

©Copyright 2023

Chris Holstrom

Controlled, In Control, and Out of Control:
The Effects of Different Forms of Vocabulary Control on the Subject
Indexing and Subject Tagging Processes

Chris Holstrom

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2023

Reading Committee:

Joseph T. Tennis, Chair

Nicholas Weber

Marika Cifor

Program Authorized to Offer Degree:

The Information School

University of Washington

Abstract

Controlled, In Control, and Out of Control:
The Effects of Different Forms of Vocabulary Control on the Subject Indexing and Subject
Tagging Processes

Chris Holstrom

Chair of the Supervisory Committee:
Joseph T. Tennis

The Information School

Subject indexing, the process of determining what a document is about and then translating that “aboutness” into a representation in an indexing language, is fundamental to library cataloging. Subject indexing has historically relied on controlled vocabularies to standardize language and to optimize information retrieval; however, controlled vocabularies have faced criticism for reinforcing systemic biases, for being slow to adapt, and for not reflecting the language and values of many people, especially those from marginalized and underrepresented groups. The emergence of folksonomies in the Web 2.0 era presented a clear contrast to rigorous vocabulary control and promised a more “democratic” form of knowledge organization that could reflect natural language, adapt quickly, and potentially avoid problems with systemic bias. Folksonomies and social tagging were criticized, though, for being too “uncontrolled” and, therefore, less effective as information retrieval indexes.

This study updates and enriches this comparison of vocabulary control in subject indexing and tagging 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 indexers and taggers. The study used 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 explored how controlled vocabularies made participants feel controlled, in control, or out of control while they indexed and tagged; measured how different forms of vocabulary control affected coextensiveness between aboutness statements and indexing terms; and analyzed how these experiences inform the criticisms of controlled vocabularies, particularly for documents about racial and social justice. The study found that the form of vocabulary control had 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 navigated complex issues related to racial and social justice. These findings should help the designers and editors of knowledge organization systems and controlled vocabularies to better and more equitably serve indexers, taggers, information seekers, and the information being indexed and tagged.

TABLE OF CONTENTS

	Page
List of Figures	v
Glossary	vi
Chapter 1: Introduction	1
Chapter 2: Literature Review	7
2.1 Subject Analysis	8
2.2 Subject Indexing	10
2.2.1 The Steps in the Subject Indexing Process	11
2.2.2 Specificity, Exhaustivity, and Accuracy in Subject Indexing	14
2.2.3 Different Orientations in Subject Indexing	17
2.3 Vocabulary Control and Criticisms of Vocabulary Control	19
2.3.1 Criticism: Controlled Vocabularies Reinforce Systemic Biases	20
2.3.2 Criticism: Controlled Vocabularies Are Too Slow to Adapt	22
2.3.3 Criticism: Controlled Vocabularies Do Not Reflect the Language of Information Seekers and Many Groups and Cultures	23
2.4 Tagging and “Uncontrolled” Vocabularies	24
2.5 “Subject Tagging” as a Specific Type of Tagging	26
2.6 Five Forms of Vocabulary Control	29
2.6.1 Minimal Vocabulary Control	29
2.6.2 Passive Vocabulary Control	30
2.6.3 Post Hoc Vocabulary Control	31
2.6.4 Flexible Vocabulary Control	31
2.6.5 Rigid Vocabulary Control	32
2.7 Forms of Vocabulary Control, Control Relationships, and Systemic Bias	33
2.7.1 Vocabulary Control and Bibliographical Control	34
2.7.2 Controlled Vocabularies, Systemic Bias, and Control of Information Seekers	35
2.7.3 Rigidly Controlled Vocabularies, Systemic Bias, and Control of Indexers	36
2.7.4 Control Relationships and Less Rigid Forms of Vocabulary Control	39
2.7.5 Controlled, In Control, and Out of Control	41

2.7.6	Control and the Ease of Use of Indexing Languages	42
2.8	Empirical Studies of the Subject Analysis and Subject Indexing Processes	43
2.8.1	Chu & O'Brien: What Document Properties Influence Subject Analysis?	44
2.8.2	Joudrey: How Do Participants Determine Aboutness of Documents?	45
2.8.3	Šaupperl: How Do Catalogers Determine Subject Descriptions?	46
2.8.4	Bertrand & Cellier: How Does Expertise Affect Indexing and What Transformations Occur During Translation?	47
Chapter 3: Research Questions		
3.1	Gaps in Existing Literature	49
3.2	Overarching Research Question	50
3.3	Three Research Sub-Questions	50
3.4	Impact: Why These Research Questions Matter	52
Chapter 4: Methodology		
4.1	Research Paradigm	54
4.1.1	My Ontological and Epistemological Assumptions	54
4.1.2	Dewey's Instrumentalist Pragmatism	55
4.2	Key Questions that Inform the Study Design	60
4.2.1	Which Documents Should Participants Index and Tag?	60
4.2.2	Is It Ethical to Ask Participants to Tag and Index Books About Charged Topics Like Racial and Social Justice and to Share Their Feelings About the Experience?	66
4.2.3	How Many Participants Should the Study Observe?	69
4.2.4	Should the Study Separate Subject Analysis and Subject Representation?	71
4.2.5	Should Participants Be Experienced with Indexing and Tagging?	72
4.2.6	Should Participant Demographics Reflect U.S. Librarian Demographics?	73
4.3	Methods	76
4.3.1	Experimental Design	76
4.3.2	Participants, Sampling Methods, and Recruiting	79
4.3.3	Experimental Materials and Apparatus	89
4.3.4	Procedures	94
4.3.5	Data Collection	97
4.4	Threats to Validity and Limitations of the Study	102
4.4.1	Threats to Internal Validity	102
4.4.2	Threats to External and Temporal Validity	104
Chapter 5: Results and Discussion		
5.1	Feeling Controlled, In Control, and Out of Control	109

5.1.1	Feeling Most In Control to Feeling Most Controlled	110
5.1.2	Feeling Most In Control to Feeling Most Out Of Control	115
5.1.3	Easiest to Use to Most Difficult to Use	123
5.1.4	Specificity, Exhaustivity, Accuracy	130
5.2	Coextensiveness of Aboutness Statements and Indexing Terms	140
5.2.1	Coextensiveness and Types of Transformations	141
5.2.2	Analysis and Discussion of Transformations from Aboutness State- ments to Indexing Terms	145
5.2.3	Analysis and Discussion of Participants' Self-Reported Degree of Co- extensiveness	161
5.2.4	Analysis and Discussion of Bibliographic Apparatus and Coextensive- ness	171
5.3	Vocabulary Control, Critical Cataloging, and Racial and Social Justice	175
5.3.1	Working with and around Problematic and Missing Indexing Terms	176
5.3.2	Interrogating Whiteness in Subject Indexing and Subject Tagging	181
5.3.3	Indexing and Tagging Books that are Critical of the Conversations around Racial and Social Justice	184
Chapter 6:	Conclusion	189
6.1	Forms of Vocabulary Control and Feelings of Control	190
6.2	Forms of Vocabulary Control and Coextensiveness	192
6.3	Racial and Social Justice in Subject Indexing and Subject Tagging	196
6.4	Validity and Limitations	197
6.5	The Role of Humans in Knowledge Organization	198
6.6	Closing Words	199
References		202
Appendix A:	Recruiting Emails	212
Appendix B:	Demographic Survey	215
Appendix C:	Script to Introduce the Study to Participants	218
Appendix D:	Consent Form	221
Appendix E:	Script to Introduce Minimal Control (Free Tagging)	224
Appendix F:	Script to Introduce Passive Control (Autocomplete Tags Enabled)	225
Appendix G:	Script to Introduce Flexible Control (<i>Sears List</i>)	226

Appendix H: Script to Introduce Rigid Control (<i>LCSH</i>)	227
Appendix I: Script for When Participants Finished Indexing or Tagging a Book	228
Appendix J: Interview Questions	229
Appendix K: Codebook	233
K.1 Codebook for Participant Responses to Interview Questions About Coexten- siveness	233
K.2 Codebook for Bibliographic Apparatus that Participants Used to Find About- ness Information	235
K.3 Codebook for Participant Reactions, Responses, and Behaviors	237
Appendix L: ChatGPT Responses to Subject Indexing and Subject Tagging Prompts	241

LIST OF FIGURES

Figure Number	Page
2.1 The Subject Indexing Process	11
2.2 Examples of Translation into Different Indexing Languages	13
2.3 Annotations in a copy of the <i>Sears List</i>	40
2.4 The Two Axes of Control	41
4.1 Dewey’s Iterative Model of Inquiry	59
4.2 Racial and Ethnic Identities of Participants	82
4.3 Gender Identities of Participants	83
4.4 Ages of Participants	84
4.5 Participant Education in Knowledge Organization	85
4.6 Participant Education in Cataloging	86
4.7 Participant Experience Level in Indexing and Tagging	87
4.8 Participant Familiarity with Racial and Social Justice Literature	89
4.9 Study Materials	90
4.10 Tagging and Indexing User Interface	93
4.11 Administrator User Interface	94
4.12 Example Aboutness Statement	99
4.13 Ranking Note Cards	101
5.1 Bimodal Distribution of In Control and Out of Control Rankings for Minimal Vocabulary Control	117
5.2 Sankey Diagrams Summarizing Transformations	148
5.3 Streamlined Aboutness Statement	169
L.1 ChatGPT Aboutness Statement for <i>Caste</i>	241
L.2 ChatGPT Library of Congress Subject Headings for <i>Caste</i>	242
L.3 ChatGPT Library of Congress Subject Headings for <i>Caste</i>	243
L.4 ChatGPT Sears List Subject Headings for <i>White Fragility</i>	244
L.5 ChatGPT Tags for <i>Mismatch</i>	245
L.6 ChatGPT Tags for <i>The End of Racism</i>	246

GLOSSARY

ABOUTNESS STATEMENT: The output of the subject analysis process (see below), an aboutness statement is a summary of what the subject indexer or tagger determines the document to be about. Aboutness statements are typically short summaries or lists of key topics.

CATALOGING: The process of assigning metadata to documents in a bibliographic record in a library catalog. Subject indexing is just one aspect of cataloging, which also involves other types of metadata such as author, title, publication date, ISBN, and so forth.

COEXTENSIVENESS: In subject indexing and tagging, the degree to which the aboutness statement produced in the subject analysis step matches the subject representation expressed as subject headings or tags at the end of the subject indexing or subject tagging process. Note that some definitions of coextensiveness focus on the degree to which subject representation matches the topic of the resource. While the definition in this study aligns with these other definitions, this study focuses on and analyzes the differences and similarities between the aboutness statement and the subject representation in subject headings and tags.

CONTROL: The power to influence, limit, or direct the behavior or experiences of oneself or others. Control often manifests as a relationship between actors who are *in control* and actors who are *controlled* by those actors or by the systems and artifacts that they create. Alternatively, control can manifest as actors feeling *out of control* because they lack the knowledge or context to guide their own actions. This study argues that these feelings of control—*controlled*, *in control*, and *out of control*—are affected by

the type of vocabulary control being used.

CONTROLLED VOCABULARY: Similar to an indexing language, a controlled vocabulary is a collection of authorized indexing terms and the rules that govern their use. Controlled vocabularies aim to standardize language and avoid ambiguity and confusion. See also: “Vocabulary Control” and “Indexing Language.”

DOCUMENT: Any information object that can be indexed or tagged, including a book, web page, journal article, video, or audio recording. In this study, a “document” and a “resource” are synonymous.

FOLKSONOMY: A portmanteau of “folks” and “taxonomy,” a folksonomy is a collection of tags and the Knowledge Organization System that they form in aggregate. Folksonomies are sometimes called “uncontrolled vocabularies” because they historically did not employ vocabulary control; however, this study distinguishes between folksonomies, which are an aggregation of tags that might or might not have some form of vocabulary control, and “uncontrolled vocabularies,” which are defined by the lack of restriction on vocabulary—not necessarily the use of tagging.

INDEXING LANGUAGE: A collection of indexing terms, authorized or not, and the rules that govern the use of these terms. Indexing languages are used to build individual indexes for specific collections. Examples of indexing languages include the *Sears List of Subject Headings*, which is a controlled vocabulary, and the MetaFilter folksonomy, which is not.

INDEXING TERM: Any word or pre-coordinated set of words that describes what a document is about, independent of the type of knowledge organization system. “Indexing term” is used in this study to refer to subject headings and tags collectively, and can refer to class names in a classification scheme too.

KNOWLEDGE ORGANIZATION SYSTEM (KOS): A broad term for many types of systems that organize and facilitate the retrieval of information, including classification schemes, ontologies, authority records, and—importantly for this study—subject heading lists and folksonomies.

SOCIAL TAGGING: See “Tagging.”

SUBJECT ANALYSIS: The process of analyzing a document to determine what it is about. The output of the subject analysis step is an “Aboutness statement” (see above). Subject analysis is typically followed by an independent step of “Subject representation” (see below) in both subject indexing and subject tagging. Some researchers use the term “Content Analysis” synonymously with subject analysis and some treat the terms as slightly different processes. For consistency, this study uses only subject analysis. Additionally, many researchers note that indexers and taggers often combine subject analysis and subject representation, so that they are not distinct steps. This study, and its research method, follow Langridge’s and Lancaster’s recommendation that these two steps remain separate [86, 85].

SUBJECT HEADING: An authorized indexing term in a subject heading list. Subject headings can be single nouns or compound indexing terms. Unless otherwise noted, “subject heading refers” to authorized (or preferred) terms, and does not include non-authorized entry terms.

SUBJECT HEADING LIST: A published list of authorized subject headings and rules for their use, typically intended for use in library cataloging. Examples of subject heading lists include the *Sears List of Subject Headings* and the *Library of Congress Subject Headings (LCSH)*.

SUBJECT INDEXING: The process of determining what a document is about (subject analysis) and then translating that “aboutness” into a representation in an indexing

language (subject representation).

SUBJECT REPRESENTATION: The process of translating the “aboutness” of a document, previously determined in the subject analysis step, into an indexing term or terms in an indexing language. These indexing terms then represent the subject of the document and are used to aid retrieval by subject.

SUBJECT TAGGING: The process of determining what a document is about (subject analysis) and then representing that “aboutness” with keywords or tags. Subject tagging is typically done on the web by members of online communities and is one of many types of tagging strategies collectively known as “social tagging” or just “tagging.”

TAGGING: The process of assigning keywords or “tags” to a document or resource. Tagging is typically done on the web by members of online communities and can involve many strategies, including tagging to manage tasks, to express judgments, and to identify the type of media. An aggregation of tags that is used as a Knowledge Organization System is called a “Folksonomy” (see above).

TRANSLATION: The process of mapping an aboutness statement into indexing terms from an indexing language. Some models of the subject indexing process treat translation and subject representation as a single, combined step while others treat translation as a separate step that precedes representation. This study analyzes the effect of different forms of vocabulary control on the translation process by studying “Co-extensiveness” (see above).

UNCONTROLLED VOCABULARY: A freely growing collection of indexing terms with no restrictions on which indexing terms can be used or added to the vocabulary. Every time an indexer or tagger uses a new indexing term—even a typo or a synonym of a term that is already in the vocabulary—the uncontrolled vocabulary grows. “Folksonomies” are sometimes called uncontrolled vocabularies because they historically

did not employ vocabulary control; however, this study distinguishes between folksonomies, which are an aggregation of tags that might or might not have some form of vocabulary control, and uncontrolled vocabularies, which are defined by the lack of restriction on vocabulary—not necessarily their use of tagging.

VOCABULARY CONTROL: A system for standardizing indexing terms to control for synonyms and homographs and to define the relationships between terms. This study argues that vocabulary control is not binary (not just controlled or uncontrolled) but that multiple forms of vocabulary control exist. See also: “Controlled Vocabulary.”

ACKNOWLEDGMENTS

I acknowledge that I conduct my research on the unceded, traditional land of the Coast Salish People, including the Duwamish People past and present. I honor with gratitude the land itself and the Duwamish Tribe.

DEDICATION

To my amazing and inquisitive daughters, Gwen and Tova. Gwen taught me that Taylor Swift is the key to teaching Knowledge Organization and Tova enthusiastically joined me in combing through past editions of the *Sears List of Subject Headings* and encouraged me to keep writing. Both girls bring hope and joy to my days and ask me lots of great questions about subject indexing and tagging.

Chapter 1

INTRODUCTION

Suppose that you want to find books in your local library about long-distance running. You can search the library's catalog for "long-distance running," and the catalog will likely return a set of relevant results about the subject. My local library returns results that include the memoir *Spirit Run: A 6,000-Mile Marathon Through North America's Stolen Land* and the training handbook *The Long Distance Runner's Guide to Injury Prevention and Treatment*. As a library patron and information seeker, I am happy to see these books and a dozen more about long-distance running made available in one convenient search result.

This type of search and retrieval experience is, of course, no accident. It is the result of work, performed by librarians and catalogers, that is called subject indexing. Subject indexing is the process of determining what a document is about and then translating that "aboutness" into a representation in an indexing language. So, when my local library acquired *Spirit Run*, a cataloger might have analyzed the subtitle, table of contents, introduction, and other parts of the book to determine what it was about. Then, the cataloger might have consulted an indexing language like the *Library of Congress Subject Headings (LCSH)* and translated their understanding of what the book was about into a set of authorized indexing terms or subject headings. In this case, my local library chose five authorized subject headings, including "Long-distance running--North America--Anecdotes" and "Indians of North America--Biography." These subject headings not only support search but also serve a navigation function because I can click on "Indians of North America--Biography" and find other works about that subject.

Importantly, these subject headings also illustrate another concept critical to library catalogs: vocabulary control. Vocabulary control is a system for standardizing indexing terms, such as subject headings, to control for synonyms and homographs and to define the rela-

tionships between terms. Because of vocabulary control, I can find all the books about long-distance running under one subject—the *LCSH* authorized subject heading “long-distance running”—instead of finding some books under “long-distance running” and others under an unauthorized synonym like “high-mileage running.” Vocabulary control for some indexing languages also tells catalogers how to represent complex subjects by using subdivisions. For example, the authorized heading “Indians of North America—Biography” represents a complex subject with multiple parts, and the way that those complex parts are combined is governed by the rules of the *LCSH*.

Finally, and most importantly to this study, vocabulary control limits the indexing terms or subject headings that catalogers can choose to a set of authorized terms—and those authorized terms have often proven to be problematic. For example, the subject heading “Indians of North America—Biography” uses the problematic phrase “Indians of North America” [13], but catalogers who want to use the *LCSH* and benefit from its controlled vocabulary must use that problematic subject heading. This study aims to study this specific situation—when catalogers might feel controlled by the indexing languages that they are using.

*

Now, suppose that you want to find posts about science fiction on a community website. On the community blog MetaFilter, you can search for or click on the tag “sciencefiction” and find posts about science fiction, such as “Hugo Award finalists include a story in tweeted images” [15] and “Can Science Fiction Wake Us Up to Our Climate Reality?” [139]. As an information seeker on the web, you would likely be happy to find these science fiction posts collected together, much like I was happy to find books about long-distance running in a convenient search result at my local library. As in that long-distance running example, this retrieval experience is the result of work that was performed, in this case, by taggers who tagged posts about science fiction with the tag “sciencefiction.”

You might, however, find some inconsistencies as you explore posts and tags on MetaFilter. You might notice that, while most science fiction posts are tagged “sciencefiction,” a

different set of posts are tagged “scifi” and others are tagged “SF.” Because the tagging vocabulary does not require taggers to use a single authorized tag, some posts are likely to be scattered in small, disjoint collections defined by just one of these tags. You might also note further complications due to the lack of vocabulary control when you investigate the tag “SF,” which is used as shorthand for science fiction but is also used to tag posts about the city of San Francisco. (San Francisco itself has many tags associated with it, including “SanFrancisco” and “San” and “Francisco” as separate tags.) The lack of vocabulary control in many social tagging systems and the resultant proliferation of orthographic variants—including typos like “sceince fiction,” different spacing like “sciencefiction” and “science fiction,” different punctuation like “science-fiction” and “science.fiction,” and nearly synonymous tags like “scifi”—have inspired much criticism [81, 97, 110] and can result in inferior recall during information retrieval [91].

However, tagging continues to be a key indexing strategy across the web, and tagging is used even in more traditional information ecosystems like library catalogs and academic journals [25, 51]. Why would someone use an “uncontrolled” vocabulary if it cannot differentiate between a major US West Coast city and a futuristic genre of film and literature? The answer lies in how tags are created and who creates them. Tagging benefits from free labor performed by people who want to collect and organize information for themselves and for their communities. That labor is free (or very cheap) because taggers, on average, are not as skilled as catalogers and because the barrier to entry for tagging is extremely low—just enter some tags, no training or specialized tools required. Because of the lower barrier to entry, tagging has been praised as not just cheaper but also more accessible and more democratic than the traditional indexing done in library catalogs [94]. Tagging vocabularies are more likely to reflect common usage [97] and incorporate new terminology more rapidly [82]. For example, a tagger would not feel compelled like a cataloger to use a tag like the *LCSH* subject heading “Indians of North of America” and could instead use “Native Americans” or “American Indians” or any tag that they felt was descriptive and appropriate. Additionally, the flexibility of tagging provides historically marginalized groups with a more visibility [70]. For example, many of the posts tagged with “sciencefiction” on MetaFilter are also tagged with “postsbypoc.” This emergent tag, which raises the visibility

of posts made by people of color in the MetaFilter community, lacks an analog in traditional subject indexing and shows the flexibility and power of tagging.

*

These two examples set the context for this study and highlight some of the key similarities and differences between subject indexing and tagging, but they do not tell the whole story—especially about vocabulary control. This study argues that vocabulary control is not a binary feature of knowledge organization systems (KOS), that vocabulary control is not either rigorously enforced (as in the long-distance running and *LCSH* example) or completely absent (as in the science fiction tagging example). Instead, many KOS have some form of vocabulary control that falls between those extremes. For example, some subject indexing languages are more flexible about vocabulary control than *LCSH*, like the *Sears List of Subject Headings*, which specifically encourages catalogers to add their own subject headings if the *Sears List* does not provide sufficiently specific headings. And some tagging systems, like the system on the question-and-answer website Stack Overflow, employ forms of vocabulary control in which trusted community members curate community-provided tags to ensure consistency [20].

This study describes five forms of vocabulary control that are used in subject indexing and subject tagging, and then proposes a mixed-methods study to understand the effect of these forms of vocabulary control on the subject indexing and subject tagging processes. The study emphasizes the people who do the subject indexing and subject tagging work that was described in the examples above and considers the connection between these experiences and social equity. The study contains the following chapters:

- **Chapter 2** is a review of literature about subject analysis, subject indexing, tagging, and vocabulary control. The chapter starts by exploring the steps involved in the subject indexing process, describing how accuracy and specificity and exhaustivity are key to achieving the objectives of subject indexing, and analyzing different orientations in subject indexing, such as document-oriented and user-oriented approaches. Next,

the chapter reviews the role of vocabulary control in subject indexing and discusses criticisms of controlled vocabularies. Third, the chapter reviews literature that compares and contrasts subject indexing and tagging, and defines “subject tagging” as a subset of tagging that is most directly compared to subject indexing in this study. Next, the chapter describes five different forms of vocabulary control and considers their similarities and differences. Then the chapter explores the relationship between these forms of vocabulary control, control relationships in KOS, and racial and social justice. Finally, the chapter reviews in detail four studies that directly observe subject analysis and subject indexing work. These studies serve as inspiration for the methods used in this study, and the questions that these studies leave unanswered motivate the research questions presented in the following chapter.

- **Chapter 3** is a short chapter that reviews questions that remain unanswered in the existing literature and presents these questions as the research questions at the center of this study. The chapter then argues why these research questions matter in the current subject indexing and tagging landscapes.
- **Chapter 4** covers the methodology that was used to answer these research questions. The chapter begins by describing how a Pragmatist research paradigm aligns with my ontological and epistemological views and fits the research questions described in this study. Next, taking cues from Deweyan Pragmatism, the chapter discusses a series of research design questions and the rationale for my research design. Third, the chapter discusses limitations to the study design. Next, the chapter describes the participants, experimental materials, and procedures for the study. Finally, the chapter details the methods for collecting data and the types of data that were collected.
- **Chapter 5** analyzes and discusses the results from the study, focusing on the effect of vocabulary control on participants’ experiences with the indexing and tagging processes, the transformations that occurred when participants translated their subject analysis into indexing languages, and the connections between critical cataloging literature and participants’ experience indexing and tagging books related to racial and

social justice.

- **Chapter 6** concludes by discussing how designers of knowledge organization systems and controlled vocabularies can better and more equitably serve indexers, taggers, information seekers, and the information being indexed and tagged.

Chapter 2

LITERATURE REVIEW

This chapter expands on the discussions, started in the Introduction, about subject indexing, tagging, and vocabulary control. This expansion begins with a review of the literature about the first step in subject indexing and subject tagging: subject analysis. Next, the section continues by discussing the steps in the subject indexing process and the role of vocabulary control in those steps. It continues by detailing three goals of subject indexing—accuracy, specificity, and exhaustivity—and how they relate to vocabulary control and to different orientations or approaches to subject indexing.

This discussion of subject indexing leads into a review of criticisms of vocabulary control and how controlled vocabularies can perpetuate systemic bias. The subject indexing and tagging processes and criticisms of controlled vocabularies were central to this study, which aimed to better understand these processes from the perspectives of indexers and taggers, to understand the effect of vocabulary on these processes, and to see how these effects inform criticisms of controlled vocabularies.

Then, the chapter introduces tagging as an alternative to subject indexing and reviews literature that compares subject indexing and tagging—defining the term “subject tagging” to make a direct comparison between these approaches. The chapter presents a set of five forms of vocabulary control that are in active use today in Knowledge Organization Systems (KOS). These forms of vocabulary control are described in detail to because the study is designed around comparing the effects that these five forms of vocabulary control have on the subject indexing and subject tagging processes.

Next, the chapter explores vocabulary control, control relationships, and racial and social. This section examines different types of control that indexers and taggers exert over documents and information seekers and the different types of feelings of control that indexers and taggers experience while indexing and tagging documents. The section then considers

how these control relationships, which are reified by vocabulary control, can perpetuate racial and social bias or can promote racial and social justice. The section establishes these connections between vocabulary control, control relationships, and racial and social justice to motivate the design of the study, both broadly and in details such as the subject domain of the documents indexed and tagged in the study.

Finally, the chapter reviews four studies of the subject analysis and subject indexing processes. These four studies directly inform the research questions and methodology of the current study. These four studies also represent a rarity in research about subject analysis and subject representation because they directly observe people while they do subject analysis and subject indexing work, instead of just analyzing the output of their work or interviewing them about their processes. This study takes inspiration from the research methods employed in these four studies.

2.1 Subject Analysis

D.W. Langridge describes subject analysis as “the first, the most important and the most difficult part of all classification and indexing” [86]. This critical first step involves analyzing a document to understand what it is about and producing an aboutness statement that can be used in subsequent steps in the subject indexing or subject tagging processes. Despite its fundamental importance, subject analysis is somewhat poorly understood. W. John Hutchins laments that while “[w]e find a great deal about the construction of indexing languages and classification systems ... we find very little about how indexers and classifiers decide what the subject of the document is, how they decide what it is ‘about’” [76]. Frohman adds to Hutchins, stating “research focuses on the ... [final] step, while the first continues to be lamented as an intellectual operation both fundamental to indexing yet so far resistant to analysis” [48]. *ISO Standard 5963: Documentation — Methods for examining documents, determining their subjects and selecting indexing terms* [77] follows this same pattern, as its flowchart summarizing the indexing process dedicates only one of 19 steps to subject analysis: “Examine document and identify significant concepts.”

One reason for this pattern of under-specifying the subject analysis process is due to a lack of studies that observe and describe the actions and thought processes of people

actually doing subject analysis work. The last section of this chapter discusses four studies that begin to fill in the gaps that Hutchins, Frohman, and others describe by observing subject analysis in action. The current study, while not primarily focused on detailing the subject analysis process, does observe participants doing subject analysis work and reports on their processes to provide context for their subject indexing and tagging processes overall.

Another reason that subject analysis tends to be under-specified is that this difficult process can be summarized so succinctly (i.e., “Identify significant concepts.”). Knowledge Organization scholars have attempted to add some detail and guidance to this succinct description of the subject analysis process. A dominant voice in the discussion of subject analysis methods is Patrick Wilson [138], who notes that indexing manuals “are curiously uninformative about how one goes about identifying the subject of writing.” Wilson seeks to remedy the lack of guidance by recommending an initial analysis step of identifying the “cast of characters”—“things named, mentioned, or referred to, and... concepts employed”—in a document and then describes four methods of analyzing that raw data as part of subject analysis:

- The **purposive method**, which aims to understand the creator’s aim or purpose for the document, whether stated directly or not.
- The **figure-ground method**, which identifies the concept or concepts that stand out in the document and asks: What is the central or most important figure of this document?
- The **objective method**, also called the counting references method or the constantly referred to method, which considers which concepts and terms are mentioned the most times in a document.
- The **cohesion method** or appeal to unity method, which aims to identify the theme that holds the document together.

