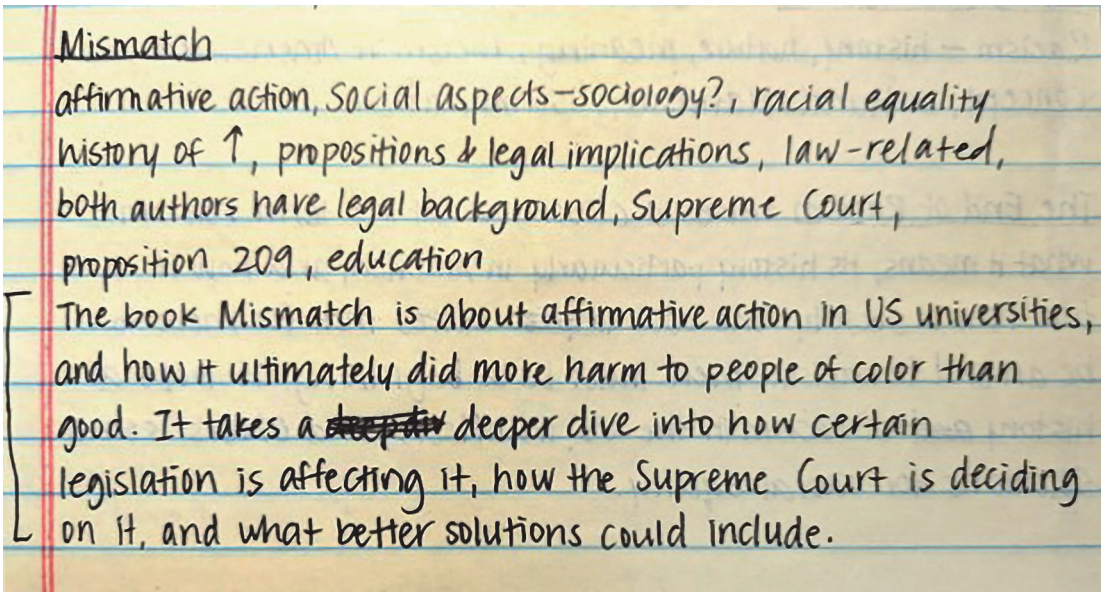

Figure 1. An example of an aboutness statement written by one of the participants to summarize what the book *Mismatch* is about. The aboutness statement itself is marked with a line in the left margin. This participant was one of a handful of participants who wrote notes (above the aboutness statement in this case) in addition to the aboutness statement itself.

Participants indexed one book for each of four forms of vocabulary control:

1) Minimal control: Free tagging with no controls and minimal guidance.
2) Passive control: Tagging with suggested vocabulary that became visible via autocomplete as users entered text into a user interface.
3) Flexible control: Subject indexing with a physical copy of the *Sears List of Subject Headings*, which allows local additions to the vocabulary if indexers follow specific rules and patterns.
4) Rigid control: Subject indexing with the LCSH, which allows only authorized subject headings. Participants used the web interface at https://authorities.loc.gov/ to find authorized subject headings.

As with the order of the books, the order of the forms of vocabulary control was randomized to avoid order effects. Participants did not use post hoc vocabulary control in this portion of the study because it was not readily re-recreated in an experimental setting. Instead, participants were asked about their reactions to post hoc editorial changes to tags or subject headings in the interview portion of the study. For the four forms of vocabulary control studied in this portion, participants entered their subject headings or tags in a simple web application, as shown in Figure 2.

3.2.4 Retrospective interview

After completing the subject analysis and subject representation portions of the study for all four books, participants were asked a series of interview questions about their experiences. The interview followed a consistent script and structure across participants, with some room allowed for follow-up questions. While the interview covered a wide range of questions, including questions about specificity, exhaustivity, and accuracy, this paper focuses on the interview questions that directly related to participants' feelings about the relationship between the form of vocabulary control, the feelings of control that they experienced, and their ability to support racial and social justice through their subject indexing and subject tagging decisions. Three parts of the retrospective interview investigated this relationship:

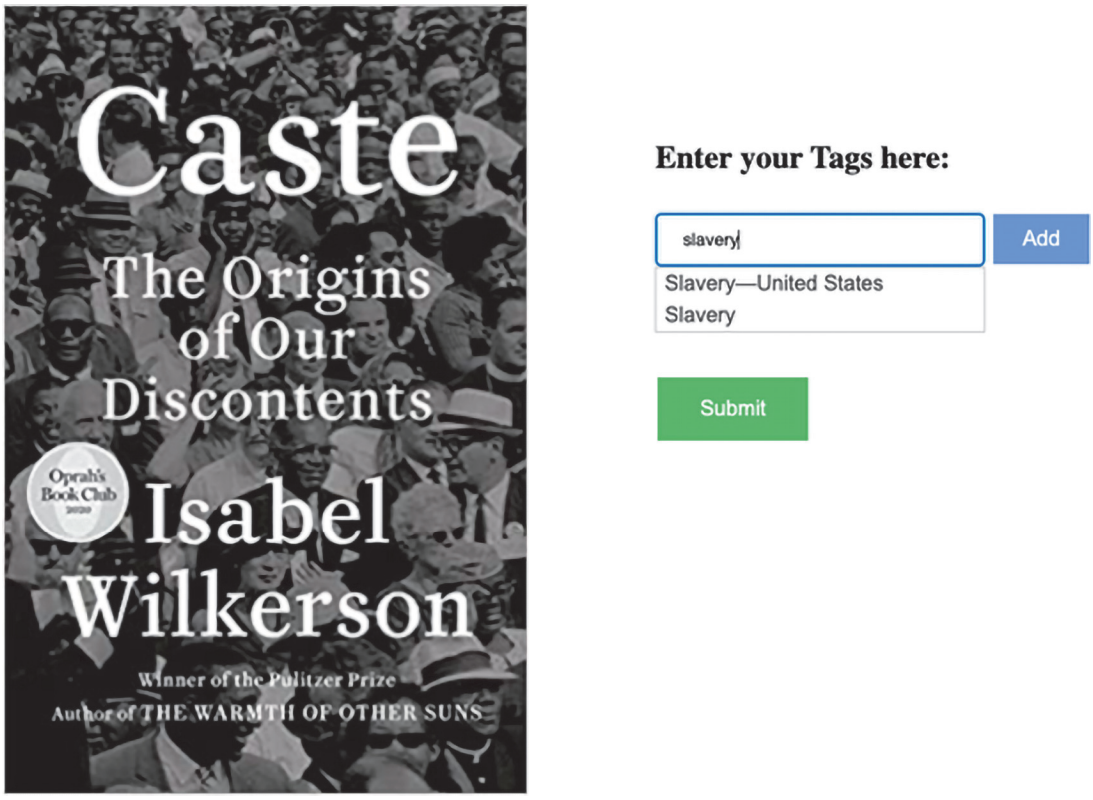

Figure 2. The user interface for collecting participants' tags and subject headings during the think-aloud protocol. This screenshot shows passive vocabulary control, as the word "Tags" is displayed instead of "Subject Headings" and the autocomplete feature is enabled. The autocomplete vocabulary included roughly 500 relevant indexing terms taken from the Sears List of Subject Headings.