While none of Wilson’s four methods is *the* definitive way to conduct subject analysis, in aggregate they provide useful guidance. Mai suggests adding one more method, which “could

be named the requirement-oriented or user-oriented method” to make Wilson’s framework more complete [92].

Langridge, who was mentioned above as a strong advocate for focusing on subject analysis, provides a framework that addresses the requirement- and user-orientation gaps that Mai identifies in Wilson’s methods. Langridge describes subject analysis as a process of asking three questions about form, topic, and audience: 1) What is it?, 2) What is it about?, and 3) What is it for? These questions serve as an almost a facet analysis and allow the indexer or tagger to not just identify a document’s topic (e.g., “economics”) but also its form (e.g., “social practice”) and its intended audience or user (e.g., “industrial managers”). F.W. Lancaster proposes a similar series of questions: 1) What is it about?, 2) Why has it been added to our collection?, and 3) What aspects will our users be interested in?” [85].

The ISO standard [77] follows the document examination and question-asking approaches outlined by Wilson, Langridge, and Lancaster. Like Wilson’s recommendation for identifying the cast of characters, the ISO standard instructs indexers to look at the title, abstract, table of contents, introduction, and other key parts of documents to understand what they are about. Then, like Langridge and Lancaster, the ISO standard encourages indexers to ask a checklist of questions, including: “Does the document deal with the object affected by the activity?” and “Does the subject contain an active concept?” These questions provide some scaffolding for the subject analysis process, but subject analysis remains a somewhat under-studied and under-specified first step in the subject indexing and subject tagging processes.

2.2 Subject Indexing

This section reviews literature about the subject indexing process, focusing on three topics that are highly relevant to this study: 1) the steps in the subject indexing process, 2) the importance of accuracy, specificity, and exhaustivity in subject indexing, and 3) different orientations for subject indexing, such as user orientation and content orientation.

2.2.1 The Steps in the Subject Indexing Process

While the literature generally agrees that subject analysis is the first step in subject indexing, knowledge organization 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 [48], Farrow [41], Smiraglia [121], and Lancaster [85] all propose a two-step model, which Lancaster summarizes as: “1) Conceptual analysis, and 2) Translation.” Miksa [99], Langridge [86], and Mai [92] posit a three-step process. Mai summarizes his version of the three-step process below:

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.

Mai also adapts a visualization from Miksa to show these three steps and the textual entities that serve as the input and output for each step, as show in [Figure 2.1](#).

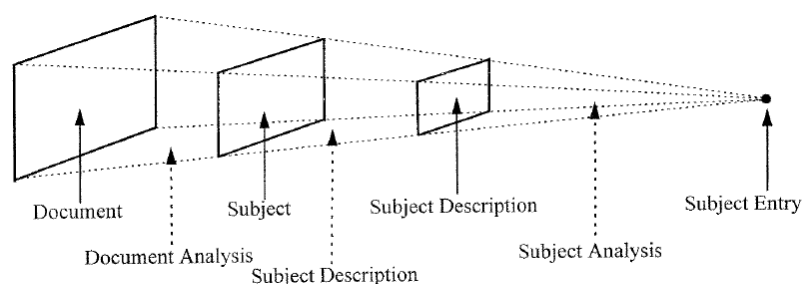


Figure 2.1: A three-step model of the subject indexing process, as depicted by Jens-Erik Mai and adapted from Miksa. Mai’s process includes the steps “Document Analysis,” “Subject Description,” and “Subject Analysis.” The steps involve increasingly summarized or focused representations of a document, ending with an indexing term or terms in a subject entry.

Note that the shape narrows from left to right, showing that subject indexing is a process of increasingly narrow summarization—moving from a whole document to a short description to eventually just an indexing term or terms in a subject entry. Note also that while

Mai confusingly uses the phrase “subject analysis process” to describe translation, these two- and three-step processes are largely in agreement and differ primarily on granularity.

A four-step process proposed by Chu and O’Brien [23] describes a similar set of steps with even more granularity:

1. Subject analysis of text.
2. Expression of the subject content in the indexer’s words.
3. Translation into an indexing vocabulary.
4. Expression of the subject in index terms.

This study considers all these models and uses a two-step model that largely agrees with the two-, three-, and four-step models discussed above. The model used in this study emphasizes inputs and outputs in defining where steps begin and end:

Step 1: Subject Analysis - The process of analyzing a document to determine what it is about.

- *Input*: Document to be indexed
- *Output*: Aboutness statement that summarizes the document

Step 2: Translation and Subject Representation - The process of translating the “aboutness” of a document, previously determined in the subject analysis step, into an indexing term or terms in an indexing language.

- *Input*: Aboutness Statement and indexing language
- *Output*: Indexing term or terms in a subject entry

F.W. Lancaster demonstrates an important aspect of translation and subject representation in [Figure 2.2](#), namely that when indexers translate the concepts that they identify during subject analysis (or concept analysis) the translation can vary depending on many factors, including the indexing language that they use [85]. Lancaster uses a classification scheme, a subject heading list, and a thesaurus as examples, but notes that the type of indexing language is not the main influence on translation. Rather, the scope and specificity

of the language influence how coextensive the aboutness statement and the indexing terms are.

Conceptual Analysis	Dewey Decimal Classification	Library of Congress Subject Headings	INSPEC Thesaurus
Industrial robots		ROBOTS, INDUSTRIAL	INDUSTRIAL ROBOTS
Artificial intelligence	670.427263 Artificial intelligence applied to robots in factory operations	ARTIFICIAL INTELLIGENCE	ARTIFICIAL INTELLIGENCE
Manufacturing operations		MANUFACTURING PROCESSES—AUTOMATION	MANUFACTURING PROCESSES
Materials handling	621.86 Materials handling equipment	MATERIALS HANDLING	MATERIALS HANDLING
Locomotion	531.112 Kinematics	ROBOTS—MOTION	KINEMATICS

Figure 2.2: Examples, provided by F.W. Lancaster, of translating concepts identified in an initial step of subject analysis (conceptual analysis) into indexing terms while using three different indexing languages during the following step, subject representation. The examples show that the terms provided in a controlled vocabulary can affect the indexing terms that an indexer assigns, even if they identify the same concept during subject analysis.

In discussing consistency between indexers, Lancaster further explores the relationship between vocabulary control, translation, and the coextensiveness between subject analysis and subject representation. He describes a thought experiment where inexperienced indexers (“a group of high school students”) are asked to index medical articles. Lancaster suggests that during subject analysis, the students “will look for words or phrases that occur frequently and/or appear in the title of other prominent places.” Lancaster would then “ask the students to translate their free text indexing into terms selected from the National Library of Medicine’s *Medical Subject Headings (MeSH)*.” Lancaster suggests that the “controlled vocabulary will have an adverse effect” because “students will have trouble selecting the appropriate control terms.” While the current study does not evaluate inter-indexer consistency, Lancaster’s thought experiment still informs the current study’s understanding of the translation process, especially for the inexperienced indexers and taggers in the cur-

rent study, and suggests that inexperienced indexers and taggers are likely to be affected by vocabulary control during the translation process, which would manifest in the degree of coextensiveness between their aboutness statements and indexing terms.

Finally, it is important to note that no matter how many steps a model has, in practice these steps can be iterative and “may take place almost simultaneously” [92]. Langridge, however, [86] argues for keeping the subject analysis step independent. Lancaster 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 [85].

Chapter 4 of this study discusses how combining or not combining the subject analysis and subject representation steps informs an understanding of coextensiveness—the degree to which the aboutness statement and the subject representation match—and influences the design of this study. In short, to best understand the influence of vocabulary control on the subject representation step, this study follows Langridge’s call to keep subject analysis and subject representation separate.

2.2.2 Specificity, Exhaustivity, and Accuracy in Subject Indexing

When indexers do subject indexing work, they are creating part of a representation or a surrogate of a primary document that can be used to make that primary document easier to find in search and enable a KOS to collocate that document with other documents about the same subject. These objectives of finding and collocating are two of the primary objectives of KOS and subject indexing [126], and these objectives rely on subject headings in surrogates being appropriately specific, appropriately exhaustive, and accurate. The following subsections explore these three concepts and how they relate to this study.

Specificity

In his *Rules for a Dictionary Catalog*, Cutter writes: “Enter a work under its subject-heading, not under the heading of a class which includes that subject.” His example is: “Put Lady Cust’s book on ‘The Cat’ under Cat, not under Zoology or Mammals, or Domestic animals” [29]. Cutter’s example summarizes one of the key principles of subject indexing: specificity, or using the most specific indexing term that is appropriate for the document being indexed. Svenonius calls the “injunction to use the most specific heading possible...the foremost rule in indexing” [126]. The *LCSH Subject Heading Manual* instructs indexers to “[a]ssign headings that are as specific as the topics they cover” and continues by noting that “[s]pecificity is not a property of a given subject heading; instead, it is a relative concept between a subject heading and the work to which it applies” [88]. Foskett reinforces the importance of using appropriately specific indexing terms: “The higher the specificity, the more likely we are to be able to achieve high relevance...If specificity is lacking, we are in fact reduced to the kind of sequential scanning that is necessary if our collections are not organized at all” [46]. Finally, an important part of the principle of specificity is that an indexer should use specific headings and not use redundant broader headings. For example:

Some students of indexing make the mistake of indexing redundantly. Having indexed an article on oranges under the term “Oranges,” they want also to assign “Citrus fruits” and even “Fruit.” This is quite unnecessary. Indeed, it is poor indexing practice. If the generic terms are assigned every time a specific term is used, it becomes difficult to distinguish general articles from specific ones [85].

While specificity is clearly one property that determines how “good” an indexing term is for a particular resource and is critical to supporting the finding and collocating objectives of subject indexing, this study does not attempt to evaluate whether participants’ subject headings and tags are “good.” Instead, this study focuses on indexers’ and taggers’ perspective by asking participants to self-report when they felt their indexing terms were appropriately specific, too specific, or not specific enough, and it tracks instances in which translation into an indexing language results in more or less specificity in comparison to the

concepts in an aboutness statement. This study pays particular attention to how different forms of vocabulary support (or do not support) participants' efforts to assign specific indexing terms related to racial and social justice.

Exhaustivity

While specificity describes how precisely indexing terms align with the concepts in a document, exhaustivity is about the level of detail and completeness with which a document is analyzed and described by indexing terms. A fully exhaustive index—or what Foskett describes as “depth indexing” [46]—would list all possible indexing terms for the concepts present in the document and could be quite lengthy. Lancaster describes the advantages of more exhaustive indexing: “Exhaustive indexing implies the use of enough terms to cover the subject matter of the document rather completely”; “The more information given, the more clearly the representation indicates the scope of the article and the more likely it is to indicate to a reader whether or not it satisfies some information need”; “The longer the representation, too, the more access points it provides” [85].

Exhaustive indexing or depth indexing, then, might sound like an inherently preferred approach to indexing. However, in practice, most subject indexing uses selective indexing instead of fully exhaustive indexing, especially for general purpose collections. Selective indexing aims to represent the most prevalent or most important concepts, following from Wilson's methods for subject analysis [138] and anticipating how users will search for the documents being indexed. Selective indexing is reflected in instructions in the *LCSH Subject Heading Manual* [88], which states the Library of Congress practice “Do not assign more than ten headings to a work” and recommends a maximum of six subject headings in most cases. The *Manual* also states the Library of Congress practice: “Assign headings only for topics that comprise at least 20% of the work.” These limits are partly an artifact of the practical limitations of catalogs being stored on physical catalog cards and partly a recognition of the labor involved in full exhaustive indexing, but, more importantly, they reflect that while fully exhaustive indexing can increase recall in information retrieval, too many assigned indexing terms can negatively affect precision in information recall.

As with specificity, this study does not evaluate the exhaustivity of participants' indexing terms. It does, however, consider the relationship between control and exhaustivity and listen to participants' perspectives about how exhaustively they were able to index the documents in the study while using different forms of vocabulary control.

Accuracy

Like specificity and exhaustivity, accuracy is critical to supporting indexing objectives. Also like specificity and exhaustivity, accuracy is a reflection of the full subject indexing process, from subject analysis to subject representation. Assigning accurate indexing terms relies on a correct analysis and an accurate—or faithful [111]—representation of the identified concepts in the indexing terms. That faithful representation can depend on how expressive the indexing language is, how well the indexer knows how to work with the indexing language, how well the indexer understands the topic of the document that they are indexing, and other factors. While accuracy might be equated to correctness, Lancaster notes that “indexing terms assigned to a document cannot be judged ‘correct’ or ‘incorrect’ in any absolute sense. That is, there is no one ‘best’ set of terms” [85]. He does, however, list types of errors that occur during subject indexing, including one type of error directly relevant to accuracy: “The indexer uses an obviously incorrect term, perhaps through lack of subject knowledge.”

As with specificity and exhaustivity, this study does not evaluate the accuracy of participants' indexing terms. Instead, this study focuses on indexers' and taggers' perspective by asking participants to self-report when they felt their indexing terms were accurate and aims to understand what causes indexers to use indexing terms that they consider accurate or not accurate.

2.2.3 Different Orientations in Subject Indexing

Another significant topic in the literature about subject indexing is the orientation of the indexing process. Orientation considers how indexers approach the process of subject indexing, what they value or prioritize in determining how to index a document, and how they

express those values through indexing terms. For example, a subject indexer might adopt a document- or content-oriented approach to indexing, prioritizing a faithful representation of the content in the document. Most of Wilson's methods for subject analysis lead to these orientations [138]. Alternatively, a subject indexer might take a user- or request-oriented approach to indexing.

User- and request-oriented indexing, as described by Soergel [123] and Fidel [44], consider how information seekers look for information and the language that they use to search for information. They also, like Beghtol [9], anticipate what topics and concepts information seekers will consider important in documents. Bates describes the challenges of this type of anticipation in user-oriented indexing: "the challenge for the indexer is to try to anticipate what terms people...might search for in those cases where the record in hand would, in fact, go part way in satisfying the user's information need. This can be seen as a very peculiar challenge" [8]. Birger Hjørland agrees that subject indexing must account for the user: "Subject analyses should establish priorities based on the specific user group served...The best subject analysis is the one that makes the best prognosis of the future use of the document" [66]. Brown et al., [19] and Hilderley and Rafferty [65] take the user-oriented approach a step further by suggesting "democratic indexing," in which users add their own indexing terms to fit their perspective, anticipating tagging approaches that are discussed later in the chapter.

While he advocates for user-oriented indexing, Hjørland also advocates for a domain-oriented approach to subject indexing, emphasizing approaches to subject indexing that specifically serve a particular domain [67]. Another orientation in subject indexing is requirement-oriented indexing, which typically focuses on business requirements for corporate collections [92]. Finally, Šauperl [117] notes that many experienced subject indexers take a collection-oriented approach to subject indexing, paying close attention to where new documents fit within an existing collection and choosing indexing terms based on that history and the desire to collocate resources in the library.

These different orientations are not necessarily mutually exclusive, but they emphasize different perspectives and motivations in subject indexing, and they can affect the indexing terms that indexers prefer and how indexers interact with vocabulary control. For example,

an indexer working more from a content-oriented perspective might prefer indexing terms that match the language in a book title or the terms used in an abstract of an academic article. In contrast, an indexer working from a user-oriented perspective might be motivated by anticipating potential use search terms or the values and language of index users. These types of different approaches—and the motivations and goals behind them—help to explain the perspective of subject indexers (and taggers) as they do their work. While this study does not directly investigate orientation in subject indexing, instead primarily focusing on the effects of vocabulary control, it does consider these orientations as important context for establishing a holistic understanding of each participants’ subject indexing and subject tagging processes and for better understanding the effect of vocabulary control on their processes.

2.3 Vocabulary Control and Criticisms of Vocabulary Control

Controlled vocabularies—collections of authorized indexing terms and the rules that govern their use—are long-established features of KOS, integral to subject indexing, and a central focus of this study. Controlled vocabularies minimize ambiguity and promote consistency [77], improve precision and recall in information retrieval [84, 124], and help to collocate related resources [126]. Popular controlled vocabularies include subject heading lists like the *Sears List of Subject Headings* and the *LCSH* and classification schemes like the *Dewey Decimal Classification (DDC)* and the *Library of Congress Classification (LCC)*. These controlled require that indexers and classifiers use authorized indexing terms and classes to ensure uniformity and uniqueness. For example, the *LCSH* authorizes the subject heading “Sofas” and does not allow indexers to use the term “Coaches.” Or, as discussed in the Introduction, the *LCSH* requires that indexers use the term “Indians of North America” instead of terms such as “Native Americans” or “American Indians.” Controlled vocabularies also provide instructions and examples for increasing specificity by, in the case of the *Sears List*, adding subject headings, or, for both *Sears* and the *LCSH* creating synthetic, pre-coordinated subject headings through the use of subdivisions. For example, the *Sears List* instructs indexers to add specific types of birds (“e.g., Birds of prey, canaries”) to local catalogs even though they are not enumerated in the *Sears List* [137]. And both the

LCSH and the *Sears List* support the use of subdivisions (e.g., “Butterflies—Migration”) to increase specificity. Despite these allowances for indexing with non-enumerated terms while following specific instructions, controlled vocabularies are still, by definition and by purpose, restrictive.

Because vocabulary control is so central to subject indexing, the Knowledge Organization community has often reflected on and critiqued controlled vocabularies, studying their value in modern KOS [43, 57]; contrasting controlled vocabularies with natural language search [113] and with uncontrolled vocabularies [128, 1]; and criticizing controlled vocabularies for reinforcing systemic biases [105, 47, 49, 10], for being too slow to adapt to change [129, 14, 11], and for not reflecting the language of information seekers and many groups and cultures [75, 38]. These criticisms, which are many and which form the primary inspiration for this study, 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 three categories of criticisms—reinforcing systemic biases, being too slow to adapt, and not reflecting the language of information seekers and many groups and cultures—in more depth and describe how leaving these criticisms unaddressed can cause continued harm to historically marginalized groups.

2.3.1 Criticism: Controlled Vocabularies Reinforce Systemic Biases

A defining characteristic of controlled vocabularies is that they have authorized terms and non-authorized (or non-preferred) 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 about the same topic. However, vocabulary control can also reflect the implicit or explicit biases of the designers and editors of the controlled vocabulary—and can reinforce and make authoritative those biases in all the systems, people, and documents that the KOS touches. Foskett offers a somewhat measured evaluation of biases in the *LCSH* and the *Sears List*, noting that these controlled vocabularies “are

designed for use in Western libraries” and “reflect the historical biases of those libraries and their users” [45]. While Foskett sees the utility of Western controlled vocabularies like the *LCSH* for Western users and libraries, many catalogers and scholars are much more critical of biases in controlled vocabularies. Sanford Berman explains why these biases are persistent and harmful:

[I]n the realm of headings that deal with people and cultures—in short, with humanity—the LC list can only “satisfy” parochial, jingoistic Europeans and North Americans, white-hued, at least nominally Christian (and preferably Protestant) in faith, comfortably situated in the middle- and higher-income brackets, largely domiciled in suburbia, fundamentally loyal to the Established Order, and heavily imbued with the transcendent, incomparable glory of Western civilization. Further, it reflects a host of untenable—indeed, obsolete and arrogant—assumptions with respect to young people and women. And exudes something less than sympathy or even fairness toward organized labor and the sexually unorthodox and “*avante garde*.” [10]

Berman also notes that indexing languages like the *LCSH* are unlikely to be superseded despite catalogers’ frustrations with their shortcomings because they are deeply embedded in our information ecosystem, and, without revolutionary change, are designed to persistently spread and legitimize harmful biases. While the work of radical catalogers like Berman and the rich work of critical cataloging scholars over decades have improved many of the problematic headings that Berman identified in *Prejudices and Antipathies*, Berman noted in a more recent interview [11] that systemic bias and colonial attitudes persist in the *LCSH* and that some of the headings that he critiqued decades earlier still remain in place.

Other scholars note that persistent bias is not tied only to terminology but also to the structure of indexing languages [105] and can even be exacerbated by attempts to remove bias [49]. Importantly, these persistent and systemic biases affect how information seekers find and perceive documents—and how they perceive themselves. For example, Howard and Knowlton [75] show that researchers looking for information about Black Studies and LGBTQ+ studies—and particularly scholars looking for information at the intersection of

these two fields—struggle to find information that reflects their identities because of system 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 analyzes a history of racism in classification schemes and calls for “taxonomic reparations” and calls for “multiple, local, community-based, and personal ways to organize knowledge and ideas” [2]. This study aims to build on the work of these critical catalogers by examining how different forms of vocabulary control affect indexers’ (and taggers’) approaches to representing documents and how different forms of vocabulary control can perpetuate or minimize bias.

2.3.2 Criticism: Controlled Vocabularies Are Too Slow to Adapt

In establishing subject ontogeny, the study of how indexing languages change over time, Joseph Tennis analyzed how the class for “Eugenics” moved throughout the DDC schedule based on changing social perceptions of the value of eugenics [129]. These types of changes to indexing languages (whether structural or terminological) typically require some type of warrant—evidence that a new term or a new usage are widespread and well enough established that the editors of indexing languages are confident that they should update their indexing language and require indexers to go to the expense of updating their indexes. The expenses associated with updates are one reason that indexing languages tend to be conservative about change and updates tend to lag behind common usage [14].

This lag can be annoying when new technologies like personal computers and smart phones are slow to appear in controlled vocabularies, but the lag can be harmful when hard-won civil rights are not reflected in authoritative indexing languages, when people’s identities cannot be accurately represented in indexing terms, and when terminology that has become offensive (or always was offensive) persists in controlled vocabularies. This type of problem is evident in the “Indians of North America” heading in the *LCSH* that was discussed in the Introduction and in the “Illegal aliens” case that is discussed below. 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 [72].

One of the goals of this study is to explore ways to mediate the harm done by outdated

and pejorative authorized indexing terms by exploring different forms of vocabulary control and better understanding, through the eyes of indexers and taggers, when they have control or lack control to use language that prevents harm.

2.3.3 Criticism: Controlled Vocabularies Do Not Reflect the Language of Information Seekers and Many Groups and Cultures

As discussed in the two previous sections, controlled vocabularies tend to reflect systemic biases, to favor those in power in society, and to be slow to change, all of which can cause harm to historically marginalized and underrepresented groups. One form that this harm can take is in 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. In addition to having practical consequences related to information retrieval, this misrepresentation, whether intentional or unintentional, can cause large symbolic and spiritual harm by echoing colonial misnaming as a form of controlling marginalized groups and limiting the visibility and legitimacy of marginalized groups, including Indigenous tribes [38].

One illustrative example of these criticisms and how they are intertwined is the case of the “Illegal aliens” headings in the *LCSH* [6, 90]. This dated and prejudiced term appeared while a student was conducting research at the Dartmouth University Library because it was a preferred term in the controlled vocabulary of the *LCSH*. 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 for a taxonomy of harm in KOS [3]) and that controlled vocabularies can control more than just indexing terms. Controlled vocabularies can also control the self-identities of information seekers and they can control the ability of indexers to represent documents in socially just and equitable way [73].

Snow and Dunbar

Because this connection between control and social and racial equity is a significant

aspect of this study, the documents chosen for participants index and tag in the study were about racial and social justice and analysis focused on

2.4 Tagging and “Uncontrolled” Vocabularies

To this point, this literature review has focused on subject indexing and the types of controlled vocabularies associated with subject indexing, like the *LCSH* and the *Sears List*. This section introduces another form of Knowledge Organization 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 some 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 [42] to establish a consensus set of tags for a resource [81, 61]. 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 [53]—two issues that the uniform and unique headings in controlled vocabularies like the *LCSH* and the *Sears List* are design to address. ¹

Early forms of tagging and folksonomies faced many criticisms. Critics most often cited the orthographic and semantic limitations of folksonomies [81, 97, 110] and how these limitations lead to ambiguity. For example, taggers working without a controlled vocabulary might create synonymous tags such as “albums,” and “vinyl” and “records” (all of which have ambiguous meanings), orthographic variations such as “vinyls,” and even typos such as “viynl.” In contrast, the *Sears List* has an entry term “Records, Phonograph” and an authorized heading “Sound recordings” that ensure a uniform heading [137]. Tagging also has problems differentiating homographs such as “bat” (for baseball) and “bat” (winged

¹Folksonomies were often called uncontrolled vocabularies because they historically did not employ vocabulary control; however, this study distinguishes between folksonomies, which are an aggregation of tags that might or might not have some form of vocabulary control, and uncontrolled vocabularies, which are defined by the lack of restriction on vocabulary—not necessarily the use of tagging.

mammal), a problem that controlled vocabularies solve by establishing unique headings for each different topic. Because they lack the uniform and unique headings that controlled vocabularies such as the *LCSH* and the *Sears List* employ, tag-based folksonomies are criticized for their weakness in key KOS objectives such as collocation of materials about the same subject and precision and recall in search. Social tagging and folksonomies are also criticized for not defining the relationships between broader and narrower terms, often resulting in problems with the level of specificity of assigned tags and creating confusion about the meaning of tags or about how works with different tags relate to each other. Folksonomies, because they are typically open to contributions from a broad community, are also susceptible to “graffiti,” intentionally harmful tags or misinformation (such as tags designed to push harmful agendas or conspiracy theories), and unintentionally inaccurate or insufficiently specific tags that result from community members lacking domain expertise.

Despite these criticisms, folksonomies were considered a revolutionary alternative 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 [7], to “directly reflect the vocabulary of users” [97], and to better “reflect the population’s conceptual model of information” [110]. Adler, in contrasting tagging with a traditional controlled vocabulary, frames this difference as a difference in 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” [1]. 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 has space in the “long tail” of tags to more exhaustively represent a diversity of 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 perpet-

uating biases of those in power. Tagging does this, in part, by not privileging—and not granting controlling power to—a single, preferred view of the world. Instead, as Shirky argues, tagging intentionally rejects the idea that there is one correct view of the world and honors many different points of view:

[I]f we are, from a bunch of different points of view, applying some kind of sense to the world, then you don't privilege one top level of sense-making over the other. What you do instead is you try to find ways that the individual sense-making can roll up to something which is of value in aggregate, but you do it without an ontological goal. You do it without a goal of explicitly getting to or even closely matching some theoretically perfect view of the world. Critically, the semantics here are in the users, not in the system [119].

While “roll[ing] up to something which is of value in aggregate” can be a challenge of coordination and organization, Shirky’s quote shows the value of diverse viewpoints and shows how folksonomies have a different and decentralized dynamic of control when compared to traditional controlled vocabularies.

Finally, it is worth noting that tagging can have a financial advantage over subject indexing. Folksonomies rely on volunteer labor and require minimal central coordination, which means that costs were quite low compared to traditional subject indexing, which is typically done by paid professionals and requires continuous coordination with the controlled vocabulary.

2.5 “Subject Tagging” as a Specific Type of Tagging

This study aimed to make direct comparisons between the types of vocabulary control associated with tagging and the types of vocabulary control associated with subject indexing. However, researchers have found that social taggers use tags for a wide variety of purposes, including many purposes other than identifying the subjects of documents, making this type of direct comparison difficult. For example, Golder and Huberman note that tags are used for many functions other than describing what a document is about [53]. Their taxonomy of tagging strategies includes:

1. Identifying what (or who) it is about
2. Identifying what it is
3. Identifying who owns it
4. Refining categories
5. Identifying qualities of characteristics
6. Self-reference
7. Task organizing

These strategies cover a wide variety of use cases, including tags whose value does not extend beyond the person who uses the tag. For example, self-reference tags might include “my stuff” and task organizing tags might include “toread,” both of which amount to noise when using tags for subject retrieval.

Munk and Mørk [103], while they view tagging as primarily an act of categorization, present a similar taxonomy of tagging strategies, with some additional strategies:

1. Tagging as content categorization
2. Tagging as media categorization
3. Tagging as categorization subject to copyright
4. Tagging as a type of categorization
5. Tagging as categorization through value judgment
6. Tagging as categorization referring to categorization
7. Tagging as categorization on the basis of jobs and processes
8. Tagging as a mark in time
9. Tagging as an exclusive personal categorization

Importantly, “tagging as categorization through value judgment” indicates that taggers can communicate their values and their personal opinions about documents through the tags that they assign.

Greenaway et al., [56] classify tagging practices on the video sharing platform YouTube and document sixteen different tagging strategies or practices, which they group into four high-level categories:

1. Generic relationship between tag and video content
2. Specific relationship between tag and video content
3. Tag only useful to a minority or users, specific individual or group
4. Irrelevant/Non Useful Tags

The first two categories differentiate between tags that describe content in general and tags that require specialized knowledge; however, both categories include subcategories that could be related to subject. From the first category:

1. Tag generically identifies what video is ‘of’
2. Tag identifies video Category/Genre

And from the second category:

1. Tag specifically identifies what video is ‘of’ (place names/events)
2. Tag specifically identifies what video is ‘of’ (people/animals/objects)
3. Tag identifies what video is ‘about’

These five categories of tags account for approximately 43.4% of the total tags that Greenaway et al., analyzed, showing that taggers use tags for many purposes other than representing the subjects of documents, create many tags for just their personal use, and submit many tags with typos or parsing problems.

These tagging taxonomies from Golder and Huber, Munk and Mørk, and Greenaway et al., all show that taggers overload the use of tags and do not restrict themselves to only subject tags the way that subject headings exclusively represent the subjects of documents. For this study to directly compare the forms of vocabulary associated with subject indexing and the forms of vocabulary control associated with tagging, this study focuses specifically on the types of tags Golder and Huberman call “identifying what (or who) it is about,” what Munk and Mørk call “tagging a content categorization,” and what Greenaway et al., describe in the five subcategories listed above. This study names this subset of tagging that focuses on the subject of documents as “subject tagging” and treats subject tagging as a very similar

process to subject indexing. Subject indexing is the process of determining what a document is about (subject analysis) and then translating that “aboutness” into a representation in an indexing language (subject representation). Similarly, subject tagging is the process of determining what a document is about (subject analysis) and then representing that “aboutness” with keywords or tags (subject representation). Because these two processes focus on the subject of documents and share the same first step of subject analysis, this study could be designed to guide participants through the subject analysis and subject representation steps and make more direct comparisons of the role of vocabulary control in tagging and indexing.

2.6 Five Forms of Vocabulary Control

The previous sections contrasted the historical use of vocabulary control in subject indexing and the historical lack of vocabulary control in tagging and subject tagging. While this dichotomy is instructive, it is an incomplete picture because tagging technologies and communities have evolved since early folksonomies. This section expands upon the comparison between subject indexing and subject tagging by naming and describing five forms of vocabulary control that are in use today. These forms of vocabulary control, while they have been studied separately and in various combinations, have not previously been studied and compared all together in a taxonomy of five forms of vocabulary control.

2.6.1 Minimal Vocabulary 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 while they work. With minimal control, the vocabulary is an amorphous collection of tags that is constantly updated with each new tag that is added, including any typos and synonyms that taggers enter. Minimal control is the cheapest type of vocabulary control to implement 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 knowledge organization research noted with early folksonomies [81, 97, 110].

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 UI 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 post. Except for a retroactive tagging effort for old blog posts [69], MetaFilter employs minimal moderation to fix abuse or other problematic uses of tags. Similarly, the “Fandoms,” “Relationships,” “Characters,” and “Other Tags” fields on the fan fiction 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.

2.6.2 Passive Vocabulary Control

Passive vocabulary control is similar to minimal control in that taggers are free to enter any tags that they like. However, the UI 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 when they tag documents. The early tagging website del.icio.us supported passive vocabulary control by displaying clickable buttons with popular tags for a given resource. Clickable buttons have fallen out of favor over the past decade and have largely been replaced with an autocomplete feature that displays suggested tags as users type into a text box. A recent study confirmed that an autocomplete feature is much more likely than clickable links to encourage taggers to use terms from a controlled vocabulary—and to avoid some of the retrieval problems associated with typos and synonym tags [71]. 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,” “#universitywashingtonhuskies,” and other popular tags. Users can choose to use or

ignore these algorithmically generated suggestions to tag their posts, but the ease of selecting a popular tag is considered a key feature—and it is a defining characteristic of passive vocabulary control.

2.6.3 Post Hoc Vocabulary Control

While minimal control and passive control function with little editorial intervention, post hoc control involves active editorial control of tags to manage synonyms, homographs, and typos. Julia Bullard studies three websites that employ post hoc vocabulary control—Stack Overflow, LibraryThing, and Archive of Our Own—and terms them “curated folksonomies” [20]. Bullard 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. [78] adopt the metaphor of “gardening” from Peters and Weller [108] to describe how they use this approach of harnessing user-generated tags with human judgement to build a structured folksonomy in a citizen science project.

Because human 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 been able to encourage qualified community members to volunteer to do this editorial work 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,” ensuring that “linux” is displayed on posts that had been tagged “linux-development,” which improves collocation, precision, and recall on Stack Overflow [107].

2.6.4 Flexible Vocabulary 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” [137], following established rules and patterns. For example, a general reference under the subject heading “Flowers” instructs indexers to add types of flowers (e.g., “Roses”). By making the creation of new subject headings an explicit and encouraged feature of the indexing language, the *Sears List* provides a flexible and adaptive form of vocabulary control that balances the benefits of standardized indexing terms with local agency for indexers. For example, indexers can, in some cases, mitigate the problem of controlled vocabularies being slow to adapt to change by adding their own subject headings. For example, my used copy of the 21st edition of the *Sears List* includes annotations for indexer-added terms for representing that LGBTQ+ community that were absent from the *Sears List* until after the *Obergefell v. Hodges* decision provided what *Sears* editors consider sufficient warrant to add more inclusive subject headings.

However, flexible vocabulary control does come 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 indexers to use an authorized heading if it is available and also requires indexer to not use non-authorized 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.

2.6.5 Rigid Vocabulary Control

With rigid control, indexers can use only those indexing terms that are authorized and in the controlled vocabulary. While rigid control provides indexers with little flexibility, it is used widely to ensure uniform and unique indexing terms and considered a gold stan-

dard 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—while technically possible—difficult and expensive to create and maintain. For example, an indexer using the *LCSH* might disagree with a subject heading like the problematic heading “Illegal aliens” that was discussed above. 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 that not all libraries have. Rigid control is also enforced by how headings can be added or changed—through centralized approval of new and updated subject headings. Unlike the flexible control of the *Sears List*, which encourages indexers to add new subject headings as needed to their local index, the *LCSH* requires an extensive review process through the Subject Authority Cooperative Program (SACO) and that committee’s approval before a heading is added or changed in the *LCSH* and available to subject catalogers.

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*, which uses flexible control) 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” and still be working within the rules of rigid vocabulary control even though “African Americans—Blogs” is not enumerated in the *LCSH*.

2.7 Forms of Vocabulary Control, Control Relationships, and Systemic Bias

Studying the five forms of vocabulary control discussed above is important, not just because they have not previously been studied together, but also because “[i]n KOS, control relationships are most often communicated through controlled vocabularies” [73] and because these control relationships are intimately connected with the criticisms of vocabulary control that were discussed above. Control relationships are relationships where one actor

has the power to influence, limit, or direct the behavior or experience of another actor. The actor who holds the controlling role in the relationship holds power and can make decisions that affect others. Conversely, the actor in the controlled role lacks agency and can make only suggestions—not decisions—that affect other actors. For example, the Dartmouth librarians felt controlled by the editors of the *LCSH* and by the authorized headings that used the phrase “Illegal aliens.” As that example shows, the dynamics of these control relationships directly inform racial and social equity in KOS. They are also particularly important because people of color and other marginalized groups have historically been underrepresented—and continue to be underrepresented—in controlling roles. This section reviews literature about the concept of control in KOS and illustrates the hierarchical structure of the control relationships between 1) information seekers, 2) indexers and taggers, and 3) the designers and editors of controlled vocabularies in these KOS, emphasizing the experience of the actors at the middle of this hierarchy, subject indexers and taggers and their role in promoting social and racial equity. The section also discusses the five different forms of 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.

2.7.1 Vocabulary Control and Bibliographical Control

In *Two Kinds of Power*, Patrick Wilson uses the term “bibliographical control” to summarize the motivation and intent for vocabulary control and to describe the foundational control relationship in KOS: information seekers’ “power over texts” [138]. Wilson further describes this control relationship as “access to knowledge about texts,” “bibliographical control over a body of literature,” and the ability to “line up a population of writings in any arbitrary order, to make the population march to one’s command.” These phrases put information seekers in a controlling role and texts in a controlled role, and his continued emphasis on information seekers shows that Wilson views the goal of controlled vocabularies and subject indexing to be one of maximizing information seekers’ bibliographical control.

Wilson, however, describes two types of bibliographical control—exploitative control and descriptive control—to show that information seekers cannot exercise complete control over

texts. 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. That degree of control over texts is unattainable according to Wilson because it requires indexers to appraise the merits and appropriateness of texts for each specific information request. Descriptive control, which is widely implemented in libraries and other KOS, is evaluatively neutral and returns to information seekers all texts that satisfy a given search criteria. While Wilson's thought experiment shows that he views information seekers' control over texts as the fundamental purpose of KOS, it also shows that their control depends on how indexers use controlled vocabularies to index texts. Put another way, an ideal indexer or tagger would enable an information seeker ideal control over texts through exploitative control, but real indexers and taggers—through their use of vocabulary control—can limit and control some information seekers' control over texts. Therefore, Wilson shows us that vocabulary control informs at least two cascading control relationships: 1) information seekers' control over texts, and 2) indexers' control over information seekers.

2.7.2 Controlled Vocabularies, Systemic Bias, and Control of Information Seekers

Wilson's conclusion that exploitative control is unattainable is based, in part, on the understanding that indexers (and taggers) cannot anticipate and satisfy the requirements of every information seeker's unique needs each time they search for information. He argues, however, that descriptive control is an effective and attainable approximation for exploitative control. Hope Olson problematizes that conclusion, arguing that effective bibliographical control—even imperfect descriptive control—is unattainable for many information seekers because of systemic biases in controlled vocabularies. These systemic biases have historically meant that members of advantaged groups—in the United States, those groups are typically male, white, straight, cis, and Christian—have likely experienced effective descriptive control while members of less advantaged class are less like to experience that type of control over text. Olson summarizes how vocabulary control privileges particular groups and critiques the assumption that controlled vocabularies are a universal reflection of reality by stating that “universal language marginalizes and excludes the other” [105]. Olson shows

how 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, Jonathon Furner [49] also 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 the “self-identity principle, which specifies that 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.” 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 and designers of indexing languages to information seekers. Howard and Knowlton [75] echo this critique, noting that information seekers lack control while searching for documents that reflect intersectional identities, such as documents at the intersection of Black studies and LGBTQIA studies, because of the design of indexing languages with rigid vocabulary control and the information retrieval tools used with them.

2.7.3 Rigidly Controlled Vocabularies, Systemic Bias, and Control of Indexers

This lack of control manifests not only in information seekers who are unable to search for documents about people like themselves but, critically for this study, in indexers and taggers who control information seekers’ experience but who often lack control to accurately rep-

resent documents because of biases in controlled vocabularies, particularly rigid controlled vocabularies. Sanford Berman's *Prejudices and Antipathies* [10] and his reflective interview more than forty years later [11] make both a brilliant critique of the *LCSH* and serve as a testament to the frustration that indexers feel in working with a controlled vocabulary that they do not control. Berman wrote detailed arguments, packed with warrant and ethical arguments, for the *LCSH* to cancel and change racist, misogynistic, and biased subject headings because he understood that standards like the *LCSH* controlled vocabulary can control what indexers do, can distort how they represent documents, can perpetuate systemic bias, and can affect how information seekers find and perceive these documents.

This dynamic of indexers feeling a lack of control because of a controlled vocabulary played out dramatically in the previously mentioned case of “Illegal aliens” headings in the *LCSH*. When Melissa Padilla encountered these headings in the Dartmouth Library, she rallied other members of the CoFIRED (Coalition for Immigration Reform, Equality, and DREAMers) student organization to demand changes to the Dartmouth Library catalog. The initial, private response by Dartmouth librarians was to ask, “Don’t these students know that we don’t control these headings?” [6]. The librarians, while sympathetic to the students’ desire to not be labeled with a pejorative term, felt that they lacked control to change the headings because the headings were in the *LCSH* controlled vocabulary. After discussing who had the power to change the headings, the Dartmouth librarians teamed with students and administrators to submit proposed changes to the Subject Authority Cooperative Program (SACO) at the Library of Congress. This proposal, echoing Berman’s proposals in the past (and eventually gaining his endorsement), shows that indexers can make *suggestions* to change rigid controlled vocabularies but their place below designers and editors of indexing languages in the hierarchical structure of control relationships in KOS means that they cannot make *decisions*. Eventually, the Library of Congress made a compromise change, but the replacement headings still use the word “illegal” and the change took years of back-and-forth, support from the ALA, and significant political wrangling. At no point during this process were librarians or information seekers in control of the terms used to index and retrieve documents about undocumented immigrants. This lack of control yields a dynamic where biases can spread and perpetuate and where indexers have little

recourse but to suggest changes or to work, often ineffectively, around widespread bias.

One promising way for indexers to work around control relationships and rigid controlled vocabularies is to develop their own indexing languages to complement or supplement broad subject headings lists like the *LCSH*. Snow and Dunbar [122] frame this approach with one of the tenets of Critical Race Theory (CRT), advocating for the power of “counter-storytelling,” which Kumasi et al., explain “centralizes the voices and experiential knowledge of people of color as...legitimate and uniquely situated to describe systems and social structures from their vantage point of experiencing the effects of institutional racism on a daily basis” [83]. While this strategy does not guarantee that large indexing languages will adopt indexing terms from these alternative controlled vocabularies, the languages have value as independent entities and can serve as warrant for new and updated terms in large indexing languages. This approach is particularly relevant for marginalized groups that historically have not been in controlling roles in KOS. 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” [33, 24]. That cutting-edge perspective comes from people within the LGBTQ+ community who have lived experiences and domain expertise that inform their vocabulary decisions. Similarly, the First Nations House of Learning (FNHL) subject headings, which were developed at the University of British Columbia and are used at the Xwi7xwa Library, aim to “better reflect Indigenous Peoples’ self-identities, modes of understanding, and traditional knowledge” [130, 35]. The Mashantucket Pequot Thesaurus of American Indian Terminology is another controlled vocabulary built from an Indigenous perspective [89]. 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*.” 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. Of course, this approach to changing broad and widely used controlled vocabularies to promote social equity does not happen quickly and is not guaranteed to succeed, but it does represent

a meaningful path forward for indexers who feel controlled by rigid vocabulary control.

2.7.4 *Control Relationships and Less Rigid Forms of Vocabulary Control*

While criticisms of controlled vocabularies, particularly rigid controlled vocabularies, show that indexers often lack control and have minimal recourse for (often slow-moving) remedies, other forms of vocabulary control—like minimal, passive, post hoc, and flexible 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” [26]. Because Twitter has passive vocabulary control, 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 non-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 determined by their subject analysis work and as required by their local collection, within certain constraints but without any type of approval from editors or designers of the language. At least historically, these additions have been made by writing in the margins of a print copy of the *Sears List* at a local library, as shown in [Figure 2.3](#). While these handwritten annotations afford indexers little or no control over the centralized *Sears List* vocabulary, which is determined by editors and published for all libraries, 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*.

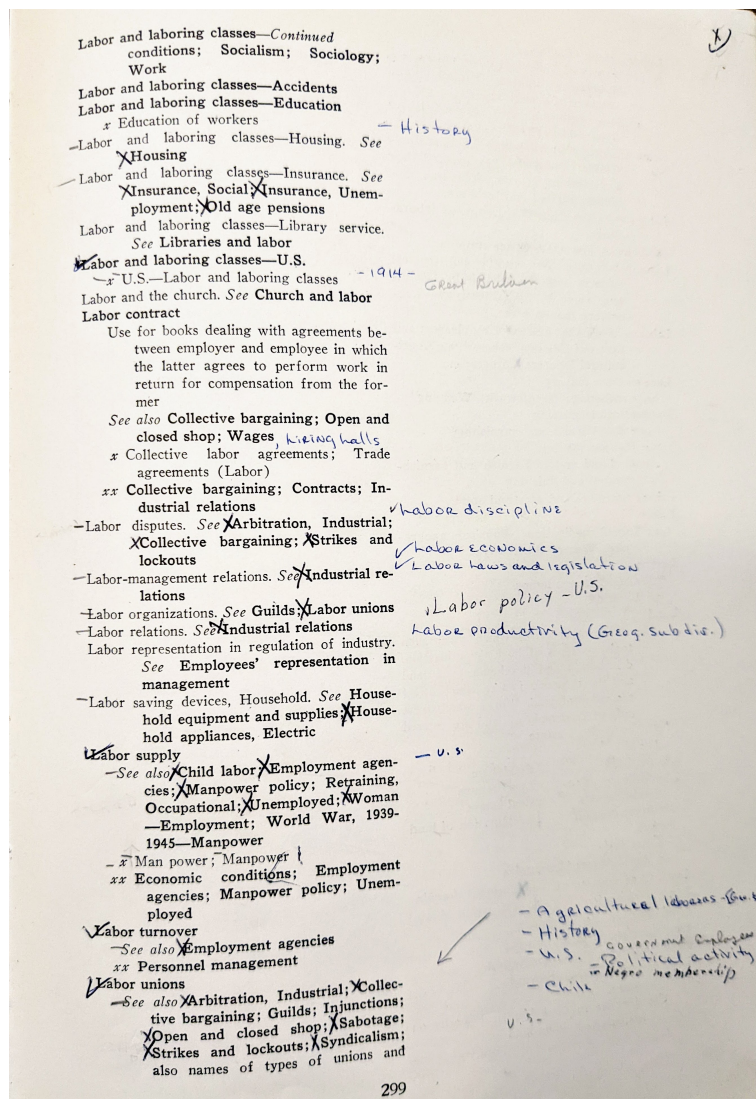


Figure 2.3: Hand-written annotations in a print copy of the 10th edition of the *Sears List of Subject Headings* indicating local changes to the controlled vocabulary. Indexers using the *Sears List*, which employs flexible vocabulary control, are encouraged to add detailed subject headings as necessary for their collections. These annotations include new subject headings, “Labor economics,” “Labor laws and legislation,” and “Labor policy—United States” among others.

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 related to racial and social justice?

2.7.5 *Controlled, In Control, and Out of Control*

So far, this section 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” headings and all the criticisms described by Berman, Olson, Furner, and Howard and Knowlton). While these examples all 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 in subject indexing and subject tagging work. The other axis also has “in control” at one end, but has “out of control” at the other end. These two axes are shown visually in [Figure 2.4](#).

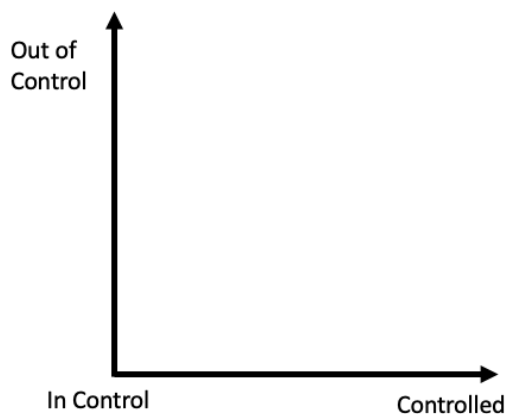


Figure 2.4: Control in this study is understood along two axes: 1) from feeling in control to feeling controlled, and 2) from feeling in control to feeling out of control. Feeling controlled and feeling out of control are both opposites of feeling in control, but they are shown as separate axes to indicate that they are different types of relationships. For example, one might expect rigid vocabulary control to appear near the bottom right corner of the graph, which would indicate that rigid vocabulary control makes indexers feel controlled instead of in control but also in control instead of out of control.

The second axis, between feeling in control and feeling out of control, while discussed less extensively in Knowledge Organization literature, is also important for this study, and it 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. Conversely, one could expect that the lack of guidance provided by minimal vocabulary control might make some indexers and taggers feel more out of control as they attempt to determine useful indexing terms without the constraints of authorized or suggested indexing terms.

Understanding the dynamics of these two ways of feeling less in control are particularly relevant for the “middle” forms of vocabulary control (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? This study aims to answer these types of questions, which span across two axes of control,

which are likely to differ for the five forms of vocabulary control, and which are likely to challenge the critical cataloging consensus that vocabulary control puts indexers in a controlled role, are central to this study and will hopefully reveal a nuanced relationship between indexers’ and taggers’ feelings of control, the five different forms of vocabulary control, and social and racial justice in KOS.

2.7.6 Control and the Ease of Use of Indexing Languages

Finally, this discussion of different forms of vocabulary control, control relationship, and feelings of control is incomplete without addressing the relationship between control and ease of use, particularly for inexperienced indexers and taggers. This study asked inexperienced indexers and taggers to index and tag books using different forms of vocabulary

control, including forms of vocabulary control that were represented by complex and possibly intimidating indexing languages like the *Sears List of Subject Headings* and the *LCSH*. Indexers and taggers with limited exposure to these indexing languages, their long histories, and their many rules might find them difficult to use and might encounter confusion and roadblocks unrelated to the focus on this study, the inclusion or exclusion of specific terms in the controlled vocabulary. For example, some participants might be confused by how to construct 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. (Then again, other participants might find these experiences intuitive and helpful.) This study used a series of questions and ranking activities to understand how feelings of control relate to the form of vocabulary control and how feelings of control relate to general ease of use, with the aim of homing in on the effect of vocabulary control on the subject indexing and subject tagging processes. [Chapter 4](#) discusses interview questions and other techniques for isolating and understanding these differences.

2.8 Empirical Studies of the Subject Analysis and Subject Indexing Processes

The previous sections in this chapter broadly explored literature about subject analysis, subject indexing, tagging, and vocabulary control to set context, establish some of the key discussions in Knowledge Organization (KO), and to motivate the research questions and design of the current study. This section narrows the focus by reviewing four specific studies that examine subject analysis and subject indexing and that closely match the methodology of the current study. The four empirical studies that are presented in this section focus on the subject analysis or subject indexing processes; use methods that allow researchers to observe these processes in action; and use methods like think-aloud protocols, interviews, and surveys to understand the participants' thought processes as they analyzed and indexed documents. Each study is presented in two parts: 1) First, the methods and findings of the study are described and 2) then the how the study informs and relates to the methods for the current study are discussed, with particular attention to identifying research gaps that inform the research questions in current study.

It is important to note that the absence of tagging studies in this section is not an

oversight. Studies of social tagging and folksonomies have emphasized analysis of the tags that users create to understand their behaviors [59, 61, 54, 80] and have not observed taggers in action. One of the novel contributions of the current study is that it observes tagging in action.

2.8.1 Chu & O'Brien: What Document Properties Influence Subject Analysis?

In their 1993 study, Clara Chu and Ann O'Brien [23] argued that the first step in subject indexing, subject analysis, is often overlooked and poorly understood. They were particularly interested in how the properties of a document influence the subject analysis process. To answer their questions about the influence of document properties, Chu and O'Brien asked 104 library science students to write aboutness statements for three articles and to complete a survey about their experience. Chu and O'Brien found that five properties of the articles influenced the subject analysis process and participants' subject analysis performance: 1) the subject discipline of the document (science, social science, or humanities), 2) whether the document was factual or subjective in nature, 3) the complexity of the subject (single subject or multiple subjects), 4) the clarity of the text, and 5) bibliographic apparatus (such as a title, abstract, headings, etc.). This last property, bibliographic apparatus proved to be particularly important, as the participants in the study focused their analysis on these features in the documents.

Chu and O'Brien's study provides research design guidance and leaves open questions relevant to the current study. The design of Chu and O'Brien's study shows that library science students are a viable recruiting pool for study participants and that novices are able to do subject analysis work in an experimental setting, suggesting that the current study can effectively recruit from the locally available pool of library science students, even if they do not have tagging or indexing experience. The notable difference is that the current study asks participants to do subject analysis and subject representation work. Chu and O'Brien's study design also shows the importance of retrospective data collection—in the case of their study, a survey—to understand what participants experienced while doing subject analysis work. The current study takes inspiration from this design by using

a retrospective interview to understand participants' experiences and thought processes. Thirdly, Chu and O'Brien find that bibliographic apparatus influences the subject analysis process; therefore, any study that involves subject analysis should either intentionally treat bibliographic apparatus as a variable or should control for bibliographic apparatus in the documents that are analyzed in the study. The current study takes the latter approach, as described in more detail in [Chapter 4](#). Finally, while the Chu and O'Brien study serves as a useful design model, its emphasis on the subject analysis part of subject indexing means that it does not address translation, subject representation, or vocabulary control. Any study that wants to understand the effect of vocabulary control must include translation and subject representation in its procedures.

2.8.2 Joudrey: How Do Participants Determine Aboutness of Documents?

Daniel Joudrey, in his 2005 dissertation *Building Puzzles and Growing Pearls: A Qualitative Exploration of Determining Aboutness* [79] also focuses on the subject analysis step in the subject indexing process but takes a more qualitative approach than Chu and O'Brien and aims to more deeply understand the subject analysis processes of fewer participants. Joudrey's research questions are: "How do participants determine the aboutness of an item? What activities are involved? What are the observable patterns in the aboutness determination process?" and "What bibliographic, content, or visual features are key to the conceptual analysis process?" These research questions are very much about *processes*. Joudrey uses a think-aloud protocol and a retrospective interview to understand these processes. He asks participants to voice their thoughts and actions while developing aboutness statements for three books and then asks interview questions to gain a deeper understanding of the process of determining what a document is about. Two models of subject analysis emerge in Joudrey's study: 1) a holistic approach where participants' initial perceptions evolve as layers of information are added, which Joudrey terms "pearl growing," and 2) an atomistic approach, which Joudrey terms "puzzle building," where participants look for discrete pieces of information about the document and determine how they fit together.

Joudrey's study provides insight into designing a study about subject analysis but, like

Chu and O'Brien's study, does not explore the full subject indexing process, leaving out translation and subject representation. Also like Chu and O'Brien, Joudrey has success working with novice subject analysts, collecting rich data about their processes, which further suggests that the current study can use library science students who are inexperienced with indexing and tagging as a recruiting pool. Unlike Chu and O'Brien, Joudrey uses a think-aloud-protocol and a retrospective interview and shows that these methods are effective for deeply understanding a person's thoughts and process as they do subject analysis work. Joudrey analyzes the qualitative data collected during the think-aloud protocol and interviews, writes detailed accounts of each participant's process, and finds patterns in these processes. The current study takes cues from Joudrey's approach methods for collecting qualitative data with a think-aloud protocol and retrospective interviews; however, the current study, unlike Joudrey's study analyzes translation, subject representation, and vocabulary control, and uses quantitative and qualitative methods to understand participants' experiences.

2.8.3 Šauperl: How Do Catalogers Determine Subject Descriptions?

In her 1999 dissertation, *Subject Determination During the Cataloging Process*, Alenka Šauperl [116] observes 12 experienced catalogers as they examine and catalog new acquisitions in their libraries. Similar to Joudrey, Šauperl asks participants to think aloud as they work, and then interviews the participants afterward. Šauperl finds that subject determination is a gradual and iterative process that differs for each cataloger, and she documents these complex processes in detailed, individualized flow charts for each participant. Šauperl also finds that experienced catalogers thoroughly examine books, especially difficult-to-catalog books, and then identify tentative headings for the books based on their examination and their knowledge of the indexing language with which they are working. She also finds that experienced catalogers consider document-oriented indexing, request-oriented indexing, and the existing collection at their libraries during subject determination. Finally, Šauperl finds that experienced catalogers are much more likely than inexperienced catalogers to consider subject representation from the start of their process instead of treating subject analysis as

a separate and prerequisite step to subject representation.

Šauperl’s dissertation and her subsequent work [117] give us the most detailed observations of the subject indexing process in a library setting. They also, like Joudrey’s study, show that think-aloud-protocols and interviews are highly effective methods for understanding the process of subject indexing, not just the outcomes. Šauperl’s observation that experienced indexers do not separate subject analysis and subject representation, as Langridge and Lancaster argue they should [86, 85], is particularly relevant to the current study because isolating the effect of different forms of vocabulary control is difficult when the vocabulary control is integrated into the process from the start. Put another way, any study that aims to measure the effect of vocabulary control by measuring the coextensiveness of aboutness statements produced during subject analysis and indexing terms assigned during subject representation—as the current study does—must consider the experience of participants in the study because Šauperl shows that experienced indexers do not separate these steps.

2.8.4 Bertrand & Cellier: How Does Expertise Affect Indexing and What Transformations Occur During Translation?

In their 1995 study, Annick Bertrand and Jean-Marie Cellier set out to “complement ... Chu and O’Brien [who] exclusively centered on the first stage of indexing and on the effects of textual factors” by studying not just that first stage—subject analysis—but also the transformations that occur during translation into indexing languages [12]. Bertrand and Cellier also aim to understand the effect of expertise—both indexing expertise and domain expertise—on indexing behavior. In their study, Bertrand and Cellier ask 25 experienced and novice catalogers to index eight books each (4 in their domain, 4 outside of their domain) while vocalizing their thoughts in a think-aloud protocol. Bertrand and Cellier find that subject analysis and subject representation are largely not coextensive, primarily due to suppression of concepts that are not available in the controlled vocabulary. They also find that experienced indexers spend much more time analyzing documents while novice indexers are more likely to look at the title, abstract, and table of contents briefly and rely more

heavily on the indexing language to determine the subject of documents—suggesting that controlled vocabularies can provide guardrails to keep inexperienced indexers from feeling out of control. Although Bertrand and Cellier, unlike Šauperl, intentionally separate subject analysis and subject representation in their experimental design to study coextensiveness, they still arrive at a similar conclusion to Šauperl—that indexers who are experienced with the indexing language that is being used are more likely to combine analysis and representation steps.