1) Coextensiveness between aboutness statements and subject headings and tags: Participants were asked a series of questions that investigated whether they felt that the form of vocabulary control for each book that they tagged or indexed caused them to change the language that they used in their aboutness statements during translation into subject headings and tags – and whether they perceived those changes as helping or harming their ability to support racial and social equity. For example, participants were asked: "Did you feel that your subject analysis (the first step for each book) matched with the subject headings or the tags that you ultimately chose? Why or why not?" and "What influenced your choices of tags or subject headings in relation to your aboutness statement?"
2) Feelings of being "in control," "controlled," and "out of control": Participants used note cards to rank each of the five forms of vocabulary control 1) from easiest-to-use to hardest-to-use, 2) from making them feel most in control to most controlled, and 3) from making them feel most in control to most out of control. Participants were then asked to explain what "in control," "controlled," and "out of control" meant to them and to discuss how those feelings affected their indexing and tagging work and the subject headings and tags that they chose and did not choose.
3) Racial and social justice: Participants were asked open-ended and follow-up questions about how the coextensiveness between their aboutness statements and their subject headings and tags and about and how their feelings of control related to their ability to support racial and social justice.

As with the previous portions of the study, the interview used audio recording to capture participant responses, which were later transcribed and coded.

3.3 Analysis

This mixed-methods study produced a significant amount of quantitative and qualitative data. This section focuses on how the following data was analyzed: 1) aboutness statements provided by participants, 2) subject headings and tags assigned to books by participants, 3) transcripts and notes collected during the think-aloud protocol, 4) participants' answers to questions about the coextensiveness of their aboutness statements and their subject headings and tags, 5) participants' rankings of the five forms of vocabulary control on scales from easiest-to-use to hardest-to-use, making them feel most in control to most controlled, and from making them feel most in control to most out of control, and 6) participants' answers to follow up interview questions about racial and social justice.

3.3.1 Analyzing feelings of control

This study aimed to understand indexers' and taggers' feelings related to control as they worked with different forms of vocabulary control. Participants were asked to rank the five forms of vocabulary control on scales from easiest-to-use to hardest-to-use, making them feel most in control to most controlled, and from making them feel most in control to most out of control, and to explain their rationale for their rankings. The aggregate rankings for each form of vocabulary control were calculated for each scale and the participants' explanations for their rankings were used to contextualize the findings. Rankings were collected with and without post hoc control, which was not included in the think-aloud protocol, but which was explained and discussed during the interview.

3.3.2 Analyzing the effect of vocabulary control on coextensiveness

Coextensiveness was defined in this study as the degree to which the aboutness statement that was produced in the subject analysis phase matched the subject headings or tags that were finalized in the subject representation phase. This study investigated coextensiveness to better understand how vocabulary control affected indexers and taggers' choices when they translated their aboutness statement into indexing terms and how those choices aided or hindered social and racial equity. The aboutness statements and subject headings and tags were analyzed and concepts in each were recorded to determine the coextensiveness of aboutness statements and subject headings and tags. For example, the aboutness statement shown in Figure 1 includes the following concepts: "Affirmative action," "US," "universities," "harm," "people of color," "legislation," "Supreme Court," and "solutions." These concepts were then compared to the subject headings or tags that the participant assigned to that book.

This analysis of the coextensiveness of aboutness statements and the subject headings and tags assigned to books used a taxonomy of transformations developed by Bertrand and Cellier (1995), which is organized into subtractive transformations, additive transformations, and other transformations.

Subtractive transformations were determined as follows:

1) Suppression was determined by identifying a topic or concept that was present in the aboutness statement but absent in the indexing terms. For example, a participant's aboutness statement expressed the concept "Blackness" but their indexing terms did not include an equivalent term.
2) Simplification was determined by identifying complex concepts or topics (i.e., constructed with multiple nouns

or with a noun and an adjective) that were made less specific by removing detail from the original concept or topic. For example, a participant's aboutness statement might have expressed the concept "Elementary school libraries," but they might have used a separate indexing term "Libraries" in the subject representation step. Simplification was not observed in this study.

3) Generalization was determined by identifying a specific concept in the aboutness statement that was expressed with a more general indexing term. For example, one participant's aboutness statement described the concept "caste" and they used the more general subject heading "social classes" in the representation step.
4) Dissociation begins with a complex (multi-part) concept and splits it into two indexing terms. For example, the concept "Art of Black people" might be represented by two dissociated indexing terms: "Art" and "Blacks." This study anticipated that dissociation (for example, separating "Black" and "Lesbian" instead of representing the intersectionality in a single indexing term "Black lesbians") would be highly relevant to this study because critical cataloging discusses intersectionality extensively (Olson 2001; Fox 2016; Howard and Knowlton 2018). However, dissociation was not observed in this study.

Additive transformations were determined as follows:

1) Addition was determined by identifying indexing terms that were used in the final subject representation but that were not part of the original aboutness statement. For example, a participant did not have the concept "Discrimination" in their aboutness statement, but they included "Discrimination" in their subject headings for that book. Addition proved to have a complex relationship with vocabulary control. Addition with minimal and passive vocabulary control was typically driven by the ease of adding tags. Addition with passive, flexible, and rigid vocabulary control was typically driven by participants "finding" or "discovering" additional indexing terms as they worked with the autocomplete user interface or the apparatus for the Sears List or the LCSH.
2) Precision was determined by identifying an indexing term that was made more specific than the concept expressed in an aboutness statement through the addition of a noun, adjective, or subdivision. For example, one participant's aboutness statement expressed the general concept "Education," which they made more precise with the indexing term "Higher Education." Precision is considered additive not because it adds more subject headings or tags but because it adds specificity.