Bertrand and Cellier’s study is particularly instructive for the current study because it not only considers the full subject indexing process but also provides a robust model for evaluating coextensiveness by comparing aboutness statements with subject headings. That model includes a taxonomy of different types of transformations that occur during translation: suppression, simplification, generalization, disassociation, addition, precision, and substitution. This taxonomy of transformations is discussed at length in [Chapter 5](#) and used as a key framework for analyzing the effect of different forms of vocabulary control on coextensiveness. In addition to providing a taxonomy for analyzing coextensiveness between aboutness statements and indexing terms, Bertrand and Cellier reaffirm the effectiveness of think-aloud protocols for understanding the indexing process and skillfully connect quantitative data about subject headings and coextensiveness with qualitative data from the think-aloud protocol—yet another model for the current study to follow. Finally, it is important to note in the context of the current study that while Bertrand and Cellier study more than one indexing language, they do not consider the effect of different forms of vocabulary control on the indexing process.

Chapter 3

RESEARCH QUESTIONS

This chapter summarizes gaps in the existing literature, focusing on previous research that empirically examined subject indexing, presents a high-level research question and three sub-questions that aim to fill those research gaps by examining the effects of different forms of vocabulary control, and describes the motivation for and intended impact of this study.

3.1 Gaps in Existing Literature

While the literature about subject indexing and tagging and the critiques of controlled vocabulary are extensive, only a small collection of studies have empirically examined the processes of subject analysis and subject indexing in action and in depth. This collection of studies, which were highlighted in the previous chapter, serve as models for this study, and the gaps in these studies serve as motivation for the research questions in this chapter. To review, the gaps are as follows:

- Chu and O'Brien and Joudrey studied only the first part of the subject indexing process: subject analysis.
- Šauperl did not compare different forms of vocabulary control and did not isolate subject analysis and subject representation to measure the effect of different forms of vocabulary control.
- Bertrand and Cellier isolated analysis and representation to measure coextensiveness and compared different indexing languages for specific domains but did not consider different forms of vocabulary control, including forms that are common in social tagging environments.
- None of the empirical studies engaged in critical cataloging literature or considered whether different forms of vocabulary control were suitable to ameliorate problems

such as systemic bias. These questions at the intersection of critical cataloging and racial and social justice are discussed extensively in the broader Knowledge Organization literature but not in studies that observe subject indexers or and taggers in action.

3.2 Overarching Research Question

This study aims to fill gaps in knowledge about subject indexing and tagging by considering five different forms of vocabulary control and by connecting an empirical study of the subject indexing and subject tagging processes to issues of racial and social justice. These gaps in knowledge and my research goals motivated one overarching research question that drives the overall study and three related sub-questions that home in on different aspects of the overarching question. The overarching research question of this study is:

How do different forms of vocabulary control affect the subject indexing and subject tagging processes?

Notably, this overarching research question focuses on *processes*—how subject indexers and taggers do their work. It also focuses on the subject indexers and taggers themselves and what they experience while they do that work. Finally, and importantly, it foregrounds forms of vocabulary control as the main variable in this research.

3.3 Three Research Sub-Questions

The overarching research question for this study frames the following specific sub-questions:

- (A) **How do different forms of vocabulary control affect indexers’ and taggers’ *experiences* as they index and tag documents? When do they feel controlled? In control? Out of control?**

These questions, which are referred to collectively as “research sub-question A” throughout this study, consider the effects of vocabulary control from the perspective of indexers and taggers. This study chose to focus on the perspective of indexers and taggers

because they are at the center of the control relationships in Knowledge Organization Systems [73] and because when indexers and taggers feel controlled, that feeling of being controlled often manifests in cataloging decisions and often cascades down the hierarchy of control relationships to information seekers. This research sub-question aimed to develop a more fleshed out understanding of the model of five forms of vocabulary control described in Chapter 2, to better differentiate the properties of these forms of vocabulary control, and to better understand their connections to control relationships and feelings of control.

(B) How do different forms of vocabulary control affect the coextensiveness of aboutness statements and indexing terms?

Research sub-question B is about the relationship between vocabulary control and 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. While Bertrand and Cellier [12] studied coextensiveness, they did not consider the effects of five different forms of vocabulary control on coextensiveness. Understanding this relationship is important to a study about control because coextensiveness is a concrete and measurable expression of how vocabulary control can change or leave unchanged the subject analysis of indexers and taggers as they translate their aboutness statements into subject headings and tags. In other words, coextensiveness is one concrete measure of control. This research sub-question about coextensiveness presents a rich opportunity for both quantitative study (to understand the amount and type of differences between aboutness statements and indexing terms) and qualitative study (to better understand why indexers and taggers make these changes during their process). This study pursues both approaches to answer this research sub-question.

(C) How do indexers' and taggers' experiences while working with different forms of vocabulary control inform criticisms of controlled vocabularies, particularly for documents about racial and social justice? This research sub-question aimed to add to the rich critical cataloging literature which identifies

systemic bias in Knowledge Organization Systems. It is important to note that this research sub-question did not directly ask: Which form of vocabulary control best supports racial and social justice? Instead, inspired by Dartmouth librarians who wanted to help student researchers but felt controlled by the *LCSH*, this research sub-question asked how and when indexers and taggers experienced feelings of control when indexing and tagging documents about racial and social equity. This research sub-question depends on having participants index and tag documents that address racial and social justice, so part of the methodology includes selecting books in this domain and asking participants to reflect on their experiences working with these books.

3.4 Impact: Why These Research Questions Matter

While this study was motivated in part by the gaps in existing research discussed above, the study was also motivated by the impact that I hoped it would have on contemporary Library and Information Science. This study—and its emphasis on control—would not have happened without drawing inspiration from case of the “Illegal aliens” subject heading in the *LCSH* and the lack of control that librarians and information seekers experienced in that case.

This The aspect of this study that was most novel in Knowledge Organization research was its approach to comparing five different forms of vocabulary control instead of studying just one or two forms of vocabulary control. Considering more forms of vocabulary control—some of which have developed in recent years as social tagging has matured—should provide a richer view of how vocabulary control affects subject indexing and subject tagging and should modernize our understanding of subject analysis and subject representation.

I also hope that this study impacts how we view indexers and taggers and reaffirm their importance in KOS. The study intentionally aligns itself with the perspective of indexers and taggers, paying attention to their thoughts and their reactions as they do tagging and indexing work. This alignment was intentionally chosen because indexers and taggers serve a critical middle role between the information seekers that they support and the editors and designers of indexing languages. One goal of this research is that by highlighting indexer

and tagger feelings about control, it might serve to empower indexers and taggers, improve subject analysis and representation processes, and make KOS more just and equitable.

Finally, I think that this study's focus on *people* is, maybe counter-intuitively, more important than ever in an age when machine indexing has become more and more prevalent. Systemic bias in algorithm-based knowledge organization and retrieval is well documented [104] and many machine-based indexing systems use training data provided by human indexers and taggers [63]. What are these human indexers and taggers experiencing while they do their work? Do they feel in control, out of control, or controlled? Do they feel empowered to promote racial and social justice? Or do they feel controlled by an indexing language or algorithm to do the opposite? These questions show that the importance of the research described here is likely to be magnified, not diminished, by the emergence of machine indexing and algorithm-based KOS.

Chapter 4

METHODOLOGY

This chapter describes the methodology used answer the research questions in [Chapter 3](#). This chapter begins by explaining how a Pragmatist research paradigm aligns with my ontological and epistemological views and fits with the research questions in this study. Next, taking cues from Deweyan Pragmatism, the chapter discusses a series of research design questions and my rationale for the research design that I chose. Third, the chapter describes the methods for the study, including the materials, participants, procedures, and data collection methods. the chapter discusses strategies employed to ensure internal, external, and temporal validity, and the limitations of the study.

4.1 Research Paradigm

This study, like much knowledge organization research, took a Pragmatist approach to answering its research questions [\[34\]](#). Because there are many flavors of Pragmatism, including those developed by Charles Sanders Peirce, William James, and John Dewey, this section describes my ontological and epistemological assumptions and then discusses how these assumptions and the socially minded research goals of this study align with Deweyan Pragmatism and its focus on social action. Finally, this section addresses how Dewey's iterative model of inquiry inform the design of this research and how I interpreted the results.

4.1.1 My Ontological and Epistemological Assumptions

All research is informed by the ontological and epistemological assumptions of the researcher. My assumptions and my view of the nature of reality center on helping people and solving problems. If an action helps people and does not cause harm, it is good and it represents a truth that I can incorporate into my worldview and use moving forward. Similarly, if an action solves a problem and does not cause harm, it represents a true solution to that type

of problem and forms part of my view of reality. Based on these assumptions, my research goals are about understanding which approaches to vocabulary control, subject indexing, and subject tagging can help people solve problems and can promote positive social action while minimizing harm.

While my research goals are about finding approaches that help people, that solve problems, and that create social good, I acknowledge that my research will not find universal approaches to achieving these goals or a universal, indelible truth about vocabulary control, subject indexing, and subject tagging. Instead, I believe that truth depends on social and historical context, and what is true in one context is not necessarily true in another context. My approach to understanding truth and reality, then, is to test my beliefs in new contexts. Does what helped before help in this context? Does the solution that I found before solve the problem in this other context? Those epistemological and ontological assumptions mean that I am continuously updating my understanding of what is helpful, what is a solution, and what is real.

Finally, I assume that my continuously updated understanding of reality—based on my experiences in different contexts, my interpretations of those experiences, and my values—is likely different from other people’s reality, and that one reality should not be privileged over another. In my research, I aim to empathize with other people’s realities and to observe their approaches to solving problems as they interact with that reality. I believe observing and understanding others’ perspectives adds to my understanding and makes any findings from my research more robust and more likely to lead to social good.

4.1.2 Dewey’s Instrumentalist Pragmatism

My ontological and epistemological assumptions, as described above, lie adjacent to Constructivism, which is pluralist and assumes that truth and reality are created by people [87]. However, my ontological stance aligns more completely with Pragmatism, especially the Instrumentalist and socially focused Pragmatism that John Dewey introduced as an update to Peirce’s more Objectivist approach to Pragmatism. Deweyan Pragmatism is often described as viewing truth as “what works” and doing “what works” [31]. This pithy

phrase is commonly used to summarize Deweyan (and Jamesian) Pragmatism, and while it captures much of the spirit of Pragmatism, it does not fully describe the philosophical commitments or the research design principles of Pragmatism. This section aims to expand on “what works” by discussing the philosophical assumptions of Deweyan Pragmatism, by illustrating how Deweyan Pragmatism is particularly focused on social action, and by describing how Deweyan Pragmatists establish and constantly renegotiate truth through an iterative model of inquiry.

Philosophical Assumptions of Deweyan Pragmatism

Deweyan Pragmatism re-frames the fundamental ontological question about the nature of reality and asks its own fundamental question: What is the nature of human experience? [31]. This re-framing puts Dewey at odds with established metaphysical views about the nature of reality by focusing on how beliefs about reality are related to human action. Morgan summarizes that relationship in “two inseparable questions: What are the sources of our beliefs? And, what are the meanings of our actions?” [102]. These questions have an inseparable relationship because Deweyan Pragmatism calls on thinkers to reflect on their beliefs to choose actions and then reflect on their actions (and the consequences of those actions) to choose which beliefs constitute truth.

Knowledge Organization scholar Thomas Dousa [34] describes this relationship between belief and action in more detail:

Pragmatists seek to establish knowledge claims with reference to human action in, and experience of, the ambient world—that is to say, to determine which beliefs count as knowledge by considering how they work when put to the empirical test of practice. To put a concept or belief to the test is to inquire about its adequacy in light of experience.

Importantly, this continuous cycle of belief and action shows that Pragmatism views truth as fallible, contingent, and constantly open to reinterpretation and renegotiation based on new evidence. Dewey reinforces this concept of fallibility by using the phrase “warranted assertion” [30] instead of truth claim.

The Pragmatists view of truth is also pluralist and context-dependent, as different people in different contexts might interpret the same actions and evidence differently. Pragmatic pluralism does not mean that “anything goes,” though. Because pragmatists care about the consequences of acting on beliefs, Pragmatists are not free to believe anything that we want about the world. Instead, Pragmatists evaluate the truth in a belief based on “the experiential or practical consequences of its action” [60]. This emphasis on practical consequences aligns well with my view that truth is what helps people and what solves problems.

Deweyan Pragmatism and Social Action

While James and Dewey both developed practically minded and subjectivist varieties of Pragmatism, Dewey’s variety is differentiated by his focus on community life and social action. Dewey was an educator and a progressive activist who viewed inquiry as not just part of academic research but as part of larger social life. Deweyan Pragmatism, therefore, situates research problems in social contexts and aims to solve problems to promote the social good. This focus on the social good aligns with my values, with the goals of this research, and with ameliorating the social injustice observed in cases like the “Illegal aliens” subject headings in the *LCSH*.

While values of Deweyan Pragmatism might seem ideally aligned with pursuing social justice through research, many modern scholars view Dewey’s approach to Pragmatism as overly simplistic and optimistic about social action and, by extension, social justice. Mills [100] criticizes Dewey for relying too much on his personal experience with small-scale political discourse in New England towns. Feminist scholars Seigfried and Whipps and Lake note that while Dewey argues for social justice and supports some Feminist causes, he does so in a dated way that does not account for our contemporary understanding of prejudice and systemic bias [118, 135].

While these criticisms are well founded, Pragmatism is contextual and fallible. Therefore, I believe that Feminist critiques of and improvements upon Pragmatism can be readily adopted as new beliefs in a Pragmatic worldview. As a researcher aiming to solve problems,

help people, and promote social justice in our modern context, I embrace both Dewey's original vision to use inquiry to drive social action and progress and the Feminist updates to that vision that more fully account for systemic bias and diverse viewpoints when identifying "what works."

Dewey's Iterative Model of Inquiry

Deweyan Pragmatism's focus on human experience and social action aligns with how I view the world, and it fits naturally with the study of human systems and processes such as subject indexing, subject tagging, and vocabulary control. This research aims to study these human experiences—these realities—from the perspective of subject indexers and taggers. What do subject indexers and taggers experience while they do their work? When do they feel controlled, in control, and out of control? How does the social, cultural, and historical location of their experience inform their reality?

Dewey provides a concrete model of inquiry for answering these types of questions about human experience, specifically aimed at solving problems and driving social action. [Figure 4.1](#) shows the steps in this iterative model. The current study can be viewed as a single iteration of this model, as described step-by-step below:

1. **Indeterminate situation:** Subject indexers report that they lack control over subject indexing decisions. This situation is evident in critical cataloging literature and in specific cases like the "Illegal aliens" case.
2. **Recognizing the problem:** One view of this problem is that controlled vocabularies are restrictive and control the indexers and catalogers that use them.
3. **Determining a solution:** This study proposed that other forms of vocabulary control, such as those used in tagging, might grant subject indexers and taggers with more control.
4. **Reasoning to develop ideas and evaluate potential actions:** Conducting an experiment that compares the effect of different forms of vocabulary control on the

subject indexing and subject tagging processes.

5. **Experimental proof:** Communicating the finding of the studies and recommending social action, which in this case might be new approaches to vocabulary control in specific contexts.

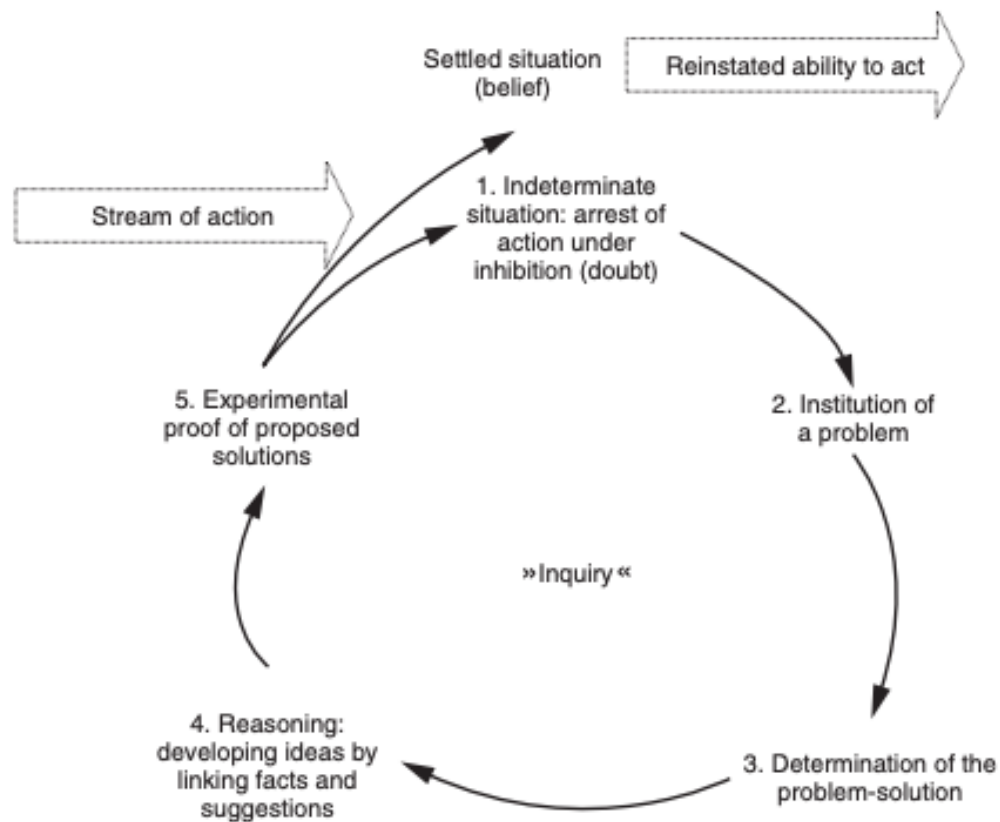


Figure 4.1: Dewey’s five-step iterative model of inquiry, as depicted in Jörg Strübing’s “Research as Pragmatic Problem-solving” [125]. The five steps in Dewey’s model of inquiry are 1) Indeterminate situation, 2) Recognizing the problem, 3) Determining a solution or a possible course of action, 4) Reasoning to develop ideas and evaluate potential actions, and 5) Reaching experimental proof of the proposed actions. If the experiment yields a settled belief, action can move forward until another doubt arises. If not, the process of inquiry repeats.

4.2 *Key Questions that Inform the Study Design*

As David L. Morgan states, thoughtful decision-making is a central tenet of Pragmatism as a research paradigm: “[P]ragmatism emphasizes that all aspects of research inherently involve decisions about which goals are most meaningful and which methods are most appropriate” [102]. Morgan goes on to detail the types of questions that Pragmatism asks of researchers:

How do researchers make choices about the way they do research? Why do they make the decision they do? And what is the impact of making one set of choices rather than another?

This section embraces the Pragmatic spirit of thoughtful decision-making in research design by describing key decisions that I made while designing this study. Each decision involved focusing on the study’s research questions and was informed by my epistemological and ontological assumptions. In many cases, the four studies discussed in [Chapter 2](#) provided helpful guidance on which choices to make. In other cases, I consulted the broader literature, tested my plans in a pilot study, and consulted with advisors to arrive at appropriate design choices.

4.2.1 *Which Documents Should Participants Index and Tag?*

The question of which documents participants should index and tag has two parts. The first part of the question asks: What should the properties of the documents be? For example, should the study use books, articles, or other types of documents? Additionally, should the documents have tables of contents, index, and other bibliographic apparatus? The second part of the question asks: What subject domain should the documents cover?

Chu and O’Brien [23] tell us that the properties of documents (especially bibliographic apparatus) and the “discipline of subject” of documents (i.e., science, humanities) both affect the subject analysis process. Bertrand and Cellier [12] confirm that discipline or subject domain affects the subject indexing process—as does the indexer’s familiarity with the domain covered in the document. Because the research questions for this study do not directly ask about the effects of a document’s properties or subject domain on the indexing

and tagging processes, this study aimed to keep the properties (including bibliographic apparatus) and the subject domain consistent across all the documents chosen for the study.

Properties of the Documents to Index and Tag

Based on the goals of the study, I chose the book format over articles, magazines, or other formats because books are widely indexed and tagged in both libraries and on websites that use social tagging. Books are also a familiar format and have somewhat standard conventions for bibliographic apparatus. For example, study participants could reasonably expect to find a table of contents near the beginning of many books. Within the book format, I chose to use non-fiction books because non-fiction books tend to have more clearly identifiable subjects than fiction books. In the pilot study, I found that the books with minimal bibliographic apparatus were difficult for participants to analyze, index, and tag. This observation matched Chu and O'Brien's findings about bibliographic apparatus and led me to choose books with clear titles, useful summaries, and informative chapter titles and tables of contents to make the subject analysis part of the study less frustrating for participants.

Subject Domain of Documents to Index and Tag

Next, I considered the subject domain for the documents. Subject domain was intimately related the goals of the study because the study situates itself in critical cataloging literature and because subject domain connects directly to research sub-question C, which aims to inform criticisms of controlled vocabularies related to racial and social justice. The choice of subject domain was also important to me personally. My interest in vocabulary control grew out of reading criticisms of systemic bias in indexing languages by Hope Olson [105], Sanford Berman [10], and others [75, 47, 49]. These criticisms show how controlled vocabularies can harm historically marginalized groups and limit the ability of indexers to properly represent these groups and documents about these groups and their culture. As a Pragmatist, I saw these criticisms as problems to be studied, understood, and eventually addressed by proposing informed interventions.

To understand these problems of systemic bias in controlled vocabularies, to answer research sub-question C, and to pursue research in a domain that I feel strongly about, I chose documents that deal with issues of racial justice and social justice. All the documents that were considered for the study address topics like systemic racism, racism in the justice system, violence against people of color, and how to respond to systemic racism as an individual and as a society. These are not easy topics, and the ethical considerations of working with documents about these topics are discussed below. It was important for the design of the study, however, that the documents are not about neutral topics. For example, a study about whether a controlled vocabulary allowed indexers or taggers to use “sofa” instead of “couch” might elicit some discussion, but it would likely not elicit strong feelings of control related to important topics like self-identity and representation of marginalized groups.

Choosing Books for the Study

Based on these research goals, I built a list of candidate books in the domain of racial and social justice, and tested them during the pilot study:

- *Between the World and Me* by Ta-Nehisi Coates
- *Caste: The Origins of Our Discontents* by Isabel Wilkerson
- *The Fire Next Time* by James Baldwin
- *How to Be An Antiracist* by Ibram X. Kendi
- *Just Mercy* by Bryan Stevenson
- *March* by John Lewis, Andrew Aydin, and Nate Powell
- *The New Jim Crow* by Michelle Alexander
- *So You Want to Talk About Race* by Ijeoma Oluo
- *White Fragility* by Robin DiAngelo

While these candidate books all address racial and social justice and are well-known and influential books in the domain, the pilot study showed that some of the books were

not a good match for the study. I eliminated *Between the World and Me* by Ta-Nehisi Coates and *The Fire Next Time* by James Baldwin because they proved difficult to analyze because, due to minimal bibliographic apparatus, they proved difficult to analyze without reading substantial portions of the text. I eliminated *March* by John Lewis, Andrew Aydin, and Nate Powell because, while it is non-fiction, it is narrative-driven and, like *Between the World and Me* and *The Fire Next Time* its subjects are difficult to determine without a thorough read.

Based on the pilot study, I provisionally chose to use the following books in the study:

- *Caste: The Origins of Our Discontents* by Isabel Wilkerson
- *How to Be An Antiracist* by Ibram X. Kendi
- *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* by Michelle Alexander
- *So You Want to Talk About Race* by Ijeoma Oluo

I chose these four books because they had similarly robust bibliographic apparatus and because they directly address issues of race and racism in the United States. They also shared the following similarities:

- Are written by Black American authors who convey their personal experiences as part of the narrative of their books.
- Are written for a general audience with respect to race—not targeted specifically at a Black or white audience
- Are widely read and respected, as evidenced by appearing on best seller lists and by winning awards
- Share a viewpoint that racial injustice is a pervasive and harmful part of American society and history
- Directly address painful injustices related to racism in the United States and other countries

- Offer remedies to racism, racist thinking, and racist systems

Choosing books that share all those similarities would have kept aspects like authorship, audience, and critical perspective from becoming independent variables or confounding factors in the experimental design.

However, choosing books with all those similarities also would have shaped the study in other ways—particularly in relation to the racial identity of study participants. First, if all four books in the study directly address painful injustices related to race from the perspective of Black authors, the work of indexing and tagging those books increases the likelihood that participants—especially participants of color—would suffer emotional harm. Second, having all four books be similar in authorship, audience, and critical perspective would mean that the study could not explore the breadth of indexing and tagging work done in libraries, on the web, and in other collections—or the breadth of challenges faced by indexers and taggers. Third, because this study aimed to interrogate how whiteness in librarianship relates to vocabulary control, I determined that the books in the study should include white perspectives so that the study could address how the predominantly white profession of librarianship addresses not just Black perspectives on race but also other perspectives on race, including those that might be less sympathetic to Black people. Fourth, because the study expected to have predominantly white participants, it was preferred that not all books in the study be written by Black authors, because this would reinforce a recurring and uncomfortable pattern of mostly white indexers and taggers deciding what books by Black authors are about.

Based on this analysis, I ultimately chose the following four books for the study:

- *Caste: The Origins of Our Discontent* by Isabel Wilkerson
- *White Fragility* by Robin DiAngelo
- *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit It* by Richard H. Sander and Stuart Taylor, Jr.
- *The End of Racism: Principles for a Multicultural Society* by Dinesh D'Souza

Caste, the lone holdover from the previous set of books, directly addresses racial injustice and explores the history of racism from the perspective of a Black American author [136]. While I determined that it was problematic to have all books in the study fit this description, including one such book would serve the study well. *Caste* was that book.

White Fragility, the second book chosen for the study, shares some similarities with the previously proposed *So You Want to Talk About Race?* in that it emphasizes the importance of thoughtful conversation in race relations [32]. But *White Fragility* differs in that it is written by a white author and addressed specifically at white readers. Replacing a book written by a Black author with a book written by a white author might seem counter-intuitive in a study that aims to pursue social justice, but it is important to the study for two reasons. First, *White Fragility* does not document the harm of racial injustice from the perspective of a Black author, so it is less likely to cause participants emotional harm. Second, because *White Fragility* addresses race from a white perspective, it helps the study achieve the goal of observing a breadth of indexing and tagging experiences related to race.

Mismatch is one of the most widely read and cited criticisms of affirmative action in higher education [114]. It is written by two white men, one an academic and the other a journalist, and it presents a significantly different perspective on race in comparison to the books previously proposed for this study. *Mismatch* makes its thesis clear in its subtitle, so participants would be unlikely to miss its controversial (and possibly upsetting) perspective, but it also presents its argument in a scholarly and scientific manner that maintains some emotional detachment. *Mismatch* was important to include in the study because it represents a different perspective on race and because it presents an argument that many people would view as not sympathetic to people of color. Both these properties broaden the tagging and indexing experiences in the study.

The End of Racism is a best-selling book written by Indian American immigrant and conservative firebrand Dinesh D'Souza [37]. The title suggests an optimistic path towards an "end of racism" in American society, but the book instead focuses on erasing and minimizing a history of racism and slavery. *The End of Racism* is an intentionally provocative book that was highly controversial when it was first published, and participants were expected to have similarly strong reactions while indexing and tagging it. However, potential

emotional reactions to this book should differ in kind from the empathetic emotional harm that participants might have felt while indexing and tagging a series of four books that all confronted racial harm from the personal perspective of Black authors. Instead, this book presents an antithetical perspective that broadens the coverage of the study.

4.2.2 Is It Ethical to Ask Participants to Tag and Index Books About Charged Topics Like Racial and Social Justice and to Share Their Feelings About the Experience?