Finally, substitutive transformations are neither subtractive or additive:

1) Substitution was determined by identifying a concept from an aboutness statement that was changed in form but that maintained the same or close to the same meaning. For example, "Racism" changing to "Race discrimination," "Universities" changing to "Colleges and universities," and "American History" changing to "United States – History." Bertrand and Cellier (1995) state that substitution "results in neither gain nor in loss." However, substitution is hardly a transformation where nothing happens, and might even be the most important type of transformation in the current study. For example, the critical differences between "undocumented immigrants" and "Illegal aliens" could be classified as substitution, but these similar terms are significantly different in their connotation and in how indexers, taggers, and information seekers experience them.

In addition to analyzing transformations from aboutness statements to subject headings and tags, this study asked participants questions about their perceptions of the coextensiveness of their aboutness statements and for explanations of why their aboutness statements and indexing terms were coextensive or not for each book that they tagged or indexed. Responses were coded as "Strong coextensiveness," "Somewhat strong coextensiveness," "Somewhat weak coextensiveness," and "Weak coextensiveness."

Finally, transcripts and notes from the think-aloud protocol were coded, and codes related to coextensiveness were identified. For example: "Frustration over vocabulary control limiting choices" and "Satisfaction with vocabulary control providing clarity" were used to mark instances when vocabulary control affected indexing and tagging choices and to mark key quotes.

3.3.3 Analyzing connections to racial and social justice

While analyzing how vocabulary control affects coextensiveness and participants' feelings of control during subject indexing and subject tagging informs our understanding of the relationship between vocabulary control and racial and social justice, these data do not directly address racial and social justice. To make these connections, participants were asked open-ended and follow-up questions about how the different forms of vocabulary control and their feelings about control during the process informed their ability to support racial and social justice. Their answers were transcribed and coded, and representative quotes are shared in the Results and Discussion section.

4.0 Results and discussion

In this study, 18 participants tagged and indexed four books each, using four different forms of vocabulary control. They wrote 72 total aboutness statements and assigned 568 total subject headings and tags. Participants also answered 15 interview questions (with many sub-questions) each and ranked different forms of vocabulary control on three different scales related to control. Their thoughts while they indexed and tagged and their answers to interview questions were captured in 29 hours of audio recordings. This section reports the quantitative and qualitative results from analyzing this data.

4.1 Feeling most in control to feeling most controlled

This study hypothesized that participants would feel more controlled by more rigid forms of vocabulary control and more in control when working with less rigid forms of vocabulary control. Table 1 shows the 18 participants' mean rankings on the scale from most in control (1) to most controlled (5) for each of the five forms of vocabulary control in the study. In line with the hypothesis, minimal vocabulary control made participants feel most in control overall (1.33 mean ranking) and rigid vocabulary control made participants feel most controlled (4.11 mean ranking). A non-parametric Friedman test found that the difference in participants' rankings was significant (chi-square=36.3, p-value<0.05). Based on this statistical analysis, the study concludes that novice and advanced beginner subject indexers and taggers feel most controlled by rigid vocabulary control and feel increasingly in control with flexible vocabulary control, post hoc vocabulary control, and passive vocabulary control.

Qualitative analysis added context to this finding and informed how it relates to racial and social justice. For example, while using minimal vocabulary control, one participant noted that she had seen the subject heading "Whites" in the LCSH but that she preferred the indexing term "White people" because it matched her own contemporary vocabulary and because "Whites" sounded problematic to her. She chose to use "White people" instead "because I can do that [with minimal vocabulary control]." The freedom to use any tags that they wanted was a recurrent theme when participants explained their ranking for minimal vocabulary control.

When asked what "controlled" meant to them, participants focused on four themes: 1) feeling constrained or restricted, 2) feeling that another agent had control over their actions, 3) feeling that a lack of appropriate terms in the vocabulary controlled them, and 4) feeling that rules or processes controlled them. One participant went so far as to say that the LCSH "owns me," while other participants described post hoc control as a "looming editor" and "like putting words in my mouth." The most common theme related to feeling controlled was the lack of authorized terms that participants wanted to use. For example, one participant said, "Ugh! I was looking for 'White Supremacy' and it comes up with 'White Supremacy Movements' and that's the only tag [sic]. That's not really what this book is about." Another participant dramatically communicated this theme of feeling controlled by a vocabulary: "Oh, man, with the LCSH, you gotta use some terms that hurt your soul sometimes." Yet another participant lamented that the LCSH lacked the term that they wanted to use ("Antiblackness") and commented on how slow the LCSH is to change.

Based on these quantitative and qualitative analyses of participants' rankings of forms of vocabulary on a scale from in control to controlled, this study found: 1) that novice indexers and taggers feel significantly more in control when using minimal and passive vocabulary control and significantly less in control when using post hoc, flexible, and rigid vocabulary control, 2) that being controlled by a form of vocabulary control is a feeling of being constrained or constricted by rules and processes associated with a form of vocabulary control, by terms missing from a vocabulary, by problematic terms in the vocabulary, and by someone or something being able to exert control over the indexer or tagger, and 3) that feeling controlled directly affected some participants' ability to represent marginalized groups using language that they viewed as supporting racial and social justice.

4.2 Feeling most in control to feeling most out of control

When participants were asked to rank forms of vocabulary control from those that made them feel most in control to those that made them feel most out of control, their rankings had a bimodal distribution that clouded the relationship be-

	Minimal	Passive	Flexible	Post Hoc	Rigid
Mean Ranking	1.33	2.44	3.33	3.78	4.11

Table 1. The mean rankings for all five forms of vocabulary control on the scale from participants feeling most in control (1) to participants feeling most controlled (5), showing that participants felt most in control with minimal vocabulary control and most controlled with rigid vocabulary control.

tween vocabulary control and feelings of being in control and out of control. The bimodal distribution was likely based on differing interpretations of the phrase "out of control." Some participants described being out of control in very similar terms to how they described being controlled: a controlled vocabulary was dictating the indexing terms that they could use. These participants tended to rank minimal vocabulary control as the form of vocabulary control with which they felt most in control. In contrast, other participants described being out of control as lacking guard rails and not really knowing what they were doing. These participants tended to rank minimal vocabulary control as the form of vocabulary control with which they felt most out of control.