One of the goals of this study was to promote racial and social justice in KOS by better understanding indexers' and taggers' experiences as they index and tag books about racial and social justice. While the study's experimental design focused on determining the effect of different forms of vocabulary control on those experiences, it was also critical to consider the effect that a study that involved indexing and tagging books about charged subjects (and then reflecting on that experience) would have on the well-being of participants. Would participants feel safe while working with and reflecting on the content of the books described above? Would participants feel safe sharing their emotions with a researcher whom they do not know? Would participants—especially participants of color—feel safe talking about race, racism, and social justice with a researcher who happened to be a straight, white, cis male? Are the potential benefits of researching vocabulary control and racial and social justice greater than the risks to participants? These questions were difficult to answer definitively, but reflecting on these questions during the design phase and aiming to minimize risk of harm to participants was an important prerequisite to conducting an ethical study.

One framework for reflecting on these ethical questions and for minimizing the risks of harm to research participants and to society comes from the Belmont Report [131]. While the Belmont Report focuses on medical and psychological research and was developed by researchers and practitioners in those fields, it contains generalizable principles that were applied to this study. Listed below are some of these principles and how I addressed them to minimize the risk of causing harm to participants and to maximize the benefits of the study.

- *Respect for persons*: “Respect for persons incorporates at least two ethical convictions:

first, that individuals should be treated as autonomous agents, and second, that persons with diminished autonomy are entitled to protection” [131]. This study respected the autonomy and self-determination of participants by allowing—and encouraging—them to use and act on their own judgment in subject analysis and subject representation tasks and by respecting and never repudiating their determinations or judgments. Part of my messaging during the think-aloud protocol stated that I would not judge or evaluate their analysis or representation choices, but that I would focus my evaluation on the forms of vocabulary control with which they are working. I also openly shared information about the goals and design of the study with participants so that they could make informed choices as they participated in the study. The second part of respect for persons, protecting those with diminished autonomy, primarily considers the elderly, disabled, juvenile, and imprisoned, but can be broadly applied to include people of color—a key demographic for this study—because people of color have historically experienced reduced autonomy and self-determination in American society. Based on this interpretation of the Belmont principles, I paid particular attention to creating a safe space [62] for participants of color.

- *Beneficence and assessment of risk and benefit:* A major aspect of designing an ethical study is maximizing the benefits and minimizing the risk of harm. This study did not involve the type or the degree of risk that medical studies addressed in the Belmont Report involve, but there was still some risk of participants being harmed emotionally or psychologically when they indexed and tagged books about racial and social justice and reflected on that experience. This study also did not have the type or degree of anticipated benefit that a study on, say, a promising life-saving drug might offer; however, the study did present the potential to help us better understand subject indexing and subject tagging from the perspective of indexers and taggers and to promote racial and social equity. The relatively low “stakes” of this study were confirmed when it gained quick IRB approval with exempt status from the UW Human Subjects Division; however, I treated that approval as a decision about legal risk for the University and not necessarily as an ethical endorsement. To ensure that the study maximized

beneficence and was ethically sound, I aimed to maximize the potential benefit by directly addressing one of the most critiqued aspects of controlled vocabularies—the treatment of minority and underrepresented groups in subject indexing. To minimize the risk of harm, I created a safe space for participants—especially participants of color—to do subject indexing and subject tagging work. For example, after determining that they might cause emotional harm, I replaced interview questions from the pilot study that were designed to elicit strong and personal emotional responses with questions that asked participants to engage with less personal concepts like accuracy, specificity, and exhaustivity.

- *Justice and selection of subjects:* The Belmont Report considers justice broadly and then focuses on justice related to the selection of subjects. Broadly, the questions of justice revolve around who the study benefits. This study aimed to benefit society as a whole, but was particularly focused on promoting justice in KOS for groups that have been historically marginalized by these systems. When consider justice more specifically, the Belmont Report discusses who should participate in a study. Why should some groups be included, excluded, or emphasized? How should vulnerable populations be treated during selection? This studied considered these questions at length and determined that the race, ethnicity, and gender of the participant sample should approximate the race, ethnicity, and gender of the American Library Association (ALA) population. This decision meant that this study did not emphasize underrepresented voices, particularly racial and ethnic minorities. However, it did keep people from historically marginalized groups from bearing an undue responsibility for the study and it more likely that study findings could be applied to the current dominance of whiteness in librarianship.
- *Information and comprehension for informed consent:* To ensure that participants were well-informed of the risks of the study, I clearly state in the recruiting materials (see [Appendix A](#)) and the informed consent form (see [Appendix D](#)) that the study would address racial and social justice so that participants know what to expect during the study. I provided a digital copy of the consent form before participants arrived at

the study and gave them ample time to review the consent form before they signed it.

- *Voluntariness*: To ensure that participants did not feel compelled to participate in any part of the study involuntarily, I told participants verbally and in the consent form that the study was voluntary and that they could stop the study at any time or skip any portion of the study with no negative repercussions. For example, I told participants that they could keep their participant gift card just for showing up, even if they opt out before participating in any part of the study. I repeated messaging around voluntariness when racial and social justice and regularly “checked in” on how participants were doing during study sessions. Finally, I remained aware of my status as a member of privileged classes in American society, my position of authority as the researcher, and of my inability to fully understand the emotions that people of color experience while talking about racism and social justice. I listened carefully, deferred to participants’ experience, and provided a safe space for participants to express or not express their emotions as they required.

To summarize, this study posed some risk of participants feeling uncomfortable and experiencing emotional harm; however, the design of this study meant that the risk level was relatively low. I further reduced risk and the possibility of harm by communicating openly, being supportive of participants, and creating a safe space—especially for participants of color. Because the risk of harm was low and because I had plans to mitigate this risk, including having participants opt out of parts or all of the study, I considered it ethical to 1) have humans, including people of color, participate in the study, 2) have participants work with books about charged topics like racial and social justice, and 3) have participants reflect on their experiences, as all three of these elements were fundamental to the design and the potential benefit of the study.

4.2.3 How Many Participants Should the Study Observe?

Determining the appropriate number of participants for a mixed methods study was an inexact process. Unlike exclusively quantitative studies, statistical models do not determine the necessary number of participants to reach significant findings. A major body of liter-

ature [133] recommends that qualitative studies collect data until they reach “saturation,” a concept that Glaser and Strauss developed as part of grounded theory [52] but that has spread to many methods of qualitative research. Saturation is generally taken to mean the point when further data collection will no longer reveal new information or generate new codes [55]; however, the definition and operationalization of saturation varies greatly [115]. Because of these varied definitions, determining the exact number of participants required to reach saturation for this study was not practical.

Braun and Clarke, who conduct research with and write about Thematic Analysis (TA) methods, argue against using saturation as a rationale for determining the sample size for studies. Part of this critique focuses on how the value of saturation is particular to the grounded theory approach to determining meaning and is employed in other qualitative research methods “without good justification” [17]. Furthermore, they argue that “meaning is *generated* through interpretation of, not excavated from, data, and therefore judgements about ‘how many’ data items, and when to stop data collection, are inescapably situated and subjective, and cannot be determined (wholly) in advance of analysis” [18].

Because of this difficulty in determining the appropriate number of participants and because of Braun and Clarke’s critiques of using saturation to predict this number, this study used two types of precedent to set expectations for the number of participants required by the study. First, a meta-study by Mason analyzes 560 qualitative PhD studies and determines that the mean sample size was 31 [96]. This finding does not necessarily mean that 31 is a good number of participants for a primarily qualitative study—the sample sizes seem to be heavily influenced by best practices in quantitative studies and by preferences for round numbers (10, 20, 30, etc.)—but it does offer some guidance because of the volume and variety of studies analyzed and also informs the number of participants required for the quantitative data in this study.

Second, previous studies of subject analysis and subject indexing provide some guidance that more directly relates to the current study. Two of the most similar studies are PhD dissertations by Šauperl [116] and Joudrey [79] that were discussed in [Chapter 2](#). Šauperl studied 12 catalogers indexing documents in their work settings and detailed the experiences of each participant. Following the approach of Šauperl, Joudrey had 12 MLIS students

each write aboutness statements about three documents. Like Šauperl, Joudrey’s analysis included in-depth profiles of each participant. Other studies discussed in [Chapter 2](#) had more participants. Bertrand and Cellier [\[12\]](#), whose approach to studying vocabulary control and coextensiveness are quite similar to the current study, asked 25 indexers to index four documents each and did not write detailed profiles of each participant. Chu and O’Brien [\[23\]](#) studied significantly more participants in their study of subject analysis, having 104 students write one-sentence aboutness statements for three articles. This range of sample size—from 12 to 104 participants—did not set a clear precedent for the correct number of participants; however, the goals, design, and planned depth of data analysis for the current study were most similar to Bertrand and Cellier’s study, which had 25 participants, and Šauperl and Joudrey’s studies, which both had 12 participants, suggesting that the appropriate number of participants might be toward the lower end of the overall range discussed here. Ultimately, the study had 18 participants, which allowed for statistically significant findings for quantitative research questions and, echoing Braun and Clarke, for significant meaning to be generated through interpretation of qualitative analysis.

4.2.4 Should the Study Separate Subject Analysis and Subject Representation?

Langridge and Lancaster argue that subject analysis should be independent of subject representation, and that merging the two steps results in inferior indexing. Langridge states:

[S]ubject analysis is frequently not seen clearly as a distinct activity: it tends to merge with translation and results inevitably suffer. Translation, by definition, is related to a particular scheme of classification on subject headings, but subject analysis is independent of any scheme [\[86\]](#).

Langridge’s argument is borne out in practice by Šauperl, who reports that experienced catalogers mix analysis and representation together in iterative processes [\[116\]](#). Based on Šauperl’s findings, one could argue that separating subject analysis and subject representation in an experimental setting would create an artificial process and compromised results—even if Langridge and Lancaster recommend that approach.

However, Bertrand and Cellier [12] show that separating the analysis and representation steps is critical to understanding the how vocabulary control influences the translation from aboutness statement to subject headings and tags. Like Bertrand and Cellier, the current study asked questions about vocabulary control and coextensiveness. For example: How do different forms of vocabulary control affect the subject indexing and subject tagging processes? and How do different forms of vocabulary control affect the coextensiveness of aboutness statements and indexing terms? To answer these questions, the design of this study separated the subject analysis and subject representation steps to create clear outputs after each step that could be studied and compared, even though that separation required introducing artificial scaffolding to guide the indexing and tagging processes. In practice, that scaffolding involved asking participants to analyze a book and write an aboutness statement without knowing which form of vocabulary control or indexing language that they would then use to represent that book. The scaffolding helped to guide participants, fit Langridge’s and Lancaster’s call for subject analysis to be “independent of any scheme,” allowed for direct observation of the under-observed subject analysis process, and—most importantly—allowed the study to compare distinct outputs of subject analysis and subject representation. However, this decision also meant that results could not necessarily be generalized to the iterative and co-mingled subject indexing that Šauperl and others observe in practice.

4.2.5 Should Participants Be Experienced with Indexing and Tagging?

Šauperl [116, 117] and Bertrand and Cellier [12] show that experienced and inexperienced indexers work differently. Jens-Erik Mai summarizes this difference: “novice indexers rely more heavily on guidelines than expert indexers and ... novice indexers break down the indexing process to sub-processes more often [than expert indexers]” [92]. This tendency of novice indexers to break down the process into sub-processes, or at least to not combine the steps of subject analysis and subject representation as experienced indexers do, makes novice indexers good matches for the research questions in this study. In contrast, experienced indexers would be more likely to resist the design decision that I made to explicitly separate

subject analysis and subject representation in the experimental design to better understand coextensiveness and, thus, would be less likely to isolate the effects of vocabulary control on their processes. In addition to inexperienced indexers and taggers being more likely to work well with this particular study, there was precedent of using similar participants in a similar study, as Joudrey [79] had successfully used novice and advanced beginners as participants in a similar empirical study about subject analysis and subject indexing. Finally, inexperienced indexers were also a good fit for this study for practical reasons, as they were easier to recruit from a pool of potential participants in the University of Washington Masters in Library Science (MLIS) program.

I did not recruit a mix of experienced and novice indexers and taggers as participants because my research questions were not about understanding the effect of expertise on the indexing process. Having settled on just inexperienced indexers and taggers as participants, it is important to describe what “novice” and “advanced beginner” mean. Dreyfus and Dreyfus [36] identify five stages of development for adult skill acquisition: novice, advanced beginner, competence, proficiency, and expertise. Mai describes the novice stage specifically for subject indexing: “The novice indexer has typically read manuals and guidelines and applies these strictly to every indexing situation. True novice indexers are probably only found in introductory classes in library schools” [92]. Mai’s description of the advanced beginner stage is also relevant to this study, as I recruited some advanced beginners in the UW MLIS program and used them as participants in the study: “The advanced beginner has learned the basic rules and indexed some documents and is able to engage in discussions of the indexing of particular documents in a particular system” [92].

4.2.6 Should Participant Demographics Reflect U.S. Librarian Demographics?

Because this study aimed to examine and ameliorate issues related to racial and social justice in subject indexing and tagging, the demographics—particularly the race and ethnicity demographics—of study participants was particularly important. Should the study raise up voices from historically marginalized groups by purposively sampling people of color? Or should the study reflect the current race and ethnicity demographics of people who do

indexing work in the United States so that findings could be more readily applied to the current reality in the field?

In answering these questions, it is useful to understand the demographics of people doing indexing work in the United States. The best and most recent data for understanding who does this work comes from the 2017 ALA (American Library Association) Demographic Study [112]. According to this survey, 81% of ALA members identify as female and 86.7% of members identify as white. In other words, the population of librarians in the United States skews heavily female and heavily white.

The more than 4:1 ratio of female to male librarians in the United States (the ALA survey considered only two genders) is noteworthy; however, gender is not directly related to the race-focused subject matter of the books to be tagged and indexed in the study and is not directly related to the research sub-question about how different forms of vocabulary control inform criticisms of controlled vocabularies related to racial justice. For these reason, it was determined that recruiting a study sample that either reflected the gender composition of the ALA or that was more gender-balanced would suit the study and gender would not be treated as a variable the study design. However,

While gender demographics were not be a particular focus for this study, race and ethnicity demographics were especially important because the books that participants would index and tag focused on racial and social justice and because one of the research questions in the study focused on the relationship between vocabulary control, critical cataloging, and racial justice. Because of this focus, the lack of racial diversity in the library field—86.7% of ALA members identified as white and only 4.4% of members identified as Black or African American in the 2017 ALA Demographics Study—presented an important question: Can a sample of predominantly white participants (which would reflect the librarian population in the United States) produce useful findings about vocabulary control and how books about racial and social justice are represented in indexes and folksonomies?

Answering this question—and determining what findings could be considered “useful”—came down to identifying the goals of the study. One potential goal was to give voice to historically marginalized groups and people with lived experiences closely related to books about racial and social justice by including more people from these marginalized groups in

the study population. While this goal was considered a laudable goal—many studies cite the importance of minority representation in population sampling [27, 4]—it also presented a shortcoming related to my research goals because it would have meant studying an aspirational state for indexing (where historically marginalized voices are emphasized) instead of the current state (where those voices are marginalized). And that current state, where whiteness dominates librarianship and indexing languages, was what raised my research questions to begin with.

To fully understand the goals of this study, it is important to dig deeper into the phenomenon of whiteness in U.S. librarianship. Whiteness is not simply the racial demographics of librarianship, but also the culture, attitudes, practices, and default assumptions in libraries and librarianship [64]. Jody Nyasha Warner [134] argues that whiteness pervades all aspects of librarianship in North America, including collections, staffing, and cataloging. Branche agrees, stating that “Whiteness and white normativity are embedded in U.S. library culture” [16]. While the American Library Association has recognized whiteness as a problem in the profession and called for systemic change [5], Black librarians argue these efforts are unlikely to succeed because they are rooted in whiteness and treat racial disparity in librarianship as a problem to be solved. This approach perpetuates issues around recruitment, retention, identity, and feelings of belonging for librarians of color [22, 50, 64].

While whiteness pervades librarianship, Todd Honma [74] details how this pervasiveness and its impacts largely go unnoticed and can be almost invisible. The impacts of whiteness were often hidden in the past by the view that libraries were neutral and that libraries were passive conduits of information. Honma describes how in more recent decades, whiteness has been hidden by muddled rhetoric about diversity and multiculturalism. Marable [95] criticizes positive sounding but ineffective or even harmful diversity initiatives in LIS: “Instead of just celebrating diversity, we must theorize it, interrogate it” [95]. Espinal et al [40] echo the need to interrogate and directly address race, specifically citing the need to address the role of whiteness in librarianship: “unless we address whiteness, unless we identify and name it, many of the problems that plague us collectively and as individual librarians of color will continue.” These arguments suggest that a study of how subject indexers and taggers represent race while using different forms of vocabulary control must

directly address whiteness, interrogate it, theorize it, and understand the role of whiteness in the experience of indexers and taggers. Addressing this often-invisible reality of the current state of whiteness in librarianship is a prerequisite to future studies that analyze how greater diversity affects the processes of subject indexing and subject tagging and is a critical contribution to how the LIS field understands itself and its relationship to whiteness and race still today. A recent study by Crist and Clark [28] reports that the phenomenon of whiteness in libraries is still pervasive, still harmful, and still requires more critical study.

Therefore, I chose to aim for a participant sampling that roughly matched the race and ethnicity demographics of the ALA to better address and interrogate the role of whiteness in aspects of control in subject indexing and tagging.

4.3 Methods

This section describes how the study was designed and conducted, including sections about the experimental design, recruiting and working with participants, the materials and apparatus that were used, the procedure for each study session, how data was collected, and how the study addressed threats to validity.

4.3.1 Experimental Design

This empirical study employed a mixed-methods approach, combining both qualitative and quantitative data analysis. Mixed methods research is well-aligned with the Pragmatists viewpoint that drove this research and is well-suited to exploring the diverse research questions in this study. Using this mixed methods framework, the two central research methods in this study were: 1) a think-aloud protocol, as developed by Ericsson and Simon [39] and described by Charters [21], and 2) a retrospective interview, which was conducted with each participant after they completed indexing and tagging tasks. The think-aloud protocol was used during the subject indexing and subject tagging portion of the study. In addition to audio recordings of participants' thoughts, this portion of the study generated indexing and tagging data, specifically aboutness statements and subject headings and tags. The retrospective interview (see [Appendix J](#) for the interview script) 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 participants' rankings of the five forms of vocabulary control on different axes of control.

Because the study collected both quantitative and qualitative data aimed at answering multiple research questions, it had a somewhat complex experimental design. The following subsections describe the salient aspects of the experimental design by first discussing those aspects aimed at supporting quantitative data collection and analysis and then by discussing those aspects aimed at supporting qualitative data collection and analysis.

Experimental Design for Quantitative Data

The study used a within-subjects design with one primary independent variable—form of vocabulary control—which had four levels: minimal vocabulary control, passive vocabulary control, flexible vocabulary control, and rigid vocabulary control. Each participant was exposed to all four forms of vocabulary control. In addition to these four levels, which were tested throughout the study, all participants were asked interview questions about a fifth form of vocabulary control: post hoc vocabulary control. ¹

There were multiple dependent variables in the study. First there were rankings, across three different scales describing control and ease of use, which participants provided during the retrospective interview. The quantitative analysis of these ranking data were a primary focus in answering research sub-question A, which asked about how vocabulary control affected participants' feeling of being in control, out of control, and controlled. Next, the study tracked and counted the number of concepts in participants' aboutness statements, the number of subject headings that participants provided, and the coextensiveness of these aboutness statements with subject headings and tags, and the types of transformations that occurred during translation from aboutness statement to indexing terms. These data were analyzed as dependent variables to understand research sub-question B, which asked about the effect of vocabulary control on the subject indexing and subject tagging processes

¹Post hoc control was studied via interview questions based on feedback received during the pilot study. Pilot study participants indicated that the study sessions were too long with five forms of vocabulary control and that the post hoc control treatment was largely the same as minimal control with a series of questions after the participant had finished tagging the document.

and on the coextensiveness of subject analysis and subject representation. In addition to these two main sets of dependent variables that were used for quantitative analysis—rankings to describe feeling of control and ease of use and aboutness statement, indexing term, and coextensiveness data—the study also collected and analyzed quantitative data about the time to complete tasks with different forms of vocabulary control and collected, coded, and counted participants' use of bibliographic apparatus during subject analysis and subject representation. The bibliographic apparatus data were analyzed quantitatively and qualitatively; see [Appendix K](#) for the bibliographic apparatus codebook.

The study did not treat participant demographics (age, gender, indexing experience, domain knowledge, etc.) as independent variables for quantitative analysis, for reasons described in section below about participant sampling, recruiting, and demographics. The study also did not treat the four books that all participants tagged or indexed as independent variables, as they were chosen because they were in a similar domain and had similar bibliographic apparatus. Finally, the study was designed to minimize order effects through randomization, as described in the section about threats to validity, so order was not treated as an independent variable for quantitative analysis. Qualitative analysis did, however, produce some discussion in which participants reported that they learned to be more efficient with subject analysis during their study session.

Experimental Design for Qualitative Data

The study employed both quantitative and qualitative methods, with the qualitative methods aimed at more fully understanding and describing participants' experiences as they indexed and tagged books. When and why were they confused or frustrated? When and why were they satisfied? What worked well for them and why? When were they able to assign the indexing terms that they wanted to assign and why? When were they able to assign accurate, specific, and exhaustive indexing terms and why? Answering these types of qualitative questions was critical to understanding participants' *experiences* working with different forms of vocabulary control and to getting at the subjective types of data that described those experiences.

Much of the qualitative data came from participants' answers to questions in the retrospective interview, which was designed to have participants reflect on their experiences and explain their views of those experiences, and to add context to the quantitative analysis in the study. For example, the ranking portions of the interviews that were described above also included questions about why participants chose the rankings that they did, resulting in responses that were analyzed for themes and that are quoted in the Results and Discussion chapter. Similarly, participants' reflections of the coextensiveness of their aboutness statements and indexing terms were used to contextualize, explain, and expand upon the quantitative analysis of transformations. Responses to other interview questions—including those about specificity, exhaustivity, and accuracy—and participants' thoughts and reactions that were captured during the think-aloud protocol were also analyzed for themes, and key quotes are presented in the Results and Discussion chapter to represent those themes.

Research sub-question C, which focused on criticisms of vocabulary control and racial and social justice was particularly geared toward qualitative data because it is difficult to quantify how well a form of vocabulary control supports racial and social justice. Instead of trying to quantify these properties, this study identified moments when participants addressed the role of race and social issues in their subject indexing and subject tagging processes—moments when participants discussed how forms of vocabulary control supported or did not support how they wanted to represent race and social issues—and discussed the themes that emerged from these moments in the context of critical cataloging.

By using this mixed methods approach and by emphasizing the thoughts and words of participants in qualitative analysis, this study aimed to reach a deep understanding of inexperienced indexers' and taggers' experiences working with different forms of vocabulary, feeling different types of control, and thinking about and reacting to the role of vocabulary control in promoting or harming racial and social justice.

4.3.2 Participants, Sampling Methods, and Recruiting

The study was conducted in two phases with different participant pools for each phase. The first phase was a pilot study that tested the materials, procedures, and data collection

for the study. Participants in the pilot study ($n=2$) were PhD students in the University of Washington Information School who were studying knowledge organization and related topics. These participants had somewhat similar academic backgrounds to the main study participants and were particularly helpful with providing feedback during the pilot study. The data from the pilot study was not used in the analysis for this study; however, because the profiles of the pilot participants were sufficiently different from the main study participant profiles and because the study design changed after the pilot study.

The second phase—the main study—involved many more participants ($n=18$). To ensure that subject analysis and subject representation were treated as separate steps, the study recruited participants who were informed novices and advanced beginners with minimal experience in indexing and tagging. These participants were recruited from the University of Washington Masters of Library and Information Science (MLIS) program, a population that was both convenient and a good match for the level of indexing and tagging experience required for the design of the study.

Recruiting Participants

Participants for the main were recruited from the University of Washington Masters of Library and Information Science (MLIS) program, a population that was both convenient and well suited to the goals of the study. Participants were recruited through MLIS student mailing lists and through visits to MLIS class sessions to advertise the study. Recruiting materials explained that the study would be about subject indexing, tagging, and vocabulary control. The recruiting materials framed participation in the study as an opportunity to learn more about these topics through hands-on practice. Recruiting materials noted the study's connection with racial and social justice. Finally, recruiting materials told participants that they would receive a \$20 gift card. See [Appendix A](#) for the text of two recruiting emails.

Participant Demographics: Gender, Race and Ethnicity, and Age

Participants were recruited from a large Library and Information Science program—without targeting specific demographics within the program—with the intent of recruiting a participant pool that somewhat resembled the demographics of the members of the American Library Association (ALA). This section describes the gender, race and ethnicity, and age demographics of study participants and compares these demographics to the demographics of ALA members.

Each participant in the study completed a demographic survey (see [Appendix B](#)) in which they indicated their gender, race and ethnicity, and age. Much like the members of the American Library Association (ALA), participants in the study identified as predominantly white and predominantly female. The 2017 ALA Demographic Survey reports that 81% of ALA members identify as “Female” and that 86.7% of ALA members identify as “White” [\[112\]](#). Similar to ALA membership demographics, 77.8% of participants in the study identified as “Female,” “Woman,” “Cis Female,” or “Ciswoman.” However, study participants were not as overwhelmingly white as the ALA membership. While most study participants self-identified as “White” or “European-American,” 38.9% of participants in the study self-identified as members of non-white or Hispanic racial and ethnic groups. As expected for a study that intentionally recruited novice and advanced beginning indexers and taggers, study participants were younger overall (median age range of 25-29) than members of the ALA (only 20.9% of members in the 25-34 age range).

[Figure 4.2](#) summarizes the racial and ethnic identities of all participants in the study. It is worth noting that while study participants were not as white overall as ALA members, the percentage of study participants who identified as “Black or African American” (5.6%) was roughly equal to the percentage of ALA members who identified similarly (4.0%). It is also important to note that the percentages displayed in the bar chart add up to more than 100%. That seeming discrepancy is accounted for by the survey question design. Participants could self-identify with multiple races or ethnicities, and one participant responded with both “White” and “Hispanic, Latino, or Spanish origin.”

Study participants were asked to “Enter any gender or genders you identify with.”

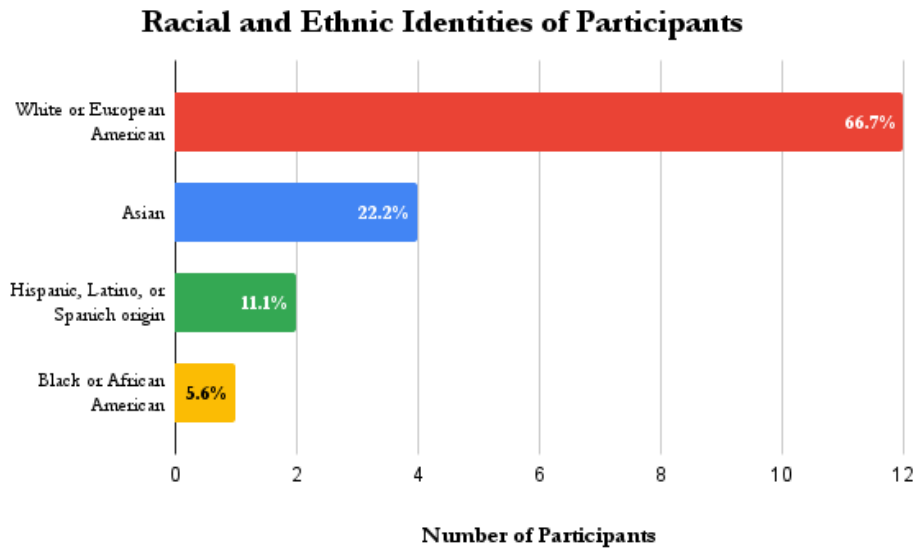


Figure 4.2: A bar chart showing the racial and ethnic identities of study participants. Like ALA members, study participants were predominantly white, although not as overwhelmingly white as the ALA membership. Note that the percentages in the bar chart add up to more than 100% because participants could self-identify as more than one race or ethnicity.

Figure 4.3 summarizes the responses for all participants in the study. As with the chart for racial and ethnic identities, the chart for gender identities adds up to more than 100% because participants could self-identify as more than one gender. Like ALA members, study participants identified predominantly as female (77.8%). The gender demographics of study participants differed from ALA members for participants who identified as non-binary, queer, or genderqueer (27.8%) and for participants who identified as male (5.6%). The ALA Demographic Study, which assumed that gender was binary, reported that 19% of ALA members identified as “Male.”

Finally, Figure 4.4 summarizes the age ranges of participants. All but two participants were between the ages of 21 and 34, and the median age range was 25-29, which accounted for 61.1% of total participants. Study participants were much younger than ALA members, which was not surprising because ALA membership spans the full career cycle while the novice and advanced beginner indexers who were recruited for the study were all at or near

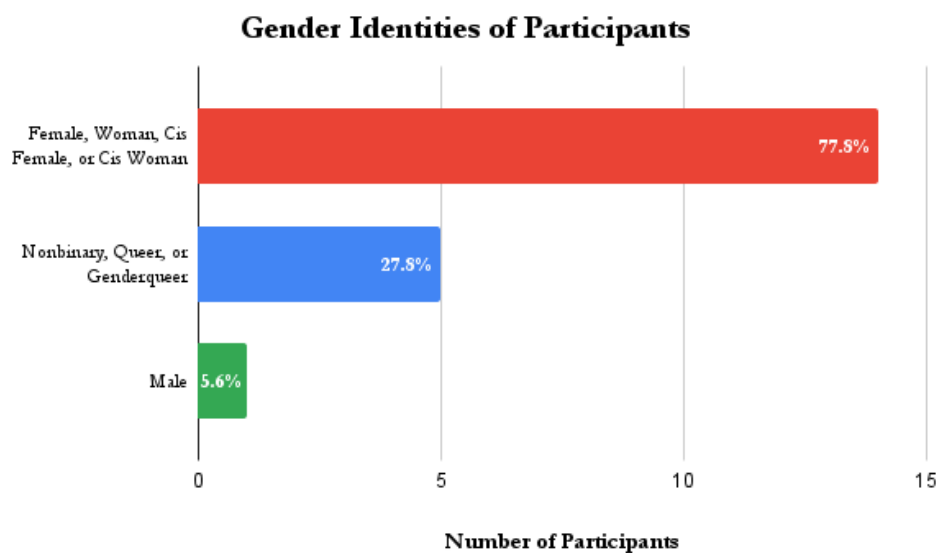


Figure 4.3: A bar chart showing the gender identities of study participants. Like ALA members, study participants were predominantly female. Note that the percentages in the bar chart add up to more than 100% because participants could self-identify as more than one gender.

the beginning of their careers.

Participant Demographics: Education and Experience Related to Indexing and Tagging

While this study aimed to recruit participants that resembled the race, ethnicity, and gender demographics of the ALA, the study did not want to recruit participants with the range of experience present in the ALA population. Instead, the design of the study, with a clear separation between subject analysis and subject representation, called for participants who were novice or advanced beginner indexers and taggers. I determined participants' level of experience—whether they were categorized as novices or advanced beginners—based on their educational experience and their experience working with subject indexing and tagging. For a participant to be categorized as an advanced beginner, they needed to have taken multiple courses related to indexing and tagging and to have some experience with or exposure to indexing and tagging outside of their coursework. Other participants, with

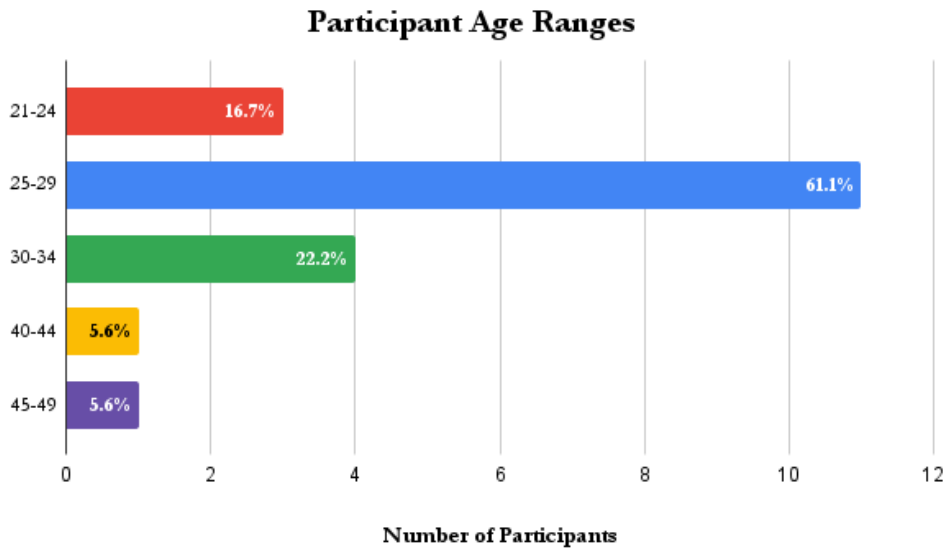


Figure 4.4: A bar chart showing the age ranges of study participants. The median age range of participants was 25-29 years old.

one exception, were categorized as novices.

To understand participants' educational background in indexing and tagging, participants were asked in a survey to "list any post-secondary degrees or certificates that [they] have completed and any academic programs in which you are currently enrolled" and were also asked to indicate specific courses that they had completed or were currently taking in the UW MLIS program. Participants reported undergraduate degrees in a wide variety of disciplines, including English, History, Computer Science, Music, Psychology, Communication, and Anthropology. None of these degrees indicated undergraduate training in indexing or tagging, so I focused on participants' MLIS coursework. I asked about many of the courses in the program, but the two courses that were most relevant to this study were 1) LIS 530: Organization of Information in Resources, which provides students a theoretical and practical introduction to Knowledge Organization, and 2) LIS 531: Catalogs, Cataloging, and Classification, which is a more advanced course that teaches students the details of cataloging, including subject indexing. [Figure 4.5](#) shows how many participants

had completed, were currently enrolled in, and had not taken the Knowledge Organization course, LIS 530.

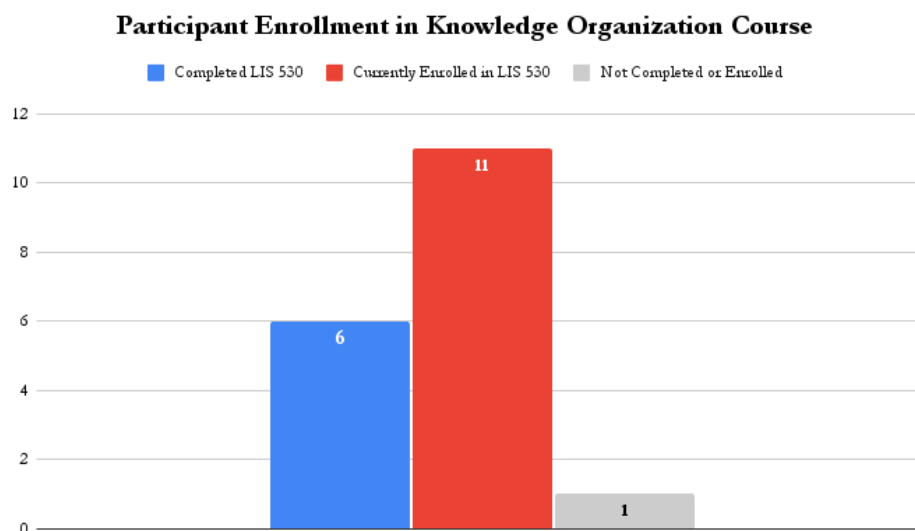


Figure 4.5: A bar chart showing how many participants had completed, were currently enrolled in, or were not enrolled in or had not completed a Knowledge Organization course, LIS 530: Organization of Information and Resources.

[Figure 4.6](#) shows how many participants had completed or had not taken the cataloging course, LIS 531.

Based on these responses, I provisionally categorized six of the participants as advanced beginners and 12 of the participants as novices. The two participants who were most difficult to categorize had completed LIS 530 but not LIS 531. I noted that both these participants had completed a course in constructing indexing languages, so provisionally categorized them with the advanced beginners.

Next, the survey asked participants to “briefly describe [their] experience with indexing, tagging, and cataloging.” Most participants (n=12) indicated that they did not have experience with tagging or indexing outside of coursework. Responses included: “Absolutely none!” or “little to no experience.” A smaller group of participants (n=5) indicated some exposure to or experience with cataloging, tagging, and indexing languages. For example: “I

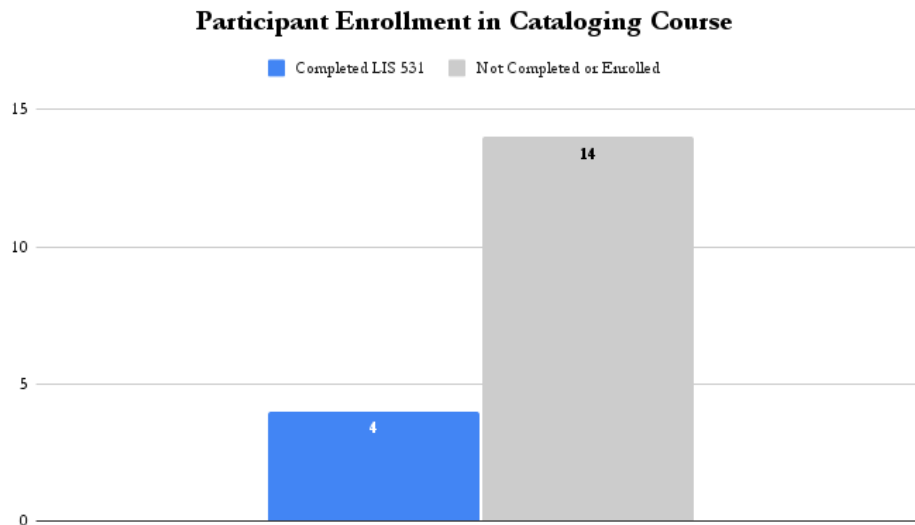


Figure 4.6: A bar chart showing how many participants had completed or were not enrolled in or had not completed a Cataloging course, LIS 531: Catalogs, Cataloging, and Classification.

currently work in an academic library so have light experience with *LCSH*” and “I scanned and tagged archival materials for my undergraduate special collections.” One participant indicated two years of experience with cataloging.

Based on the criteria for categorizing a participant as an advanced beginner—that the participant had completed multiple courses in indexing or tagging and had some experience with indexing or tagging outside their coursework—I cross-referenced the responses to the questions about LIS courses and the responses to indexing and tagging experience outside of coursework. I found that only two participants (Participant 11 and Participant 21) met both criteria and categorized them as advanced beginners. I also noted that while Participant 10 had taken only one course directly related to indexing and tagging, they had two years of experience working with cataloging. Therefore, I included Participant 10 in the advanced beginner category despite not strictly meeting the criteria.

[Figure 4.7](#) summarizes how many participants were categorized as novices and advanced beginners. Because the advanced beginner group was so small and because the difference in

indexing and tagging experience between novices and advanced beginners was not dramatic, this study did not treat experience level as a variable in the study. Instead, this demographic data about participants' education and experience was used to confirm that participants fit the profile required by the experimental design for this study.

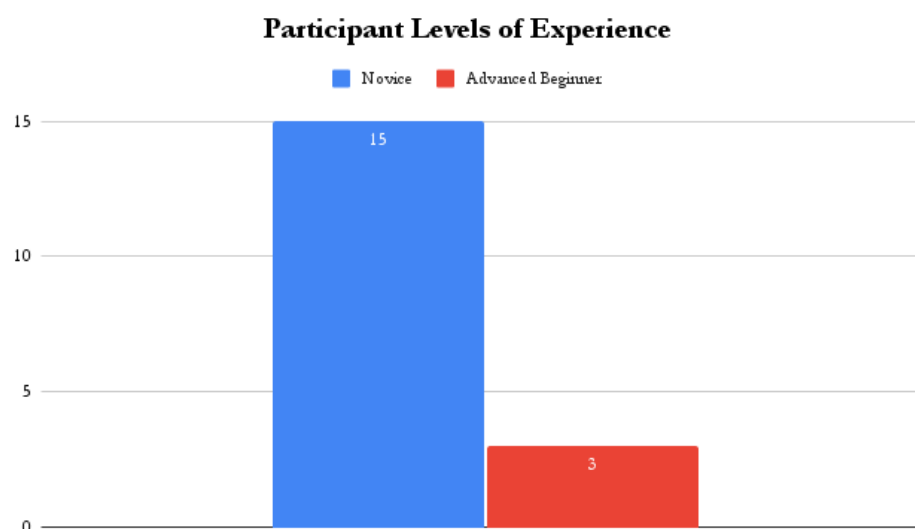


Figure 4.7: A bar chart showing how many participants were categorized as novice indexers and taggers and how many participants were categorized as advanced beginner indexers and taggers.

Participant Demographics: Familiarity with Racial and Social Justice Literature

Bertrand and Cellier find in their study that domain expertise affects indexing behavior and proficiency [12]. This finding agrees with the literature on domain analysis, which “focuses on the importance of subject knowledge” [68]. Applied directly to the current study, one might expect participants with more domain expertise (or subject knowledge) in literature about racial and social justice to exhibit different behaviors than those who do not while tagging and indexing books about racial and social justice. To determine participants' familiarity with and expertise in racial and social justice, the survey asked participants to indicate which of a variety of books they had read or were familiar with or were not familiar

with. This series of questions focused on literature instead of participants' overall familiarity with the domain of racial and social justice because questions about the domain as a whole might have devolved into some form of "How antiracist are you?" and lead to unreliable data and because literature was the most relevant aspect of social and racial justice to this study. The books listed in the survey included books about racial and social justice that were used in the study, books about racial and social justice that were not used in the study, and books about other topics. (The full list is available in [Appendix B](#).) Notably, only two participants had read any of the books that were used in the study. Participant 25 had read *White Fragility* and *Caste*. Participant 21 had read *White Fragility*. None of the participants had read *Mismatch* or *The End of Racism*. Furthermore, none of the participants indicated that they had heard of *Mismatch* or its authors.

I scored the participant's response for each book based on the values in [Table 4.1](#) and then summed up the total score to calculate an overall domain familiarity score for each participant.

	"I have read this book or something else by this author."	"I have heard of this book or author and might know a bit about them."	"I have not heard of this book or author or I am very unfamiliar with them."
Book in the domain and in the study	2	1	0
Book in the domain but not in the study	2	1	0
Book not in the domain or the study	0	0	0

Table 4.1: Matrix for calculating participants' familiarity with literature about racial and social justice. Books that were not in the domain of racial and social justice did not count toward participants' familiarity score.

[Figure 4.8](#) shows the distribution of calculated scores for all participants. The minimum score was one, which was calculated for a participant who indicated that they were familiar with one and had read none of the 15 books about racial and social justice asked about in the survey. The maximum score was 17, which was calculated for a participant who

indicated that they were familiar with seven and had read five of the 15 books. The mean score was 9.39, the median was 10, and the standard deviation was 4.64. The scores were spread broadly across the range, although they had a somewhat bimodal distribution with a gap between 6 and 9, with more of the scores falling to the higher side of that split.

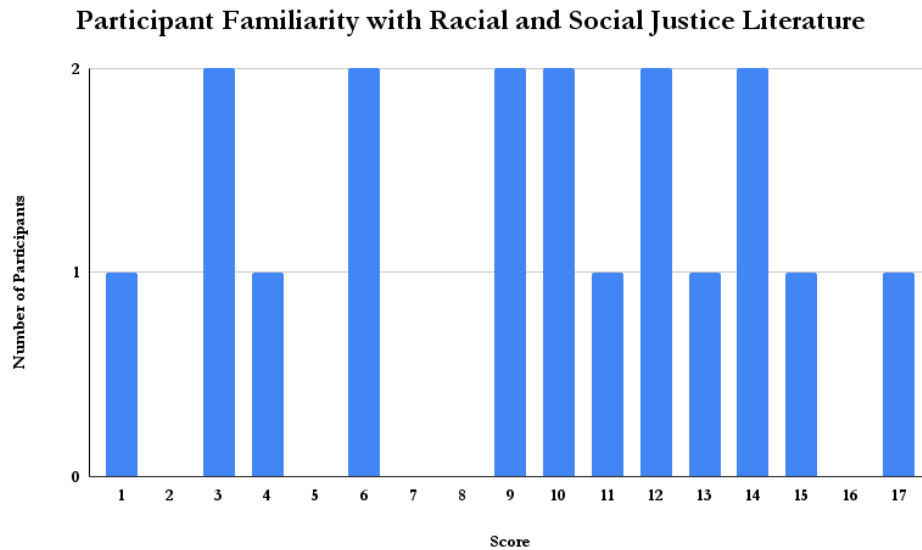


Figure 4.8: A bar chart showing the distribution of scores for participants' familiarity with racial and social justice literature. Higher scores indicate more familiarity. For example, the score of 1 is for a participant who was familiar with one book and had read none of the 15 books about racial and social justice asked about in the survey. The score of 17 is for a participant who had read five of the books and was familiar with another seven of the 15 books.

4.3.3 *Experimental Materials and Apparatus*

The study was conducted in-person in a conference room at the University of Washington. [Figure 4.9](#) shows many of the materials and apparatus that were used in the study as they were set out on the conference room table before the start of a study session. This section describes the materials and apparatus that were used in the study, including scripts and forms, data collection apparatus, and books and materials that participants worked with.



Figure 4.9: The study materials and apparatus as they were arranged before a participant arrived for a study session. The researcher sat in front of the script to the right in the photograph. The participants sat in the chair behind the laptop computer. This arrangement of the researcher sitting to the side of participants was determined during the pilot study to be good for making the researcher available but not directly in the participants' line of sight and not looking over the participants' shoulder, both of which were considered distracting.

Demographic Survey

See [Appendix B](#) for the survey questions and <https://forms.gle/JfN3n8nVhRAx97cy8> for the survey form. The survey collected data about participants' age, gender, race and ethnicity, education, experience with indexing and tagging, and familiarity with racial and social justice literature. Each survey response was connected to a Participant ID; the survey did not collect the participant's name or other personally identifiable information.

Researcher Scripts

I used scripts to maintain consistency when I introduced the overall study and when I explained each subsection of the study to participants. See [Appendix C](#), [Appendix E](#),

[Appendix F](#), [Appendix G](#), [Appendix H](#), and [Appendix I](#) for the scripts that I used in the study.

Consent Form

See [Appendix D](#). Participants received a digital copy of the consent form when they signed up for a study session and reviewed and signed a paper copy at the beginning of the study session.

Writing Materials

I provided paper, pencils, and pens for participants to write notes and to write down their aboutness statements.

Audio Recorder

I used a digital audio recorder to record the think-aloud protocol and the interview. The specific model of audio recorder was chosen because it was small and unobtrusive.

Books

The study asked participants to index and tag four books, which are discussed in detail above:

- *Caste: The Origins of Our Discontents* by Isabel Wilkerson
- *White Fragility: Why It's So Hard for White People to Talk About Racism* by Robin DiAngelo
- *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit* by Richard Sander and Stuart Taylor Jr.
- *The End of Racism: Principles for a Multiracial Society* by Dinesh D'Souza

Controlled Vocabularies

I provided participants with a print copy of the *Sears List of Subject Headings* and directed participants to access the Library of Congress Authorities at <https://authorities.loc.>

gov when they were working with flexible vocabulary control and rigid vocabulary control, respectively.

Web Application

I built a custom-made web application hosted at <https://students.washington.edu/cholstro/sio2/ready.php> for participants to enter their tags and subject headings. The user interface (UI) for the web application, shown in [Figure 4.10](#), presented participants with the correct book and the correct form of vocabulary control for each of the four indexing and tagging activities in the think-aloud protocol. A MySQL back-end collected and stored the subject headings and tags for each book. The UI for the web application was similar for each of the four forms of vocabulary control, with the biggest difference being the autocomplete feature which was enabled for passive vocabulary control only.

The web application had two limitations, which I explained to participants verbally. First, the web application did not have a delete function. If a participant wanted to delete a tag or subject heading, I asked them to verbalize their intent during the study and removed the tag or heading during analysis. (Participants verbally communicated their intent to delete a tag or subject heading three times during the study.) Second, the web application did not have automated error checking or other means of enforcing authorized headings for the *Sears List* or the *LCSH*. I strongly encouraged participants to work exclusively from authorized headings when working with the Library of Congress Authorities to prevent the use of unauthorized headings.

In addition to the participant UI, the web application had an administrator UI hosted at <https://students.washington.edu/cholstro/sio2/admin.html> and shown in [Figure 4.11](#). I used the administrator UI before each study session to configure the order of the books that the participant would tag and index and the order of the forms of vocabulary control that they would use. Participants could not see the configured order in the web application until they reached the appropriate stage in the study.

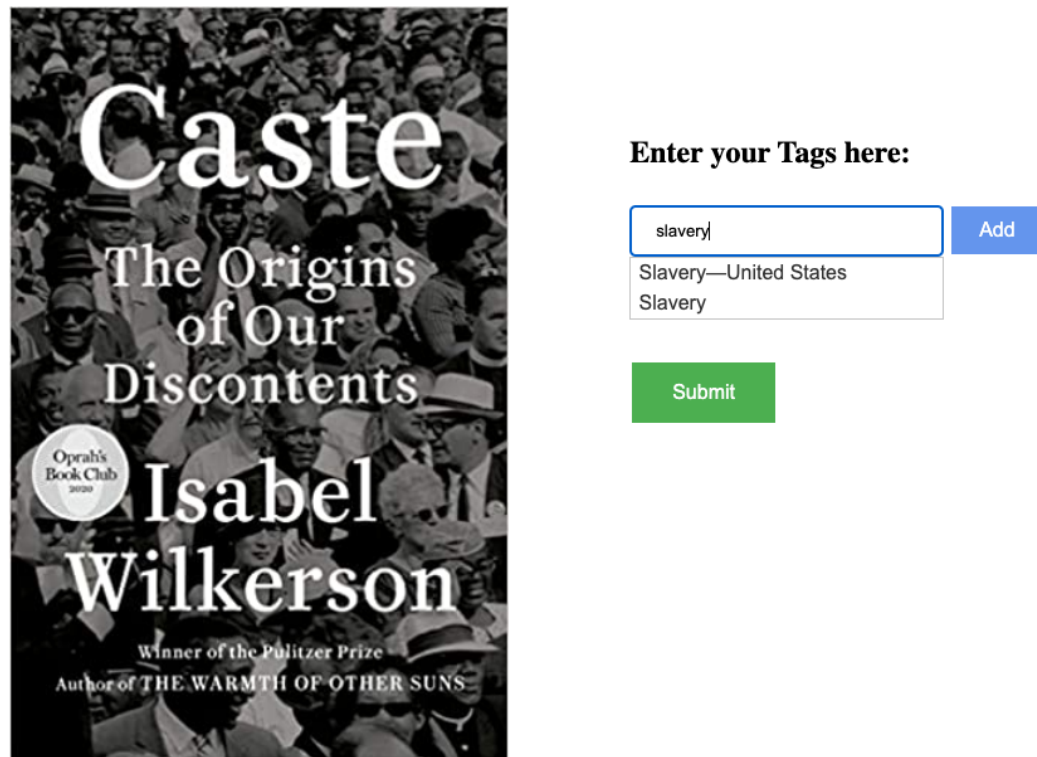


Figure 4.10: The UI for collecting participants’ tags and subject headings during the think-aloud protocol. This screenshot shows passive vocabulary control, as the word “Tags” is displayed instead of “Subject Headings” and the autocomplete feature is enabled. The autocomplete vocabulary included roughly 500 relevant indexing terms taken from the *Sears List of Subject Headings*.

Interview Questions

See the full interview script at [Appendix J](#). I developed a set of core questions for the retrospective interview that asked participants about their experience indexing and tagging books. The interview questions asked participants to identify which books were easiest and most difficult to tag or index; to address the concepts of accuracy, specificity, exhaustivity, and coextensiveness; and to rank the different forms of vocabulary control on three scales related to control and ease-of-use. These ranking were done with the note cards described below. As noted in the interview script, I randomized the order of the ranking questions to avoid order effects.

Hello! This is the admin page that the researcher uses to set up each study session.

If you're here by mistake, let the researcher know. If you're the researcher, you are awesome!

User ID:

	Book	Condition
1	<input type="text" value="Caste"/>	<input type="text" value="Flexible Control"/>
2	<input type="text" value="White Fragility"/>	<input type="text" value="Passive Control"/>
3	<input type="text" value="Mismatch"/>	<input type="text" value="Free Tagging"/>
4	<input type="text" value="The End of Racism"/>	<input type="text" value="Rigid Control"/>

Figure 4.11: The administrator UI for the experimental apparatus. At the start of each study session, I used the administrator UI to enter the User ID and to configure the order of books that the participants would tag and index and the order of the forms of vocabulary control that they would use.

Ranking Note Cards

I used a set of five note cards for participants to rank the different forms of vocabulary control from easiest-to-use to hardest-to-use, from making them feel most in control to most controlled, and from making them feel most in control to most out of control. The note cards were labeled as: Free Tagging, Autocomplete Tagging, Tagging with Editors, *Sears List of Subject Headings*, and *Library of Congress Subject Headings*. I first asked participants to rank the four forms of vocabulary control that they worked with during the subject representation part of the study. I then introduced post hoc control and asked participants to revisit their original rankings and add the tagging with Editors card to their rankings.

4.3.4 Procedures

The study was conducted in three parts: 1) a survey to collect demographic data and information about participants' education and experience, 2) a think-aloud protocol—also known as a cognitive interview—to capture participants' thoughts while they determined

the subjects of four documents, and 3) a retrospective interview that was designed to help participants reflect on the subject indexing tasks that they had just completed. Study sessions were conducted in a conference room at the University of Washington, Seattle campus. Each session lasted approximately one-and-a-half to two hours.

Research sessions were conducted with one participant at a time and used the following procedure for each participant:

1. **Demographic Survey:** Before the participant arrived at the study, I sent them a link to a demographic survey by email. The participant provided basic demographic data (age, gender, race, and ethnicity), education and experience data, and data about their familiarity with books and authors related to racial and social justice. See for [Appendix B](#) the demographic survey.
2. **Welcome and Consent:** When the participant arrived for the study, I briefed them on the study, including how long it would take, what the participant would do, and the goals and background of the study. I also asked them to complete a consent form. See [Appendix C](#) for the researcher script for this introduction and [Appendix D](#) for the consent form.
3. **Subject Analysis:** I handed the participant a book to index or tag. The study used four books, each given to the participant directly before they indexed or tagged that book. The books were all nonfiction books that address race and society, primarily in the United States:

- *Caste: The Origins of Our Discontents* by Isabel Wilkerson
- *White Fragility: Why It's So Hard for White People to Talk About Racism* by Robin DiAngelo
- *Mismatch: How Affirmative Action Hurts Students It's Intended to Help, and Why Universities Won't Admit* by Richard Sander and Stuart Taylor Jr.
- *The End of Racism: Principles for a Multiracial Society* by Dinesh D'Souza

When I gave the participant each book, I asked them to analyze the book to determine its subject or subjects and instructed them to write an aboutness statement that

summarized what the book was about. This portion of the study used a think-aloud protocol, so participants verbally described what they were doing, thinking, and feeling as they determined what the books were about. These verbalized thoughts were captured by the audio recorder along. I also took hand-written notes to capture non-verbal data like facial expressions and posture and to track which parts of the bibliographic apparatus participants used.

4. **Subject Representation:** After the participant indicated that they were satisfied with their aboutness statement, I introduced the form of vocabulary control that they would use to tag or index that book. I also provided them with the appropriate materials and user interface for that form of vocabulary control. For example, when I instructed participants to use flexible control, I told them that they were using the *Sears List*, handed them a copy of the *Sears List of Subject Headings*, and directed them to the web application to enter their subject headings. When I instructed participants to use passive control, I told them that they were using autocomplete tagging and directed them to the web application with the autocomplete feature enabled.

I 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. I also explained the “rules” of working with the form of vocabulary control presented to the participant. For example, for participants working with rigid vocabulary control, I explained that they could use only authorized headings from the *LCSH*.

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

- **Minimal control:** Free tagging with no controls and minimal guidance.
- **Passive control:** Tagging with suggested vocabulary that became visible via autocomplete as users entered text in the UI.
- **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.

- **Rigid control:** Subject indexing with the *LCSH*, which allows only indexing terms that are already in the indexing language. Participants used the web interface at <https://authorities.loc.gov/> to find authorized subject headings.

For all forms of vocabulary control, the participant entered the subjects or tags in the provided web application. As with the subject analysis part, this subject representation part of the study used a think-aloud protocol with audio recording and hand-written researcher notes. See [Appendix E](#), [Appendix F](#), [Appendix G](#), and [Appendix H](#) for the researcher scripts for introducing each form of vocabulary control.

5. **Repeat:** Steps 3 and 4—subject analysis and subject representation—were repeated for each form of vocabulary control. The form of vocabulary control and the books were presented to participants in random order to minimize ordering effects.
6. **Break:** The participants were given an opportunity to take a short break to get a drink of water or to use the restroom.
7. **Interview:** I interviewed each participant, asking questions about their experiences during the indexing and tagging portion of the study. As with the previous portions of the study, the interview used audio recording to capture participant responses. See [Appendix J](#) for the base interview script. I asked follow-up questions based on the participants' answers and the events and comments from their subject indexing and subject tagging work.
8. **Wrap-up:** I gave each participant a chance to ask any questions that they had about the study, reminded them that their information would remain private, and thanked them for participating in the study.

4.3.5 Data Collection

The study collected data through three methods: 1) a demographic survey, 2) a think-aloud protocol while participants did subject indexing and subject tagging work, and 3) a retrospective interview. The data collected was both qualitative and quantitative. The different types of data that were collected are described below.