When rankings for just the participants who described being out of control as lacking guard rails were grouped together, that subset of results revealed a clear pattern, with tagging-centric forms of vocabulary control ranked as making participants feel more out of control and controlled vocabularies ranked as making participants feel more in control, as shown in Table 2. A Friedman Rank Sum Test found that the difference in these participants' rankings was significant (chi-square=15.2, p-value<0.05).

Qualitative analysis found that novice indexers and taggers appreciated the guidance provided by flexible control, rigid control, and even the autocomplete feature of passive control. One participant stated: "When I felt the most confident on whether the terms were 'legal' and whether the terms captured what I wanted, actually, was *Caste*, even though [that book] was the hardest for me, because I had [the Sears List] to flip through. It had the terms all laid out and it had rules like, 'May subdivide geographically.'" Another participant noted that broader and narrower terms helped her to achieve specificity and found that rigid control made her feel in control, even when it limited her vocabulary choices: "It's easier when you have a language to work with because it gives you bounds within which to think. Okay, I am working with the Library of Congress. They might not have this more modern word that I am leaning toward. They might not have that, so I'll just cross that off my brain and go with something else. So, when you have a language to work with, I feel like that reins you in a little bit." Passive vocabulary control was cited for balancing support and freedom. For example, one participant liked how the autocomplete feature helped him to come up with terms, and another participant liked that she could "work within an existing system, but still had the freedom to choose [her] own tags."

In contrast, participants described working with minimal control as "too much autonomy. I don't know if what I'm doing is right," "flying solo," "shouting into the void," and "choice paralysis," suggesting that free tagging did not provide sufficient support and guidance. Another participant used a colorful description, describing minimal vocabulary control as being "left to the wind," "having no direction," and needing to "come up with terms on my own." Yet another participant said that they felt out of control when they had "no context of what's appropriate," which made them nervous and not confident in their indexing terms. Finally, another participant stated, "Using free tagging meant I had to rely on my own understanding, which was not that helpful."

Based on these quantitative and qualitative analyses of participants' rankings of forms of vocabulary on a scale from in control to out of control, this study finds: 1) that novice indexers and taggers can feel out of control in two opposing ways: because they are restricted by the form of vocabulary control and lack agency or because they have too much autonomy and the form of vocabulary control lacks support and feedback to help them, and 2) that minimal control is more likely make novice indexers and taggers feel out of control if they equate being out of control with having too much autonomy. These findings, especially the finding that indexers and taggers feel out of control when they have too much autonomy and not enough guidance, are likely specific to novice participants. This study suggests additional research to understand the relationship between control, autonomy, and minimal guidance for more experienced indexers and taggers. Additionally, this study suggests that any future research on feelings of control use a phrase other than "out of control" or fully clarify that phrase with participants to ensure a consistent interpretation.

4.3 Coextensiveness of aboutness statements and subject headings and tags

Coextensiveness was defined in this study as the degree to which the aboutness statement that was produced in the

	Flexible	Rigid	Passive	Post Hoc	Minimal
Mean Ranking	1.57	2.57	2.71	3.42	4.71

Table 2. The mean rankings for all five forms of vocabulary control on the scale from participants feeling most in control (1) to participants feeling most out of control (5), calculated based on the subset of participants who described being out of control as lacking guard rails and not really knowing what they were doing.

subject analysis phase matched the subject headings or tags that were finalized in the subject representation phase. This study investigated coextensiveness to better understand how vocabulary control affected the choices that indexers and taggers made when they translated their aboutness statement into indexing terms and how transformations during that translation process aided or hindered racial and social justice. This study found that there is not a simple and direct correlation between coextensiveness and control, but that there is a complex relationship between coextensiveness, feelings of control, and vocabulary control. The following subsections explore that relationship by 1) quantitatively analyzing the coextensiveness of aboutness statements and indexing terms and 2) using participants' responses to interview questions about coextensiveness and thematic analysis to expand upon that quantitative analysis.

While there was not a simple relationship between vocabulary control and coextensiveness, some large trends did emerge in the quantitative analysis. Table 3 summarizes the frequency of different types of transformations for each form of vocabulary control, including "unchanged," which represents complete coextensiveness or no change in a concept from aboutness statement to subject heading or tag. The table shows that suppression and generalization transformations increased as vocabulary control became more rigid and addition transformations and "unchanged" (no transformation) concepts increased as vocabulary control became less rigid. These data suggest that more rigid forms of vocabulary control are more likely to cause novice indexers to omit concepts that they identified in their aboutness statements. They are also more likely to represent concepts as they originally identified and expressed them when they use less rigid forms of vocabulary control. Taken together, these findings suggest that novice indexers have more control over their subject headings and tags when they use less rigid forms of vocabulary control.

However, precision transformations were highest with passive control and flexible control, suggesting that these forms of vocabulary control best help novice indexers and taggers assign more specific subject headings and tags. In contrast, these two forms of vocabulary control also have the highest rates of substitution transformations, suggesting either that they force indexers and taggers to use alternative terminology or that they make it easy to identify preferred alternative subject representations. Based on the lower rate of substitution transformations for rigid vocabulary control, many of the substitution transformations for passive and flexible control were participants' choices, indicating that they were more in control while using these forms of vocabulary control. This finding was evidenced by two techniques: 1) typing the roots of terms (i.e., "rac...") in an autocomplete user interface, and 2) looking for "nearby" terms in the Sears List to find indexing terms. Finally, the forms of vocabulary control associated with tagging (minimal control and passive control) had the highest rates of addition transformations, suggesting that assigning additional terms that did not appear in the aboutness statement is easiest with less rigid forms of vocabulary control.