Demographic Data

I collected demographic data from participants via a survey that they completed before arriving at their study session. The survey collected the following data about each participant:

- Age, gender, and race and ethnicity
- Educational background and level of experience with indexing and tagging
- Level familiarity with literature about racial and social justice

The survey data was collected using Google Forms, which provides a wide variety of analysis and export options. See [Appendix B](#) for the survey questions and <https://forms.gle/JfN3n8nVhRAx97cy8> for the survey form that participants used.

Subject Indexing and Subject Tagging Activities

I collected much of the data for this study while participants did subject indexing and subject tagging work. I asked participants to verbalize their thoughts, reactions, and actions as they worked, and recorded and transcribed their verbalizations. I also collected participants' aboutness statements and the subject headings and tags that they used to represent the books that they indexed and tagged. Finally, I took notes during this portion of the study to capture key events and data that was not verbalized or otherwise captured, such as tracking which bibliographic apparatus participants looked at while working.

- **Aboutness Statements:** For each book that a participant indexed or tagged, I first asked them to write an aboutness statement to summarize what they thought the book was about. I told participants that aboutness statements could be paragraphs, sentences, or lists of terms. I collected and marked aboutness statements (i.e., Participant #7, Task #3) at the end of each study session. [Figure 4.12](#) shows an example of an aboutness statement that also includes participant notes.
- **UI and Database Data:** Participants entered their subject headings and tags into the custom-made web application. The web application wrote to the database twice

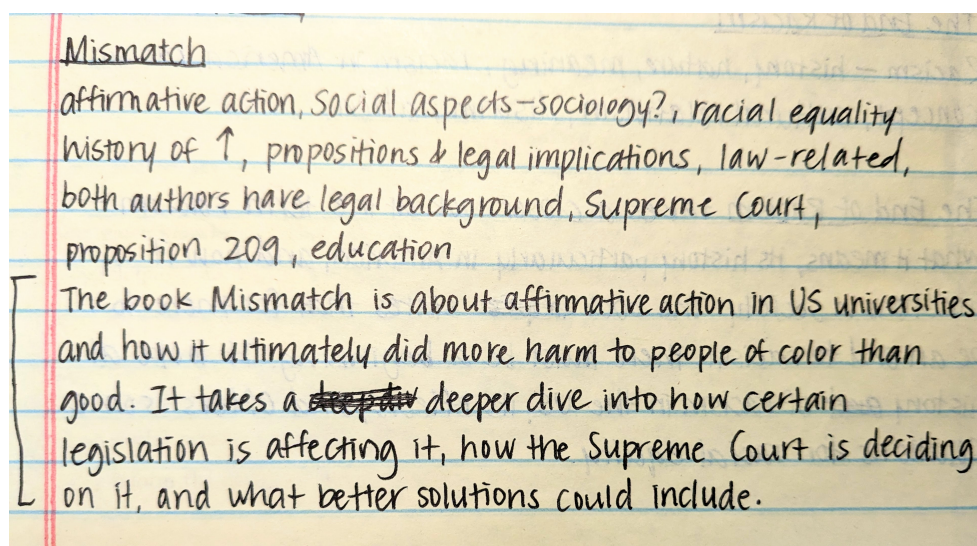


Figure 4.12: An aboutness statement from Participant 24 that describes the book *Mismatch*. The aboutness statement itself is marked with a line in the left margin. Participant 24 was one of a handful of participants who wrote notes—above the aboutness statement in this case—in addition to the aboutness statement.

for each task. The first write captured the starting timestamp for the tagging or indexing task. The second write captured the completion timestamp for the task and other data, including the tags or subject headings that the participant entered for that task. The data was stored in a MySQL database with one table that had the following columns:

- **userid**: A value, unique to each participant, that I provided participants when they filled out the demographic survey and that I entered in the administrator UI at the beginning of each session.
- **book**: The title of the book that the participant tagged during the activity.
- **condition**: A string that indicated the form of vocabulary control that the participant used during the activity. The possible values were **free**, **passive**, **flex**, and **rigid**.
- **sequence_num**: An incremented number that indicated the ordering of the task in the set of four tasks that each participant completes. For example, **sequence_num**

was 2 for the second book that a participant indexed or tagged.

- **tags:** The tags or the subject headings that the participant entered for each task. (Tags and subject headings were both stored in the same column even though that column’s name was `tags`.) The tags or subject headings were stored as a comma-separated list prefaced with “Tags” or “Subject Headings.”
- **timestamp:** A machine generated timestamp that captured the time when a row was written to the database, with the precision of one second. Each activity had two rows and two timestamps—one for when the activity began and one when the participant submitted their subject headings or tags. Unfortunately, this timestamp data would prove unusable as participants were inconsistent about clicking of the “Get Started,” “Submit,” and “Next Book” buttons.
- **Audio Recording:** While participants did subject indexing and subject tagging work, I asked them to verbalize their thoughts, actions, and reactions. This information and any questions or discussion with the researcher were be recorded on a small digital audio recorder that was placed on the table near the participant. This audio recording data was transcribed, coded, and analyzed.
- **Researcher Notes:** I took notes during the think-aloud protocol to track the bibliographic apparatus that participants used while working, to note key moments to analyze in the transcript, to note expressions and non-verbal cues that would not be captured in the audio recording, and to identify events and emotions that warranted follow-up questions during the interview.

Retrospective Interview

After the think-aloud protocol, I conducted a retrospective interview with each participant to reflect on their experience with indexing and tagging. I used an interview instead of a survey or other retrospective data collection method because the interview format supported more in-depth responses, allowed for back-and-forth discussions and clarification, and allowed me to ask follow-up questions in addition to the core interview questions listed in [Appendix J](#).

I collected the following data during the interview:

- **Audio Recording:** As with the think-aloud protocols, I recorded each interview with a small digital audio recorder and transcribed, coded, and analyzed the audio data.
- **Ranking Data:** During the interview, each participant was asked to rank the forms of vocabulary control from 1) easiest-to-use to most-difficult-to-use, from 2) making them feel most in control to most controlled, and 3) making the feel most in control to most out of control, by placing indexing cards in rank order. I captured and analyzed these rankings and the participants' explanations for their ranking choices. [Figure 4.13](#) shows an example of the ranking note cards arranged in order.

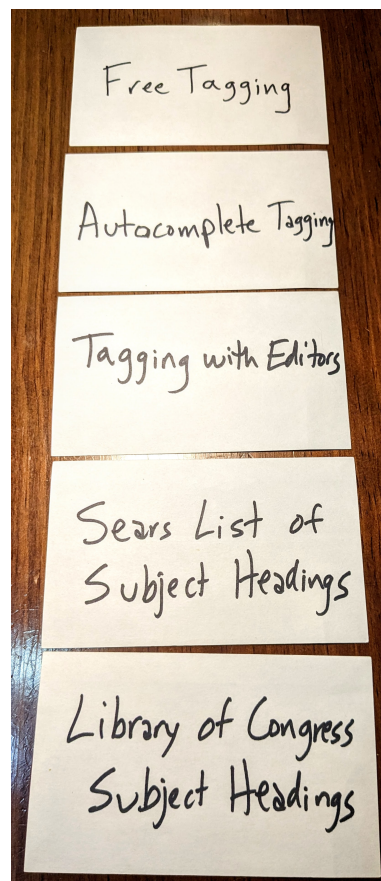


Figure 4.13: The ranking note cards, arranged to communicate a participants' ranking of the five forms of vocabulary control in the study.

- **Researcher Notes:** As with the think-aloud protocol, I took notes during the interview to note key moments to analyze in the transcript, to note expressions and non-verbal cues that were captured in the audio recording, and to track any questions that I asked outside of the core interview questions.

4.4 Threats to Validity and Limitations of the Study

This section discusses threats to internal validity, threats to external and temporal validity, and the resultant limitations to the claims that this study was able to make about the relationship between vocabulary control and the subject indexing and subject tagging processes.

4.4.1 Threats to Internal Validity

Good research design attempts to maximize internal validity, the extent to which a study actually observes the cause-and-effect relationships that it intends to observe, instead of accidentally observing the effects of other factors. This study faced threats to internal validity because of competing priorities in the design of the study. This section details some of these threats to internal validity and discusses how I minimized or acknowledged these threats.

A major goal of this study was to understand when participants felt in control, out of control, and controlled, and to understand how those feelings related specifically to the five different forms of vocabulary control. To do that effectively the research design needed to isolate aspects that were adjacent to but not fundamental to the form of vocabulary control. For example, participants who were unfamiliar with the Library of Congress Authorities website (<http://authorities.loc.gov>) were somewhat likely to feel controlled by the mechanics of searching and navigating that site, but that user interface experience was not an inherent property of rigid vocabulary control. Separating the effects of the fundamental differences between forms of vocabulary control from possible confounding effects of adjacent aspects was a challenge. I attempted to minimize the effect of adjacent aspects by making the user interface and instructions for each of the different forms of vocabulary control as similar as possible and by providing aid with adjacent features while participants worked.

I also made notes and asked clarifying questions to better understand whether participants were affected by aspects to the form of vocabulary control or by fundamental features of that form of vocabulary control.

Another common threat to internal validity is experimenter bias. Experimenter bias does not need to be intentional or consciously applied to be a threat. Because I was the lone researcher conducting sessions and analyzing the data and because that analysis required significant interpretation, my bias had a strong possibility of threatening the internal validity of the study. This threat was balanced, however, by consistency. For example, all participants interacted with the same researcher, so different researchers conducting study sessions was not a potential confounding variable. Still, my biases, especially as a straight, cis, white male conducting research about racial and social justice in subject indexing and subject tagging was a significant threat to internal validity. I aimed to mitigate this threat by reviewing my research methods with peers and advisors, by remaining aware of my potential biases, and by using established data analysis techniques. For example, I used a taxonomy developed by Bertrand and Cellier [12] to categorize types of transformations during the translation from subject analysis to subject representation. Still, one limitation of this study is that it I designed the method, ran the study sessions, and analyzed the data, and any biases that I hold were difficult to separate from the study.

This study used a think-aloud protocol, which is widely considered to be an effective method for understanding participant thought processes and feelings while they complete tasks [21]. Participants, however, can hide or distort their thought processes and emotions, either intentionally or unintentionally for a variety of reasons, including discomfort and shame. To minimize this threat to internal validity, I emphasized to participants that I was not evaluating them but that I was evaluating the forms of vocabulary control. I also created a safe space where participants could feel as comfortable as possible sharing their thoughts by being friendly and thankful during study sessions. I also aimed to collect redundant data, so I did not rely on solely the think-aloud protocol. For example, the retrospective interview provided an additional way to collect data about each participant's experience. Still, I acknowledge that think-aloud protocols are imperfect and might have limited the internal validity of this study.

Finally, ordering effects can distort cause-and-effect relationships for empirical research that, like this study, involves multiple iterations of similar procedures. For example, if one condition (e.g., rigid control) always came last in the study design, participants might have learned about the subject indexing and subject tagging processes in general by that time and applied their knowledge to only that form of vocabulary control. Conversely, they might have grown fatigued by the time they reach that last form of vocabulary control and struggled with or become more frustrated with that form of vocabulary control. To minimize the threat of ordering effects, I rotated the order in which participants experienced the different forms of vocabulary control and the order in which they worked with the four books in the study. This design also minimized the threat of confusing the effect of a specific book with the effect of a form of vocabulary control by rotating which book is tagged or indexed with each form of vocabulary control.

4.4.2 Threats to External and Temporal Validity

While internal validity is about whether a study's design observes the cause-and-effect relationships that it claims to observe, external validity is about how readily those observations can be generalized into other contexts, situations, and people. Temporal validity is a type of external validity focused on how well a study's findings can be generalized to other times. This section discusses some of the threats to external and temporal validity in this study.

A big threat to external validity for this study and all empirical research conducted in a lab setting is that the findings will not generalize to "real-world" settings. Šauperl minimizes this threat to external validity by observing catalogers in their regular work setting [116], but this study was conducted in an artificial setting—a conference room. The conference room setting helped to control for potentially confounding variables that might have been encountered in various work settings, but it limited my ability to claim that the findings could be broadly applied to "real world" work settings, where much subject indexing and tagging work is done.

Another potential threat to external validity stems from the sampling methods that I chose. In part because the experimental design was best suited to in-person sessions, I used

convenience sampling to recruit local students from the UW MLIS program. These students were clearly not a random sampling of the population, so I must limit any truth claims by noting that the sample population was not representative of the broader population. As described above, however, the study's sample population was similar to the ALA population in race and ethnicity and somewhat similar to the ALA population in gender. These similarities suggest that the findings might be generalizable to the population of librarians in the United States.

The study population had a significant difference from the ALA population, though. To support an important goal of the study—separating subject analysis and subject representation so that I could measure coextensiveness—I recruited novice and advanced beginner indexers and taggers who would be less likely to combine subject analysis and subject representation. According to findings by Šauperl and Bertrand and Cellier, more experienced indexers—who are much better represented in the ALA population—exhibit different indexing behaviors when compared to inexperienced indexers and taggers. They also hold more knowledge about indexing techniques like constructing synthetic subject headings. Therefore, I cannot apply my findings to the broader population and must limit truth claims to inexperienced indexers and taggers.

A related threat to external validity is the somewhat artificial line that my experimental design drew between subject analysis and subject representation. Langridge [86], Lancaster [85], and others note that these two steps in the subject indexing (and subject tagging) process are often co-mingled and iterative in practice, but this experiment isolated the steps to observe and compare what happened during each step. Furthermore, artificially enforcing both steps could be considered unrepresentative of real-world tagging because tagging frequently skips in-depth subject analysis (or skips subject analysis altogether) in favor of speed.

An additional threat to external validity relates to the domain of the books that participants indexed and tagged in the study. The study intentionally focused on indexing and tagging books about racial and social justice because these critical topics have historically been poorly represented in controlled vocabularies; however, this focus on a single subject domain limits claims that I can make about the differences between forms of vocabulary

control for other subject domains, especially for domains that are not as emotionally charged as racial and social justice. Put another way, the findings in this study might be tied to heightened feelings about control based on the specific subject matter of the books in the study. A similar study that used books about furniture might elicit different responses.

Finally, this study must consider threats to temporal validity. This study is novel in part because it considered forms of vocabulary control that had not been studied together before and that did not exist until social tagging evolved new forms of vocabulary control over the past decade. While studying these forms of vocabulary control filled a research gap, the emergence of new forms of vocabulary control tied to technological advances—and the very real likelihood that machine indexing will change significantly over the next decade—suggest a threat to the temporal validity of this study.

Chapter 5

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. [Table 5.1](#) shows the most assigned subject headings and tags for each of the books in the study.

Book Title	Most Assigned Subject Headings and Tags
<i>Caste</i>	Caste (7), Racism (6), Race relations (6), Social classes (4), Nazi Germany (4), Caste systems (3), Race relations—United States (3)
<i>The End of Racism</i>	Racism (16), Race relations (6), Race relations—United States (4), Race discrimination (4), Racism—United States (3), Pluralism (3)
<i>Mismatch</i>	Affirmative action programs (9), Higher education (9), Affirmative action (7), Education (5), Colleges and universities (4), Racism (3)
<i>White Fragility</i>	Racism (16), Race relations (8), Whiteness (6), Race discrimination (4), White fragility (3), Race awareness (3), White people (3)

Table 5.1: The most assigned indexing terms for each book in the study. The numbers in parenthesis indicate how many times that exact indexing term was used for that book by all participants in the study, across all forms of vocabulary control.

Participants also answered 15 interview question (with many sub-questions) each and ranked different forms of vocabulary 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 chapter reports, analyzes, and discusses the results of the study based on these data. The chapter is structured around answering the three research sub-questions described in [Chapter 3](#):

- **Research Sub-Question A:** How do different forms of vocabulary control affect indexers' and taggers' *experiences* as they index and tag documents? When do they feel controlled? In control? Out of control?

- **Research Sub-Question B:** How do different forms of vocabulary control affect the coextensiveness of aboutness statements and indexing terms?
- **Research Sub-Question C:** How do indexers' and taggers' experiences while working with different forms of vocabulary control inform criticisms of controlled vocabularies, particularly for documents about racial and social justice?

The results and discussion for research sub-question A focus on participants' experiences with vocabulary and their reflections on those experiences. The central data used to understand this research sub-question are the participants' rankings of each form of vocabulary control from making them feel most in control to most controlled, from making them feel most in control to most out of control, and from easiest-to-use to hardest-to-use. These quantitative data are contextualized with qualitative data, including participants' reflections on why they felt in control, out of control, or controlled with different forms for vocabulary control. Finally, this section considers the relationship between control and participants' responses to questions about how vocabulary control affected the specificity, exhaustivity, and accuracy of subject headings and tags.

The results and discussion section for research sub-question B begins by comparing participants' aboutness statements with the subject headings that they ultimately chose, using a taxonomy from Bertrand and Cellier [12] to categorize transformations in that translation process. Next, the section analyzes participants' self-reported reflections on coextensiveness to better understand the relationship between coextensiveness and control. Finally, the section considers the bibliographic apparatus that participants used, particularly the apparatus that they used during translation and subject representation.

The results and discussion section for research sub-question C qualitatively analyzes 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 1) participants' criticisms of problematic or missing terms in a controlled vocabulary and how they respond in those situations, 2) participants' reflections on how their identities (including racial identities) informed their indexing and tagging processes, 3) participants'

challenges in indexing and tagging books with viewpoints on racial and social justice that they disagree with, and 4) the effect of different forms of vocabulary on participants' ability to navigate complex racial and social issues in indexing and tagging.

5.1 Feeling Controlled, In Control, and Out of Control

Control, defined as the power to influence, limit, or direct the behavior or experiences of oneself or others, was the central concept of this study. Research sub-question A asked about the relationship between vocabulary control, participants' feelings of control, and participants' subject indexing and subject tagging processes. The goal was to understand the effect of vocabulary control on indexers' and taggers' experiences. Did they feel in control? Did they feel out of control? Did they feel controlled? How did these feeling affect their subject indexing and subject tagging processes?

During the retrospective interview, participants were asked to rank the four forms of vocabulary control that they used during the subject indexing and subject tagging portion of the study and one additional form of vocabulary control, post hoc control, which was introduced during the retrospective interview. Participants communicated their rankings by arranging note cards on the conference room table and by explaining why they chose the rankings that they did. This study analyzed participants' rankings and their discussion of those rankings for the following scales:

- The form of vocabulary control that made participants feel most in control to the form of vocabulary control that made participants feel most controlled
- The form of vocabulary control that made participants feel most in control to the form of vocabulary control that made participants feel most out of control
- The form of vocabulary control that was easiest to use to the form of vocabulary that was most difficult to use

The following sections analyze and discuss the data for each of these scales by looking first at quantitative data and then by adding context via qualitative and thematic analysis, and through key quotes from participants. Finally, this section analyzes participant responses

to questions about accuracy, specificity, and exhaustivity to understand the relationship between these properties and feelings of control and vocabulary control.

5.1.1 Feeling Most In Control to Feeling Most Controlled

This study proposed that there are two axes of control related to vocabulary control: 1) in control to controlled, and 2) in control to out of control. The first axis, which is discussed in this section, is more widely discussed in Knowledge Organization literature, particularly the role of controlled vocabularies in making indexers and taggers feel controlled. For example, Dartmouth librarians and others felt controlled while struggling to change the “Illegal aliens” headings, and critical cataloging literature notes many cases in which indexers and taggers feel controlled by controlled vocabularies [11, 105, 49, 75].

This study hypothesized, then, 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. The following subsections investigate whether that was the case, first from a quantitative and statistical approach and second from a qualitative and thematic approach that explores participants’ verbal explanations for their rankings.

Quantitative Analysis of In Control and Controlled Rankings

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 controlled, their rankings fit with the expectations set by critical cataloging literature. [Table 5.2](#) shows the 18 participants’ mean rankings on the scale from in control (1) to controlled (5) for each of the five forms of vocabulary control in the study. 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; feel increasingly in control with flexible vocabulary control, post hoc vocabulary control, and passive vocabulary

control; and feel most in control with minimal vocabulary control.

A post hoc Nemenyi Test made pair-wise comparisons between forms of vocabulary control (i.e., comparing just minimal control with flexible control, just minimal control with passive control, etc.) and found significant differences between the following pairs:

- Minimal vocabulary control and post hoc vocabulary control (p-value<0.05)
- Minimal vocabulary control and flexible vocabulary control (p-value<0.05)
- Minimal vocabulary control and rigid vocabulary control (p-value<0.05)
- Passive vocabulary control and rigid vocabulary control (p-value<0.05)

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

Table 5.2: The mean rankings for all five forms of vocabulary control on the scale from participants feeling in control (1) to participants feeling controlled (5), showing that participants felt most in control with minimal vocabulary control and most controlled with rigid vocabulary control. A Friedman test found a significant effect for vocabulary control on participants’ feelings of being in control or controlled (chi-square=36.3, p-value<0.05).

These significantly different pairs highlighted how much participants felt in control with the forms of vocabulary control most associated with tagging—particularly minimal vocabulary control—and how they felt controlled by forms of vocabulary control associated with subject indexing—particularly rigid vocabulary control.

Qualitative Analysis of In Control and Controlled Rankings

While statistical analysis of participants’ ranking data showed clear evidence that vocabulary control had a significant effect on participants’ feelings of being in control and controlled, this study also considered qualitative data to get a richer understanding of that significant effect.

After participants ranked each form of vocabulary control on the scale from most in control to most controlled, they were asked to explain what “in control” and “controlled”

meant to them. The dominant theme of participants' definitions "in control" was the code: "In control meaning freedom to do what the participant wanted to do," which was coded in 12 participants' responses. For example, Participant 4 said, "I can tag what I want. I am unbounded." Participant 11 framed being in control as not needing to conform; she could "use anything" she wanted. Participant 9 identified that they felt most in control with minimal vocabulary control and said, "I have the most say in what I am doing." The theme of freedom was evident not just in the retrospective interviews, but also during the think-aloud protocol. For example, while using minimal vocabulary control, Participant 5 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, which was ranked as the form of vocabulary control that made participants feel most in control. A related code that emerged during qualitative analysis was "In control meaning that the indexing term directly reflected the participant's thoughts," which was also associated with minimal vocabulary control. This code emphasized the coextensiveness of a participant's thoughts and the indexing terms that they used. For example, Participant 7 noted that with minimal vocabulary control indexing terms "reflect what I think."

The two previously discussed codes were based on participants' experiences and feelings, and did not connect the concept of control with the specificity, exhaustivity, and accuracy of indexing terms that participants assigned. A few participants did make that connection, though. For example, Participant 17 described being in control as, "I can do the task to the best of my abilities" and Participant 22 noted that being in control meant that they could "be highly specific" with their indexing tags. Another code that was more related to how "good" the indexing terms were was "In control meaning that the vocabulary control supported the participant and helped them find additional indexing terms," which participants cited when discussing passive and flexible vocabulary control. In these cases, participants felt that they could find good indexing terms through the support of the form

of vocabulary control. For example, Participant 23 said that passive vocabulary control made him feel most in control because he could choose any terms that he wanted and because he could find more words in the autocomplete vocabulary to make his tagging more complete. Similarly, Participant 16 ranked flexible control as making her feel most in control because she had a list to support her. Despite many participants emphasizing user-oriented indexing and tagging during other parts of the study, no participant directly connected feeling in control with being able to support the objectives of index and folksonomy users.

Participants were also asked to define what “controlled” meant to them, and their definitions of being controlled focus 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. The first theme, feeling constrained or restricted, was the most prevalent with nine participants describing being controlled as “being put in a box” (Participant 16), “needing to conform” (Participant 11), and “I can’t do what I want” (Participant 13). These statements all described participants’ reactions to using rigid vocabulary control and flexible vocabulary control and explain why participants consistently ranked these forms of vocabulary at the controlled end of the scale from in control to controlled.

The next theme that participants expressed in defining “controlled” was feeling that another agent had control. For example, Participant 17 alluded to working with “someone else’s ideas...someone else’s system.” Participant 22 felt that they were being “led by” someone else and Participant 9 said that someone else or something else had “the final say in what I can use.” Participant 25 noted that the autocomplete feature in passive vocabulary control “highly influenced” her and then anthropomorphized the feature while attempting to limit its control over her, saying that it was “not the boss of me.” Participant 9 also described a “relationship” with a form of vocabulary control, saying that the *LCSH* “owns me.” The relationship with “someone else” or another agent in control was particularly noteworthy for post hoc control, which ranked significantly higher than the two other forms of vocabulary control associated with tagging (minimal and passive) for participants feeling controlled. A recurring theme for participants who felt more controlled by post hoc control was the presence of another person making a final decision. Participant 7 called this presence

a “looming editor.” Participant 9 said that a post hoc editor had too much control and that there was too much room for misinterpretation and bias—that post hoc edits were “like putting words in my mouth.”

The third theme that emerged from qualitative analysis was that of a vocabulary not having the terms that they wanted to use. Participant 20 described this as having to “use what’s available” and to “conform.” Participant 13, in speaking about flexible control, noted that “there’s a lack of choice and there’s not enough terms available.” (Interestingly, Participant 13 ranked rigid control as making them feel most in control because the *LCSH* contained so many terms.) While many participants focused on how many terms were available in the vocabulary, others felt controlled by *which* terms were available in the vocabulary. For example, Participant 4 said, “Ugh! I was looking for ‘White Supremacy’ and it comes up with ‘White Supremacy Movements’ and that’s the only tag. That’s not really what this book is about.” Participant 5 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.” This quote reverberated with the case of “Illegal aliens” headings in the *LCSH*, which were considered so problematic because they were hurtful to undocumented immigrants. Participant 21 also connected with that case, lamenting that the *LCSH* lacked the term that they wanted to use—“Antiblackness”—and commenting on how slow the vocabulary was to change.

The fourth theme related to participants feeling controlled was coded as “Controlled meaning the participant felt controlled by rules and processes,” which was identified in 4 participants’ responses. For example, Participant 24 noted that they had to follow specific paths to “get to” indexing terms while they were using flexible control and rigid control. She also noted that the book format and the size of the *Sears List* made her feel controlled, which was a recurring theme that was concurrently coded as “feeling controlled by the form of vocabulary control.” One participant even responded with a half-sarcastic, half-serious “Nooooo!” and another said, “Oh boy, that’s big!” when they were handed the *Sears List*. Participants also felt controlled by the *LCSH* because, as Participant 17 remarked, they found it confusing. These themes of being controlled, intimidated, and confused by controlled vocabularies like the *Sears List* and the *LCSH* make sense for the inexperienced

indexers and taggers in the study, but would likely not be as prevalent with experienced indexers who were familiar with these controlled vocabularies.

Summary of Findings for In Control and Controlled Rankings

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 inexperienced 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 inexperienced indexers and taggers consider post hoc control more similar to flexible and rigid control than other forms of vocabulary control associated with tagging because post hoc control has an agent that can control them, 3) 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, with terms missing from a vocabulary, and with someone or something being able to exert control over the indexer or tagger.

5.1.2 Feeling Most In Control to Feeling Most Out Of Control

The second axis of control analyzed in this study compared when participants felt most in control with when participants felt most out of control. As becomes evident in the analysis below, "out of control" was interpreted by participants in two ways: 1) as similar to being controlled—lacking agency because of the constraints of the form of vocabulary control, and 2) as being recklessly unrestrained or operating without the necessary guardrails to feel self-control. This second interpretation related to some of the criticisms of social tagging, including lack of authority control, typos and orthographic variants, excessive exhaustivity or specificity, and uninformed or errant tag assignment.

This section first quantitatively analyzes participants' rankings of forms of vocabulary control from those that made them feel most in control to those that made them feel most out of control and identifies a bimodal distribution based on the two different interpretations discussed above. The data are split into these two separate groups and analyzed, but this

section refrains from drawing conclusions about participants feelings of being in control and out of control based on quantitative analysis. It does, though, use qualitative analysis to better understand participants' interpretations of "out of control" and how, when, and why different forms of vocabulary control made them feel in control or out of control.

Quantitative Analysis of In Control and Out of Control Rankings

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 between vocabulary control and feelings of being in control and out of control. [Table 5.3](#) shows the 18 participants' mean rankings on the scale from in control (1) to out of control (5) for each of the five forms of vocabulary control in the study. Flexible vocabulary control made participants feel most in control overall (2.61 mean ranking) and rigid vocabulary control made participants feel most controlled (3.44 mean ranking); however, the mean rankings were much closer together for this scale than they were for the in control to controlled scale. A Friedman Rank Sum Test found that the difference in participants' rankings was not significant (chi-square=4.4, p-value=0.36).