Qualitative analysis agreed with and expanded upon the quantitative analysis of coextensiveness. For example, in many cases, participants suppressed terms because of vocabulary control. In one case, a participant included the con-

	Minimal	**Passive**	**Flexible**	**Rigid**
Total Concepts	200	215	186	196
Suppression	21%	20%	29.6%	39.8%
Unchanged	20.5%	14.4%	12.4%	10.7%
Generalization	0%	0.5%	3.8%	1.5%
Substitution	2.5%	7.4%	9.1%	4.1%
Precision	5%	14%	11.3%	11.2%
Addition	51%	43.7%	33.9%	32.7%

Table 3. Comparison of the frequency of different types of transformations performed by participants when they translated their aboutness statements to indexing terms while using four different forms of vocabulary control. "Unchanged" represents complete coextensiveness between a concept in the aboutness statement and that concept's representation in an indexing term. Note that suppression and generalization decrease as vocabulary control becomes less rigid and that addition and "unchanged" increase as vocabulary control becomes less rigid.

cept "public policy" in her aboutness statement; however, she did not find a "Public policy" subject heading in the Sears List and ended up suppressing that concept. Another participant searched the Sears List for "Anti-blackness" while stating that it was "probably not there" and quickly moved on after remarking, "I was right." Participants arrived at generalization transformations under similar circumstances. One participant wanted to represent the concept of "systemic racism" while working with the LCSH. When they could not find "systemic racism" in the authorized headings, they settled on the more general heading "racism." In this case, the indexing language controlled the terminology that they could use and led them to a less specific and less expressive representation of the original concept. Similarly, another participant substituted the Sears List authorized heading "Racism" for their original concept of "Antiracism." While antonyms can often represent the same concept (for example, "Illiteracy" is an entry term for "Literacy" in the LCSH), the difference between the intent to use "Antiracism" and the substitution of "Racism" is marked and suggests that the form of vocabulary control affected their subject indexing process and their ability to support racial and social justice.

However, not all transformations had a negative or controlling effect. Insead, when the workflow for a transformation met participants' expectations, they did not feel a lack of control. For example, one participant was satisfied when she searched for "Affirmative action" and found the LCSH authorized heading "Affirmative action programs." She said, "I like that one. I think that one sounds like it fits." Another participant had the concept "racial integration" in her aboutness statement and entered "Integration" in the tagging user interface. When she found "School integration," she quickly selected it and was satisfied with (and not controlled by) the substitution. Another participant identified the concept "Whiteness" while writing an aboutness statement for *White Fragility* and then found the more precise term "Whiteness (race identity)" in the LCSH and felt empowered. Similarly, another participant found the phrase heading "racism against black people" in the LCSH and chose that more precise term over her original concept of "Racism." Thematic analysis showed that participants generally felt "in control," "supported by the form of vocabulary control," and "happiness with the subject indexing and subject tagging processes"' when they identified specific terms in a vocabulary and made precision transformations.

In response to interview questions about coextensiveness, participants consistently indicated that different forms of vocabulary control affected coextensiveness and, in some cases, connected that effect to feelings of control. The two most frequently identified codes for responses to interview questions about coextensiveness were the general code "Influence of vocabulary control" (count=13) and the more specific code "Vocabulary control effect on term choice" (count=11). One participant spoke at length about how different forms of vocabulary control affected their translation and subject representation work:

> With *Caste*, when I was just free typing, I'd just pick out words and be like "that's a subject heading" or "that feels like a good representation of what this book is about," versus having a set list of words I could use. I used that same kind of thought process for all the other ones, but it was regulated more. Sometimes I'd look for something in [the] Library of Congress and it wouldn't exist, and I was just like, okay, I guess I don't have that tag [sic]. And then with the Sears List, I'd start looking for something and I'd find something else, which I really liked because I feel like when I'm told to think of something, my brain is like, "There are no words." So, it's nice, especially having the broader topics or the narrower topics listed underneath because I'd be like, "I know I want to do something with race. Let's see what this book has to say as narrower topics for that." So that made it easier to come up with accurate headings.

This quote aligned with many other participants' responses and with the findings from the quantitative analysis of coextensiveness. In this quote, the participant noted that minimal vocabulary control allowed her the freedom to choose her own indexing terms without regulation, which fit with minimal vocabulary control's low rate of suppression transformations and high rate of concepts that remained unchanged after translation. She also noted that her thought process was the same for all the forms of vocabulary control, but that rigid vocabulary control was more likely to cause suppression transformations and suppression of concepts that participants thought of during the subject representation step. Her use of the word "regulated" and frustration in saying, "I guess I don't have that tag," showed that she felt controlled by rigid vocabulary control and that the control had a negative effect on coextensiveness and subject indexing as a whole. However, not all participants connected suppression with control. Finally, this quote details the browsing technique described above that is associated with transformations for passive control and flexible control, specifically browsing the Sears List and finding narrower terms that resulted in addition and precision transformations.

Based on these quantitative and qualitative analyses of coextensiveness and different forms of vocabulary, this study finds: 1) that suppression and generalization transformations increase as vocabulary control becomes more rigid and addition transformations and "unchanged" (no transformation) concepts increase as vocabulary control becomes less rigid, 2) that there is a complex relationship between co-

extensiveness and control, not a direct relationship between feeling in control and producing high coextensiveness, 3) that novice indexers and taggers felt both in control and supported when controlled vocabularies provided what they considered positive transformations, and 4) that novice indexers and taggers use a browsing technique, either by typing the roots of terms (i.e., "rac...") in an autocomplete user interface or by looking for "nearby" terms in the Sears List, to help them find indexing terms, especially substitute indexing terms and more precise indexing terms.

4.4 Vocabulary control and racial and social justice

This section reports on qualitative analysis of participant reactions and reflections that were coded as relating to racial and social justice and considers how the emergent themes relate to critical cataloging. The section includes participants' criticisms of problematic or missing terms in controlled vocabularies, how they responded in those situations, participants' reflections on how their identities (including racial identities) informed their indexing and tagging processes, and the effect of different forms of vocabulary on participants' ability to navigate complex racial and social issues in indexing and tagging.

4.4.1 Working with and around problematic and missing indexing terms

Many of the problematic or missing indexing terms that participants encountered during the study were problematic for subtle reasons. For example, one participant began with the concept "Inequity" in her aboutness statement for *White Fragility.* She looked up "Inequity" in the Sears List and did not find it; however, she found the entry term "Inequality" because it was close to "Inequity" alphabetically. "Inequality" referred her to the authorized heading "Equality," which might seem like a somewhat close translation of the original concept 'Inequity', but Participant 11 chose to suppress the concept entirely, stating: "They have 'Equality' but they don't have 'Equity', which I've learned are different things." This case shows the importance of controlled vocabularies moving quickly to adopt contemporary language and shows the advantage of other forms of vocabulary control that would have allowed the participant to assign "Inequity" and convey the meaning that they intended instead of suppressing the concept.

In another case, a participant looked for "Multiracial" in the Sears List based on that word being used in *The End of Racism*. She found "Multiculturalism" instead, read the scope note, and saw "Pluralism (Social sciences)" mentioned: "Materials on the coexistence of several distinct ethnic, religious, or cultural groups within one society are entered under 'Pluralism (Social science).'" Based on the scope note and her misgivings about "Multiculturalism" being used to minimize minorities in some contemporary usage, she decided to assign 'Pluralism (Social sciences)' as a subject heading. This case shows the importance of scope notes in guiding indexers and taggers while they work with complex social and racial issues. The novice indexers in the study were eager for guidance as they indexed and tagged, and scope notes (if they are clearly written and give advice that aligns with racial and social justice) can provide invaluable guidance. Notably, the only indexing languages in this study with scope notes were the Sears List and the LCSH, which are often criticized for being slow to update their vocabularies and the scope notes that accompany those vocabularies. Alternatively, the forms of vocabulary control associated with tagging generally do not offer scope notes, although, for example, Stack Overflow does provide short definitions and minimal scope notes for existing tags. Based on this case, this study recommends that scope notes, especially for indexing terms related to complex social issues, be a point of emphasis for the editors of indexing languages.

Perhaps the most problematic subject heading that a participant encountered during the study was "Indians of North America," which one participant encountered in the LCSH while trying to assign subject headings related to India for the book *Caste*. The following extended quote comes from this participant's stream-of-consciousness thoughts during this part of their subject indexing process:

> I'd like to have [a subject heading] for India...Oh, this is going to be fun [sarcasm]. "Indians of North America" is so big everywhere, it has so many headings [in the LCSH]. I just want to find out about Indian people. India sub-divided by History. Oh, god, [these results are] interspersed with "First Nations" and "Indian Art---Canada." I'm not going to be able to find what I'm looking for over here. So much of "Indians of North America." Who would want to use that? What I'm looking for, I'm not going to find...something about social stuff in India.

This quote illuminates a few points. First, "Indians of North America" is such a notorious subject heading that even a novice indexer was quick to note its presence and flag it as problematic. Second, problematic headings can make indexers and taggers feel out of control, even when the indexer or tagger does not need to use the heading or a related term. In this case, the participant was upset and distracted by the term, and they had to deal with ambiguity related to "Indian" carrying multiple meanings. Third, the tool for identifying authorized headings matters, as the clunky search results page on https://authorities.loc.gov/ made "Indians of North America" stand out through repetition and hid the terms related to India that the participant

wanted to find. In this case, the participant was ultimately successful and satisfied because they noted a pattern for subdivisions under "African Americans" and followed that pattern to assign synthetic headings: "India---Social life and customs" and "India---Race identity," but they felt out of control during the process.

Based on analysis of themes related to working with and around problematic and missing indexing terms, this study found: 1) that problematic and missing headings can challenge indexers even, or especially, when what makes them problematic is subtle, 2) that problematic headings cause harm not just when they are assigned but also when they confuse indexers and taggers processes, 3) that empathetic, up-to-date, and substantive scope notes can help indexers and taggers, especially inexperienced ones, feel more in control, 4) that indexers would feel more in control of their processes and the indexing terms that they assign if controlled vocabularies were not overly cautious about updating indexing terms related to racial and social issues, 5) that eliminating vocabulary control (and scope notes) completely is not a solution to problematic indexing terms, and 6) that hybrid approaches, with multiple forms of vocabulary control and broader participation, would allow indexers and taggers to better work around problematic indexing terms.

4.4.2 Interrogating whiteness in subject indexing and subject tagging

To investigate how race, specifically white racial identity, affected the subject indexing and subject tagging processes and to understand the role of race dynamics in the relationship between vocabulary control and indexers' and taggers' feelings of control, this study coded participants' reactions and reflections that were about their racial identity and how that affected their process. This section presents cases that illustrate the themes identified in this analysis.

One theme that many participants expressed was "Feeling responsibility to represent racial and social issues equitably." This was true both for participants who identified as people of color and for participants who identified as white. While participants might have been motivated to assign accurate and equitable indexing terms in general, some participants specifically identified the subjects of the books as a reason to take particular care to do excellent work. For example, one participant, who identified as white, underlined both the importance and the challenges of thoughtfully representing documents about topics connected to racial and social justice: "You have to do this within a time constraint, but particularly when you're working with topics that affect people in really deep ways, catalogers have a responsibility to do that as carefully as possible." This attitude of care towards others, particularly those who are more likely to be harmed by problematic indexing terms, showed a form of user-orientation that went beyond trying to support users by anticipating how they might search for information, and extended to empathizing with the identities and histories of people of different races and from different backgrounds. Many participants noted that they were aware of the history of racial bias in indexing and cataloging and that they needed to "be aware of" or "control for" their own biases and biases in controlled vocabularies so they could index and tag in a way that showed care for people and works that had been harmed in the past.

Another theme that emerged, which was more specifically connected to white participants, was feeling that they were unqualified to represent books that were written from a Black perspective. Notably, this feeling was connected specifically to *Caste*, which was the only book in the study that was written by a Black author, and which was sympathetic to movements for racial and social justice. One participant, who identified as white, described how her racial identity affected her approach to indexing and tagging *Caste*: "I feel really hesitant when it comes to books like *Caste* because it is about being a person of color in a racist system, and I am not that, and so I struggle with being specific because I don't actually always know what terms are the most helpful or authentic to the book." She tagged *Caste* using minimal vocabulary control and assigned only very broad terms, such as "Race," "Racism," and "Historical" and felt that she might have been more specific if she had the support of a more rigid form of vocabulary control. In contrast, another participant said that she was confident while tagging *White Fragility* because she is white and felt that she could more authentically represent the view of the book and its author based on her lived experience.

Another participant also considered how her identity as a white woman related to the book *Caste*, focusing on how her identity related to the people who might use her indexing terms to find *Caste*: "It's really weird as a white person to think: How would people look for this book? How would white people search for this?" By focusing on information seekers (particularly white information seekers) this participant distanced herself from the Black identity and lived experience of Isabel Wilkerson, which was harder to authentically relate to, and focused on a relationship with information seekers where she could feel more authentic. She assigned relatively few tags for minimal vocabulary control (count=4), and those tags were a topic that was getting a lot of attention in public discourse at the moment ("Critical Race Theory") and broad terms that came directly from the book's description: "Caste system," "United States History,'" and "Race and racism," suggesting that she struggled identify with the book and that a form of vocabulary control without any suggestions or guidance was a poor match for a difficult task about which she expressed reservations. Another participant, who identified as Asian and also used

minimal vocabulary control to tag *Caste*, said: "Because I am unsure about this book, I am going to play it as safe as possible.... I'm uncomfortable to add any more" and assigned a sparse set of tags that lacked specificity.

Based on analysis of themes related to whiteness in subject indexing and subject tagging, this study found: 1) that indexers and taggers feel additional responsibility to assign specific, exhaustive, and accurate indexing terms to books "that affect people in really deep ways," 2) that novice indexers are aware of the history of bias in indexing and aim to control their own biases so they can counteract that history and show care and empathy to historically marginalized groups, and 3) that white indexers and taggers hold reservations about how authentically they can represent books that are presented from a different lived experience than their own, particularly when they sympathize with the viewpoint of the book and particularly when they lack guidance from a controlled vocabulary. These reservations tend to manifest in fewer and less specific indexing terms, particularly when novice indexers are not working with rigid or flexible vocabulary control.

5.0 Conclusion

Motivated by criticisms of controlled vocabularies and the emergence of new forms of vocabulary control, this study detailed five forms of vocabulary control and, by observing 18 novice indexers, examined how vocabulary control affects subject indexers' and taggers' feelings of control, the coextensiveness of the aboutness statements and subject representations that they create, and indexers' and taggers' ability to promote social and racial justice. The study found that different forms of vocabulary control have a significant effect on how novice indexers do their work, the indexing terms that they produce, and their ability to support racial and social justice goals. This study recommends that instead of simply dismissing other approaches as "uncontrolled" vocabularies, editors of more rigid controlled vocabularies explore how theseother forms of vocabulary control (and hybrid forms of vocabulary control) can promote racial and social justice and support indexers.

References

Adler, Melissa. 2009. "Transcending Library Catalogs: A Comparative Study of Controlled Terms in Library of Congress Subject Headings and User-Generated Tags in LibraryThing for Transgender Books." *Journal of Web Librarianship* 3, no. 4: 309-331.https://doi.org/10.1080/19322900903341099

Adler, Melissa, and Joseph T. Tennis. 2013. "Toward a Taxonomy of Harm in Knowledge Organization Systems." *Knowledge Organization* 40, no. 4: 266-72. https://doi.org/10.5771/0943-7444-2013-4-266

Adler, Melissa. 2017. "Classification Along the Color Line: Excavating Racism in the Stacks." *Journal of Critical Library and Information Studies* 1, no. 1. https://doi.org/10.24242/jclis.v1i1.17

Baron, Jill E., and Tina Gross. 2021. "Sorrow, Fury, Helplessness, and Cynicism: An Account of the Library of Congress Subject Heading 'Illegal aliens.'" *Dartmouth Library Staff Publications* 34. https://digitalcommons.dartmouth.edu/dlstaffpubs/34.

Bates, Jo and Jennifer Rowley. 2011. "Social Reproduction and Exclusion In Subject Indexing: A Comparison Of Public Library Opacs And Librarything Folksonomy." *Journal of Documentation* 67, no. 3: 431-48. https://doi.org/10.1108/00220411111124532

Berman, Sanford. 1971. *Prejudices and Antipathies*. Metuchen, NJ: The Scarecrow Press.

Berman, Sanford and Tina Gross. 2017. "Expand, Humanize, Simplify: An Interview with Sandy Berman." *Cataloging & Classification Quarterly* 55, no. 6: 347-360.

Bertrand, Annick, and Jean-Marie Cellier. 1995. "Psychological Approach to Indexing: Effects of the Operator's Expertise Upon Indexing Behaviour." *Journal of Information Science* 21, no. 6: 459-472. https://doi.org/10.1177/016555159502100605

Bowker, Geoffrey C. and Susan Leigh Star. 1996. "How Things Actor-net(work): Classification, Magic, and the Ubiquity of Standards." *Philosophia* 25, no 2-4: 195-220.

Bowker, Geoffrey C. and Susan Leigh Star. 2000. *Sorting Things Out: Classification and its Consequences*. MIT Press.

Bullard, Julia. 2018. "Curated Folksonomies: Three Implementations of Structure Through Human Judgment." *Knowledge Organization* 45 no. 8: 643-652. doi.org/10.5771/0943-7444-2018-8-643

Bullard, Julia, Brian Watson, and Caitlin Purdome. 2022. "Misrepresentation in the Surrogate: Author Critiques of 'Indians of North America' Subject Headings'." *Cataloging & Classification Quarterly* 60 no. 6-7, 599-619.

Charters, Elizabeth. 2003. "The Use of Think-aloud Methods in Qualitative Research: An Introduction to Think-aloud Methods." *Brock Education Journal* 12, no. 2: 68-82.

Chu, Clara M., and Ann O'Brien. 1993. "Subject analysis: the Critical First Stage of Indexing." *Journal of Information Science* 19, no. 6: 439-454. https://doi.org/10.1177/016555159301900603

Cifor, Marika and K.J. Rawson. 2022. "Mediating Queer and Trans Pasts: The Homosaurus as Queer Information Activism." *Information, Communication & Society* 26, no. 11 2168-2185.https://doi.org/10.1080/1369118X.2022.2072753

Cocq, Coppélle. 2015. "Indigenous Voices on the Web: Folksonomies of Endangered Languages" *The Journal of American Folklore* 128, no. 509: 273-285.

Duart, Marisa Elena, and Miranda Belarde-Lewis. "Imaging: Creating Spaces for Indigenous Ontologies." *Cataloging & Classification Quarterly* 53, no. 5-6: 677-702.

Ericsson, K. Anders, and Herbert Alexander Simon. 1984. *Protocol Analysis: Verbal Reports as Data*. MIT Press.

Farrow, John F. 1991. "A Cognitive Process Model of Document Indexing." *Journal of Documentation* 47, no. 2: 149-166.https://doi.org/10.1108/eb026875

Feinberg, Melanie. 2006. "An Examination of Authority Control in Social Classification Systems." *Advances in Classification Research* 17: 1-11. https://doi.org/10.7152/acro.v17i1.12490

Fidel, Raya. 1992. "Who Needs Controlled Vocabulary?" *Special Libraries* 83, no.1: 1-9.

Fox, Melodie J. 2016. "'Priorities of Arrangement?' or a 'Hierarchy of Oppressions?': Perspectives on Intersectionality in Knowledge Organization." *Knowledge Organization* 43, no. 5: 373-383. doi.org/10.5771/0943-7444-2016-5-373

Frohman, Bernd. 1990. "Rules of Indexing: A Critique of Mentalism in Information Retrieval Theory." *Journal of Documentation* 46, no. 2: 365-386.

Furner, Jonathan. 2007. "Dewey Deracialized: A Critical Race-Theoretic Perspective." *Knowledge Organization* 34, no. 3: 144–68. https://doi.org/10.5771/0943-7444-2007-3-144

Golder, Scott A., and Bernardo A. Huberman. 2005. "The Structure of Collaborative Tagging Systems." Preprint, submitted August 18, 2005. *arXiv preprint cs/0508082*.

Gross, Tina, Arlene G. Taylor, and Daniel N. Joudrey. 2015 "Still a Lot to Lose: The Role of Controlled Vocabulary in Keyword Searching." *Cataloging & Classification Quarterly* 53, no. 1: 1-39. https://doi.org/10.1080/01639374.2014.917447

Halpin, Harry, Valentin Robu, and Hana Shephard. "The Complex Dynamics of Collaborative Tagging." In *Proceedings of the 16th International Conference on the World Wide Web (WWW '07)*, New York, ACM: 211-220.

Holstrom, Chris. 2020. "The Effects of Suggested Tags and Autocomplete Features on Social Tagging Behaviors." In *Proceedings of the Association for Information Science and Technology* 57, no 1. e263.

Holstrom, Chris. 2021. "The Sears List of Subject Headings: Social and Cultural Dimensions in Historical, Theoretical, and Design Contexts." In *Proceedings NASKO 2021*: 1-11.

Holstrom, Chris. 2022. "Analyzing the Structure and Dynamics of Control Relationships in the Case of 'Illegal Aliens' in the Library of Congress Subject Headings." *Knowledge Organization Across Disciplines, Domains, Services, and Technologies*: *Proceedings of the Seventeenth International ISKO Conference 6–8 July 2022 Aalborg, Denmark*, edited by Marianne Lykke et al. Advances in Knowledge Organization 19. Würzburg: Ergon Verlag, 133-146.

Howard, Sara A. and Steven A. Knowlton. 2018. "Browsing Through Bias: The Library of Congress Classification and Subject Headings for African American Studies and LGBTQIA Studies." *Library Trends*, 67, no. 1: 74-88.

ISO. 1985. "Documentation–methods for Examining Documents, Determining Their Subjects And Selecting Indexing Terms." *International Organization for Standardization, Geneva 5963*.

Jackson, Corey, Kevin Crowston, Carsten Østerlund, and Mahboobeh Harandi. 2018. "Folksonomies to Support Coordination of and Coordination of Folksonomies." *Computer Supported Cooperative Work (CSCW)* no. 27: 647-678.https://doi.org/10.1007/s10606-018-9327-z

Joudrey, Daniel. 2005. "Building Puzzles and Growing Pearls: A Qualitative Exploration of Determining Aboutness." PhD diss., University of Pittsburgh.

Kipp, Margaret E.I. and D. Grant Campbell. 2006. "Patterns and Inconsistencies in Collaborative Tagging Systems: An Examination of Tagging Practices." *Proceedings of the American Society for Information Science and Technology* 43, no. 1: 1-18. https://doi.org/10.1002/meet.14504301178

Lancaster, F.W. 1972. *Vocabulary Control for Information Retrieval*. Washington: Information Resources Press.

Lancaster, F.W. 2003. *Indexing and Abstracting in Theory and Practice*, 3rd ed. University of Illinois, Champaign, Ill.

Langridge, D.W. 1989. *Subject Analysis: Principles and Procedures*. London; New York: Bowker-Saur.

Littletree, Sandra and Cheryl A. Metoyer. 2015. "Knowledge Organization from an Indigenous Perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology Project." *Cataloging & Classification Quarterly* 53, no. 5-6: 640-657.

Lo, Grace. 2019. "'Aliens' vs. Cataloger: Bias in the Library of Congress Subject Heading." *Legal Reference Services Quarterly* 38, no.4: 170-196.

Mai, Jens-Erik. 1999. "Deconstructing the Indexing Process." *Advances in Librarianship* 23, 269-298. https://doi.org/10.1108/S0065-2830(1999)0000023013

Mathes, Adam. 2004. "Folksonomies - Cooperative Classification and Communication Through Shared Metadata." https://adammathes.com/academic/computer-mediated-communication/folksonomies.html

Miksa, Francis L. 1983. *The Subject in the Dictionary Catalog from Cutter to Present*. Chicago: American Library Association.

Olson, Hope. 2001. "The Power to Name: Representation in Library Catalogs." *Signs: Journal of Women in Culture and Society* 26, no 3: 639-68. https://doi.org/10.1086/495624

Quintarelli, Emanuele. 2005. "Folksonomies: Power to the People." In *ISKO UniMIB Meeting*, Milan, June 24, 2005.

Rosa, Kathy, and Kelsey Henke. 2017. *2017 ALA Demographic Study*. American Library Association, Office of Research and Statistics.

Rowley, Jennifer. 1994. "The Controlled Versus Natural Indexing Languages Debate Revisited: A Perspective on Information Retrieval Practice And Research." *Journal of Information Science* 20, no. 2: 108-118. https://doi.org/10.1177/016555159402000204

Šauperl, Alenka. 1999. "Subject Determination During the Cataloging Process." PhD diss., The University of North Carolina at Chapel Hill.

Smiraglia, Richard. 1991. "Subject Access to Archival Materials Using LCSH." In *Describing Archival Materials*: *The Use of the MARC AMC Format*. Routledge, 63-90.

Soergel, Dagobert. 1985. Organizing Information: Principles of Data Base and Retrieval Systems. Orlando: Academic Press, Inc.

Stack Overflow. n.d. Tag Synonyms. https://stackoverflow.com/tags/synonyms. Accessed on June 29, 2024.

Svenonius, Elaine. 2000. The Intellectual Foundation of Information Organization. Cambridge: MIT Press.

Tennis, Joseph T. 2006. "Social Tagging and the Next Steps for Indexing." *Advances in Classification Research* 17.

Tennis, Joseph T. 2012. "The Strange Case of Eugenics: A Subject's Ontogeny in a Long-Lived Classification Scheme and the Question of Collocative Integrity." *Journal of the American Society for Information Science and Technology* 63, no. 7: 1350-1359. https://doi.org/10.1002/asi.2268

Wilson, H.W. 2022. *The Sears List of Subject Headings*, 23rd ed. H.W. Wilson Company.

Wilson, Patrick. 1968. *Two Kinds of Power: An Essay on Bibliographical Control*. University of California Press.