	Flexible	Passive	Minimal	Post Hoc	Rigid
Mean Ranking	2.61	2.72	2.83	3.38	3.44

Table 5.3: The mean rankings for all five forms of vocabulary control on the scale from participants feeling in control (1) to participants feeling out of control (5). The mean rankings were not significantly different from each other (p-value=0.36).

Based on this statistical analysis, the study did not reach an immediate conclusion about the effect of vocabulary control on feelings of being in control and out of control. However, a bimodal distribution of rankings warranted further investigation. 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 (ranking=1). In contrast, other participants described being out of control as lacking guard rails and not really knowing what they were doing or whether their indexing and tagging choices were good matches for the book that they were working with. These participants tended to rank minimal vocabulary control as the form of vocabulary control with which they felt most out of control (ranking=5). [Figure 5.1](#) shows this bimodal distribution for rankings of minimal vocabulary control. A Hartigan's Dip Test for multi-modality returned a $p\text{-value} < 0.05$, statistically confirming the bimodal distribution.

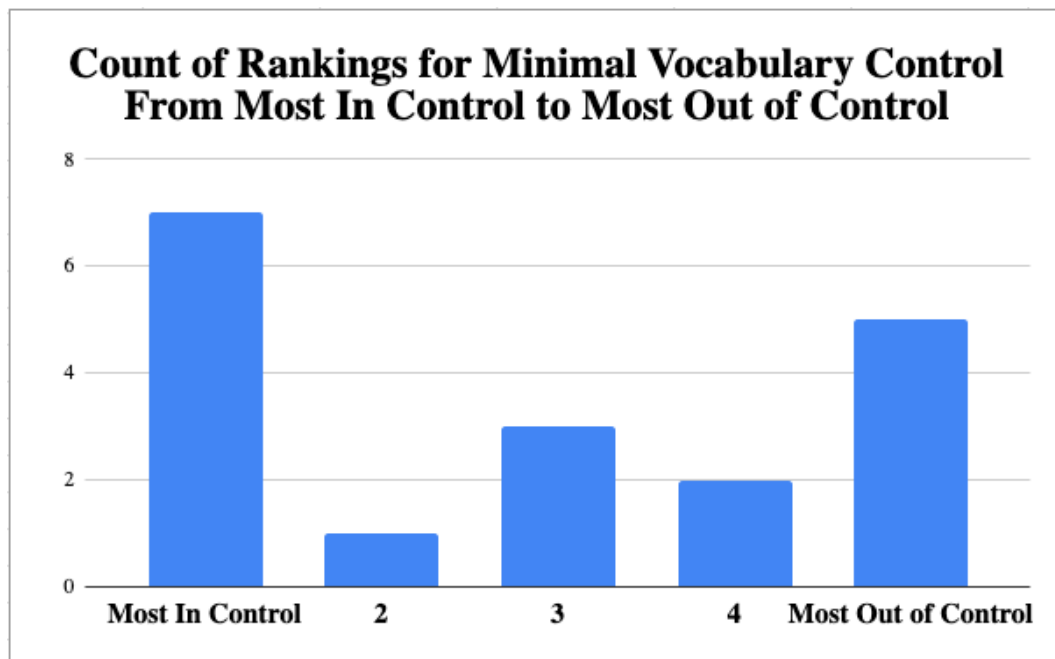


Figure 5.1: The counts of rankings, from most in control to most out of control, for minimal vocabulary control. A Hartigan's Dip Test for multi-modality returned a $p\text{-value} < 0.05$, confirming the bimodal distribution. The bimodal distribution suggested that participants interpreted "out of control" in two different and opposing ways.

When rankings for just the participants who described being out of control as lacking guard rails and not really knowing what they were doing were grouped together, that subset of results revealed a clear pattern, as shown in [Table 5.4](#). The Friedman Rank Sum Test for just the subject of data found that the difference in participants' rankings was significant

(chi-square=15.2, p-value<0.05).

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

Table 5.4: The mean rankings for all five forms of vocabulary control on the scale from participants feeling in control (1) to participants feeling 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. The mean rankings were significantly different from each other for this subset of data (p-value<0.05).

[Table 5.4](#) shows that participants with this interpretation felt significantly more out of control when using minimal vocabulary control. This was confirmed statistically by a post hoc Nemenyi Test, which showed a significant pair-wise difference between minimal vocabulary control and flexible vocabulary control for this subset of the data.

When rankings for just the participants who described being out of control as similar to being controlled were grouped together, that subset of results also revealed a clear pattern, as shown in [Table 5.5](#). The Friedman Rank Sum Test for just the subset of data found that the difference in participants' rankings was significant (chi-square=14.5, p-value<0.05).

	Minimal	Passive	Flexible	Post Hoc	Rigid
Mean Ranking	1.13	3.35	3.375	3.50	3.75

Table 5.5: The mean rankings for all five forms of vocabulary control on the scale from participants feeling in control (1) to participants feeling out of control (5), calculated based on the subset of participants who described being out of control as similar to being controlled. The mean rankings were significantly different from each other for this subset of data (p-value<0.05).

Again, participants' view of minimal vocabulary control stood out, in this case as the form of vocabulary control that made this subset of participants feel most in control. A post hoc Nemenyi Test showed a significant pair-wise difference between minimal vocabulary control and post hoc control (p-value<0.05), minimal vocabulary control and flexible

vocabulary control ($p\text{-value} < 0.05$), and minimal vocabulary control and rigid vocabulary control ($p\text{-value} < 0.05$) for this subset of the data.

Finally, it is important to state that while this split analysis of the two halves of a bimodal distribution was interesting for discussion—especially participants’ views of minimal vocabulary control—and will be explored further in qualitative analysis below, this analysis did not warrant drawing strong conclusions about the overall effect of vocabulary control on participants’ feelings of being in control or out of control because subsets of the data were intentionally selected and analyzed separately.

Qualitative Analysis of In Control and Out of Control Rankings

While participants’ bimodal interpretation of “out of control” precluded making any conclusions about the effect of vocabulary control on feelings of being in control and out of control, the qualitative analysis of participants’ explanations for their rankings proved to be a rich source of data. This section analyzes these responses by looking first at participants’ explanations for when they felt in control. These responses are somewhat similar to how participants described feeling in control in comparison to feeling controlled, but they include some useful differences and add context to a central research question in this study. Next, the section analyzes participants’ definitions of “out of control” and their explanations of when they felt out of control. This analysis is split into two codes that represent the two interpretations of out of control: 1) “Out of control meaning that the vocabulary and its rules are restrictive” and 2) “Out of control meaning that the form of vocabulary control lacks guard rails.”

Participants were asked to define “in control,” two times during the retrospective interview: once in comparison to “controlled,” and once in comparison to “out of control.” (These questions were asked in random order to avoid ordering effects.) Participants’ explanations of what “in control” meant to them were somewhat similar in both cases, so the same codes were used, but contrasting “in control” with “out of control” meant that participants emphasized different themes and raised new points.

One key difference was that more participants’ responses were coded as “In control

meaning that the vocabulary control supported the participant and helped them find additional indexing terms,” likely because many participants were contrasting “in control” with a definition of “out of control” that meant they did not feel supported. For example, Participant 24 described being in control as having a structure to follow, so she knew what to do. Participant 5 expanded on this sentiment: “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.’” Participant 16 and Participant 8 also ranked flexible control as making them feel in control because its partially enumerated list of subject headings guided them. Participant 11 noted that listed 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, Participant 23 liked how the autocomplete feature helped him to come up with terms, and Participant 21 liked that they could “work within an existing system, but still had the freedom to choose [her] own tags.” None of the participant responses identified with this code were associated with minimal control, which is discussed in more detail below in the opposite code “Out of control meaning that the form of vocabulary control lacks guard rails.”

The second most common code for this discussion of feeling in control was “In control meaning freedom to do what the participant wanted to do,” a contrast with the preceding code which focused on how the form of vocabulary control supported participants. Participant 13 described this other feeling of being in control as “free rein. You can do what you want.” Participant 25 said that they could use whatever indexing terms that she wanted, saying “I choose it.” These responses were provided by participants who treated the scale from in control to out of control as essentially the same as the scale from in control to

controlled, focusing in both cases on their autonomy and agency as indexers and taggers.

A third, closely related code, was also about participants being able to use the indexing terms that they wanted, in this case because the controlled vocabulary aligned with their ideas. Participant 22 called this alignment between ideas and vocabulary “harmonizing,” and Participant 20 described it as indexing terms being “close to the idea in my head.” These responses focused on passive, flexible, and rigid control, which enumerate tags and subject headings and were somewhat rare (count=3), suggesting that alignment between the vocabulary of inexperienced indexers and controlled vocabularies might also be somewhat rare.

At the opposite end of the scale, participants’ definitions of “out of control” were split into two codes. The first code was “Out of control meaning that the vocabulary and its rules are restrictive,” a theme that was common among participants who viewed being out of control as similar to being controlled.

The most frequently identified code related to participants feeling out of control (count=11) was “Out of control meaning that the vocabulary and its rules are restrictive,” signaling an interpretation of “out of control” that was similar to “controlled.” For example, Participant 25 felt out of control when “only specific words are allowed, and you have to apply them in a specific way.” Participant 20 was more specific in noting that they could not use the term that they wanted while working with rigid vocabulary control and felt “stuck” with “Blacks” and “Whites” from the *LCSH* instead of being able to use language that felt more contemporary and appropriate. As with the discussion of feeling controlled, many participants referred to another agent or someone else controlling them. For example, Participant 17 described the *Sears List* and the *LCSH* as “someone else’s. Even how I can explore them feels limited.” Participant 15 said that they felt like they “were not a partner” while working with rigid vocabulary control. These responses were all very similar to participants’ explanations of feeling controlled, reinforcing and adding more context to the finding that participants felt more controlled—or forced out of control—when working with more rigid forms of vocabulary control.

A smaller, but still substantial (count=8), number of participants’ descriptions of being out of control were coded as “Out of control meaning that the form of vocabulary control

lacks guard rails.” Participant 24 described this feeling as having “too much autonomy. I don’t know if what I’m doing is right.” Participant 20 compared using minimal vocabulary control to “flying solo,” and Participant 21 referred to minimal vocabulary as “shouting into the void,” both suggesting that free tagging did not provide sufficient support and guidance. Participant 23 also 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.” Participant 16 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. Participant 10 stated, “Using free tagging meant I had to rely on my own understanding, which was not that helpful.” Participant 8 called it “choice paralysis.” These quotes, contrasted with the quotes above about forms of vocabulary control that were considered too restrictive, illustrate the trade-off in forms of vocabulary control, especially for inexperienced indexers and taggers, between ensuring consistency and supporting diverse voices and emergent vocabulary.

Finally, a couple of uncoded comments about passive vocabulary control warrant a brief mention. First, Participant 7 felt that the autocomplete feature took away her feeling of control over the tag entry field because the autocomplete suggestions were not coming from her. This UI-specific comment was unique to this participant, but it showed that some indexers and taggers might view that text entry field as theirs and not want anything else influencing that space. Second, Participant 20 indicated that they might have ranked passive vocabulary control more towards feeling in control instead of out of control if they had known where the suggested terms came from. Because that vocabulary was a mystery, they felt more out of control. Again, this was a single comment, but an interesting point to consider.

Summary of Findings for In Control and Out of Control Rankings

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 found: 1) that inexperienced 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, 2) that flexible control and rigid control are more likely to make inexperienced indexers and taggers feel out of control if they equate being out of control with restrictions and that minimal control is more likely make inexperienced 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 inexperienced 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.

5.1.3 Easiest to Use to Most Difficult to Use

Ease of use is different from, but connected with, control in subject indexing and subject tagging. This study was interested in measuring ease of use for different forms of vocabulary control to understand which forms of vocabulary control were easiest and most difficult for inexperienced indexers and taggers to better understand the relationship between ease of use and control. This section does that by quantitatively analyzing participants' rankings of forms of vocabulary control from easiest to use to most difficult to use, by analyzing the time that participants spent on subject representation with each form of vocabulary control and the number of indexing terms they assigned in that time, and by thematically analyzing participants' explanations for their ease-of-use rankings and their commentary about when and why different forms of vocabulary were easy or difficult to use. Also, as discussed in the section about threats to validity, the study aimed to understand which aspects of different forms of vocabulary control were adjacent to but not fundamental to the form of vocabulary control. For example, participants who were unfamiliar with the Library of Congress Authorities website (<http://authorities.loc.gov>) were somewhat likely to feel controlled by the mechanics of searching and navigating that site, but that

user interface experience was not an inherent property of rigid vocabulary control. This section aims to identify the effects of vocabulary control that we based on adjacent aspects, particularly through qualitative analysis.

Quantitative Analysis of Ease-of-Use Rankings

When participants were asked to rank forms of vocabulary control from easiest to use to most difficult to use, their rankings showed that forms of vocabulary control with more rules and complexity were considered more difficult. [Table 5.6](#) shows the 18 participants' mean rankings on the scale from easiest to use (1) to most difficult to use (5) for each of the five forms of vocabulary control in the study. Minimal vocabulary control was considered the easiest to use (2.55 mean ranking) and rigid vocabulary control was considered the most difficult to use (3.83 mean ranking). A Friedman Rank Sum Test found that the difference in participants' rankings for ease of use was significant overall (chi-square=10.1, p-value<0.05). However, a post hoc Nemenyi Test did not find a significant effect in any pairwise comparisons. Still, this study found that vocabulary control had a significant effect on how easy or difficult novice and advanced beginner indexers and taggers found subject indexing and subject tagging.

	Minimal	Passive	Post Hoc	Flexible	Rigid
Average Rank	2.55	2.61	2.61	3.44	3.83

Table 5.6: The average ranking for all five forms of vocabulary control on the scale from easiest to use (1) to most difficult to use (5), showing that participants felt that minimal vocabulary control was easiest to use and that rigid vocabulary control was most difficult to use. A Friedman Rank Sum Test found a significant effect for vocabulary control on participants' feelings of being in control or controlled (chi-square=10.1, p-value<0.05).

Quantitative Analysis of Speed and Quantity for Indexing Terms

The study also measured how long each subject representation task took and how many indexing terms participants provided for each book that they indexed and tagged. These

measures—time on task and number of indexing terms—were treated as imperfect proxies for ease of use. For example, if a participant assigned subject headings or tags more rapidly, the form of vocabulary control was arguably easier to use because they could complete the task in less time. Similarly, if a participant assigned more subject headings or tags, the form of vocabulary control was arguably easier to use. Finally, the study calculated the number of subject heading or tags assigned per minute.

[Table 5.7](#) shows the mean time elapsed (in minutes) during subject representation, the mean count of indexing terms, and the mean indexing terms per minute for each form of vocabulary control—except for post hoc control. Post hoc control was not included in this table because, while participants ranked it during the retrospective interview, they did not perform subject indexing and subject tagging tasks with it. The table shows that minimal vocabulary control was the fastest form of vocabulary control, both in total time elapsed during subject representation and in indexing terms assigned per minute. Passive vocabulary control resulted in the highest number of indexing terms assigned per task (9.83 tags per book) and rigid vocabulary control resulted in the lowest number (6.33 subject heading per book). While participants noted how long indexing with the *LCSH* took—Participant 5 said, “Library of Congress is slow”—flexible vocabulary control took the most time (9:28 per task) and had the slowest rate of assigning indexing terms (0.68 subject headings per minute). One-way ANOVA tests found that vocabulary control had a significant effect on the time to complete subject representation (F-value=26.8, p-value<0.05), the number of indexing terms assigned per book (F-value=6.5, p-value<0.05), and the number of indexing terms assigned per minute (F-value=27.9, p-value<0.05).

While time spent on subject representation and the number of indexing terms assigned per book were not directly connected to this study’s questions about control, analysis of these data provided useful context for understanding the different forms of vocabulary control and how easy they were to use. Overall, these measures generally agreed with participants’ rankings for ease of use. Minimal control was ranked as easiest to use and it also took participants the least amount of time. Passive vocabulary control was ranked very close to minimal vocabulary control for ease of use and also had a very similar rate of indexing terms per minute. Rigid vocabulary control was ranked as the most difficult to use

	Minimal	Passive	Flexible	Rigid
Mean Elapsed Time (Minutes)	4:15	5:25	9:28	8:25
Mean Count of Indexing Terms	8.67	9.83	6.65	6.33
Mean Indexing Terms Per Minute	2.40	2.00	0.68	0.81

Table 5.7: The mean time elapsed (in minutes) during subject representation, the mean count of indexing terms assigned per book, and mean indexing terms per minute for each form of vocabulary control. ANOVA tests showed that vocabulary control had a significant effect on all the measures in the table.

and resulted in the lowest number of indexing terms assigned. However, flexible vocabulary control, which was ranked as somewhat easier to use than rigid vocabulary control took participants the longest amount of time and resulted in the lowest rate of indexing terms per minute (mean=0.68).

Qualitative Analysis of Ease-of-Use Rankings

Having established through quantitative analysis that vocabulary control had a significant effect on ease-of-use rankings, number of indexing terms assigned, time spent on subject representation, and indexing terms per minute, this section moves to a qualitative and thematic analysis of participants' discussion of their ease-of-use rankings. Participants expressed strong opinions about ease of use and made connections between ease of use and feelings of control. This section discusses those opinions first through thematic analysis of codes related to forms of vocabulary control being easy to use and then through thematic analysis of codes related to forms of vocabulary control being difficult to use.

Thematic analysis identified three codes associated with participants finding a form of vocabulary control easy to use: 1) "Easy to use meaning easy to understand, lower effort required," 2) "Easy to use meaning vocabulary matches or allows participant's terminology," and 3) "Easy to use meaning the form of vocabulary control provided useful feedback or guidance." The most identified of these codes focused on the low effort required to understand and use the form of vocabulary control. For example, Participant 4 said, "I don't have to think, it's intuitive. I know what I am doing" when they described why minimal

vocabulary control was easy to use. Participant 23 added to this description of ease of use, equating ease to the form of vocabulary control being easy to understand and making the work “quick.” Participant 9 said that “easy to use” meant that the form of vocabulary control required the least effort or thought. Participant 22 described this in terms of “the energy required” to do the indexing and tagging work. Another contributing factor that made a form of vocabulary control easy to understand was familiarity. For example, Participant 14 said that they were familiar with internet tagging, so they found minimal control and passive control easiest to use. Participant 15 ranked flexible vocabulary control as easy to use in part because they were familiar with the *Sears List*. Overall, these descriptions that connected ease of use to effort required matched with the quantitative findings for how long subject representation took to complete with different forms of vocabulary control. Minimal and passive control were easier to understand, required less effort, and took less time. Flexible and rigid control were more difficult to understand, required more effort, and requirement more time.

The second code that was associated with forms of vocabulary control being easy to use was “Easy to use meaning vocabulary matches or allows participant’s terminology,” which reflected participants’ perception of how much translation effort a form of vocabulary control required. For example, Participant 7 ranked minimal and passive vocabulary control as easy to use and indicated that they allowed for a direct link from her brain to “words,” or indexing terms. Participants 21 and 24 both described “easy to use” as being able to type in whatever they wanted or whatever they were thinking in the UI. This ease-of-use code, while primarily associated with minimal control and passive control, was also used to describe being able to find an accurate indexing term while using flexible or rigid vocabulary control. For example, Participant 25 ranked rigid vocabulary control as the second easiest to use because she was able to find “the right words” relatively easily.

The third code that emerged from participants’ explanations of what made a form of vocabulary control easy to use focused on the feedback or guidance that the form of vocabulary control provided. Unlike the previously discussed codes for ease of use, this code related to participants who found minimal control more difficult to use and preferred the support provided by more complex indexing languages like the *Sears List* and the *LCSH*.

For example, Participant 16 said that the *Sears List* and the *LCSH* because they answered the question, “Am I doing it right?” Participant 12 described this same feeling as “getting a response” from the form of vocabulary control and contrasted it with the lack of system response provided by minimal vocabulary control. Even some participants who found forms of vocabulary control associated with tagging to be easier to use, recognized the importance of system feedback. For example, Participant 20 remarked that they ranked passive vocabulary control as easier than minimal vocabulary control because the autocomplete tags were a form of support and confirmation.

This section moves next to analyzing reasons that participants found forms of vocabulary control difficult to use by discussing the following codes that emerged through thematic analysis: 1) “Difficult meaning complex or overwhelming,” 2) “Difficult meaning constrained or restricted,” and 3) “Difficult meaning the form of vocabulary control provided insufficient feedback or guidance.” While these codes were about forms of vocabulary control being difficult and not explicitly about control, this analysis showed that participants connected difficult to use forms of vocabulary control with feeling controlled and with feeling out of control.

The most identified “difficult” code was “Difficult meaning complex or overwhelming,” which was identified 11 times. The frequency of this code was not surprising for a participant pool of novice and advanced beginner indexers and taggers, particularly for those unfamiliar with more complex indexing languages like the *Sears List* and the *LCSH*. Participants noted three major ways in which forms of vocabulary control were complex or overwhelming: because the form of vocabulary control had too many rules, because the form of vocabulary had too many indexing terms, and because they could not easily find indexing terms via the form of vocabulary control. Specific to the difficulty of working with rules, Participant 14 cited the *Sears List* as complex and noted that it “required a learning process.” Participant 21 noted that both *Sears List* and the *LCSH* had “lots of rules that I didn’t know.” Participants who were overwhelmed by too many terms mostly cited the *LCSH* and the copious search results—many of them confusingly similar or overly specific—that they found when searching for a term. Participant 11 felt that she “didn’t know where to start” because there were “too many choices.” Participants also noted being

overwhelmed by the physical size of the *Sears List*.

The second most identified code that described why a form of vocabulary control was difficult was: “Difficult meaning constrained or restricted.” In some cases, this code was associated with the previous code—participants felt restricted because they did not fully understand the rules of the *Sears List* or the *LCSH*. In other cases, participants felt constrained because, as Participant 7 said about flexible control and rigid control, the indexing language “didn’t capture what I thought things were about.” She felt that the forms of vocabulary control associated with tagging were easier to use because they did not constrain her vocabulary choices. Participant 25 expressed a similar sentiment, noting that terms that they wanted to use were not available with rigid vocabulary control and not listed with flexible vocabulary control.

The third code associated with a form of vocabulary control being difficult to use was: “Difficult meaning the form of vocabulary control provided insufficient feedback or guidance.” This difficulty code—the direct opposite of the code “Easy meaning helpful feedback and confirmation”—closely connected with participants’ inexperience with indexing and tagging and suggested that indexing languages with all forms of vocabulary control could benefit from providing more targeted guidance for novice users. On one end of the control spectrum, some participants found minimal control (and to a lesser extent passive control) difficult to use because, as Participant 16 noted, she “did not get any feedback” on whether their tags were good matches for the document or well-formatted. At the other end of the control spectrum, participants found flexible and rigid control difficult to use because they did not provide enough guidance (via scope notes or other means) to help participants choose indexing terms, did not have a feedback mechanism to indicate whether an indexing term was allowed by the vocabulary, and did not provide direct enough guidance for inexperienced indexers to correctly construct subject headings using subdivisions.

Finally, while post hoc control was largely neglected in this discussion of ease of use because participants tended to equate its ease of use to minimal or passive control, two participants added some context about this form of vocabulary control. Participant 22 ranked post hoc control as easiest to use because they saw it as the same as passive control (“no extra step required of me”) with additional guard rails to ensure good tags. In contrast,

Participant 16 ranked post hoc control as most difficult to use because she knew that she would “overthink” her tags because she knew that someone else might change them.

Summary of Findings for Ease-of-Use Rankings

Based on these quantitative and qualitative analyses of participants’ rankings of forms of vocabulary on a scale from easiest to use to most difficult to use, this study found: 1) that vocabulary control has a significant effect on how easy or difficult inexperienced indexers and taggers find the translation and subject representation processes, how long those processes take, and how many indexing terms indexers and taggers assign to documents 2) that the forms of vocabulary control associated with tagging (minimal and passive) are perceived as easier to use, take less time to use, and result in more indexing terms being assigned than forms of vocabulary control associate with subject indexing (flexible and rigid), 3) that inexperienced indexers describe ease of use based on how quickly they can understand and use a form of vocabulary control, how much the form of vocabulary control includes or allows the terms that they want to use, how much feedback and support the form of vocabulary control provides, how complex the form of vocabulary control is, and how constrained or restricted the form of vocabulary control made them feel. These findings showed that while ease of use is not the same as control and ease of use is more dependent on the specific manifestation of a form of vocabulary control (e.g., a print copy of the *Sears List*), ease of use is still a useful metric for understanding how and when indexers and taggers feel in control, out of control, and controlled while doing subject indexing and subject tagging work.

5.1.4 Specificity, Exhaustivity, Accuracy

Chapter 2 established specificity, exhaustivity, and accuracy as important principles of subject indexing (and by extension subject tagging). Subject indexers (and subject taggers) aim to assign appropriately specific, appropriately exhaustive, and accurate indexing terms to represent documents, to make documents easier to find, and to collocate documents about similar topics. Their ability to assign specific, exhaustive, and accurate subject headings

and tags is related to control. Do indexers and taggers have control to assign appropriately specific tags? Do they feel in control of how exhaustive their subject headings and tags are? Do they have control to assign accurate indexing terms? Participants in the study were asked directly about the specificity, exhaustivity, and accuracy of their subject headings and tags during the retrospective interview. For example, one question asked: “For which book did your subject headings or tags most exhaustively—most completely—represent the contents of the book? Why?” This section analyzes their responses thematically, first by considering responses that focused on the properties of the books in the study and then by discussing the relationship between vocabulary control and specificity, exhaustivity, and accuracy. These themes are then discussed in connection with feelings of being in control, out of control, and controlled.

Specificity, Exhaustivity, Accuracy, and the Properties of Books

When participants were asked about specificity, exhaustivity, and accuracy, they were not directed to discuss only forms of vocabulary control, so many participants focused on the properties of the books that they indexed and on their familiarity with the books or the topics discussed in the books. For example, Participant 7 cited her familiarity with *White Fragility* and the book’s narrower scope as reasons that she was able to accurately represent the book. These book-centric responses inform the subject indexing and subject tagging processes in general, but they did not directly inform this study’s questions about the effect of different forms of vocabulary control on these processes. As such, this analysis aimed to focus on the effect of vocabulary control on specificity, exhaustivity, and accuracy; however, the properties of books and participants’ relationships with those books appeared to affect specificity, exhaustivity, and accuracy. The book-centric properties that had an effect are summarized below and the individual discussions of specificity, exhaustivity, and accuracy note when participants raised these properties as influences on their subject indexing and subject tagging work.

Participants cited the following book-centric factors when they discussed how specific, exhaustive, and accurate their indexing terms were: 1) their familiarity with the book, 2)

their familiarity with the domain of the book or the “conversation” around the book, 3) the degree to which they agreed with the viewpoint of the book, 4) how complex they perceived the book to be, 5) how long the book was, and 6) the clarity of the book’s summary or other bibliographic apparatus. Some of these book-centric factors were inherent to the books and some were about the participants’ relationships to the books. For example, Participant 17 noted the influence of an inherent quality of a book when she noted that her subject headings for *Mismatch* were accurate because she could pull indexing terms directly from the book’s bibliographic apparatus: “They gave me so many terms.” In contrast, Participant 24 felt that her tags for *Mismatch* lacked exhaustivity and accuracy because she did not know enough about education, signaling that her relationship to the book and its content was the key factor. Many participants noted their relationship to the content of *The End of Racism* as a reason that their indexing terms lack specificity, exhaustivity, or accuracy. For example, Participant 11 felt that their representation of *The End of Racism* was their least accurate because their personal biases meant that they could not reflect the book’s conservative political leaning—they “lacked words for that.” These book-centric factors are not explored in detail in this study, but they appear to relate to Bertrand and Cellier’s finding about bibliographic apparatus and indexers’ domain knowledge [12], and would be worthy of further exploration in future studies.

Specificity

Specificity is a central tenet of subject indexing going back to Charles Ammi Cutter [29] because appropriately specific indexing terms are critical to collocating documents about the same subject and essential to precision in information recall. This study did not evaluate whether participants’ indexing terms were appropriately specific for the books that they tagged and indexed. Instead, this study asked participants to discuss when they felt that the indexing terms that they assigned to a book were too specific, not specific enough, or just the right amount of specificity. Three themes emerged in participants’ responses, with one central theme that focused on vocabulary control.

The central theme that emerged was that participants consistently cited flexible control

and rigid control as helping them to assign more specific indexing terms because the *Sears List* and the *LCSH* enumerated many specific subject headings and provided broader and narrower terms, which guided them to the right level of specificity, or as Participant 14 said, “precise headings.” As inexperienced indexers, participants generally did not cite the *Sears List*’s and the *LCSH*’s support for synthesis as an aid to achieving specificity, although Participant 7 called their subject headings for *White Fragility* “spot on” and assigned both enumerated and synthetic subdivided headings while using the *LCSH*: “Anti-racism – United States,” “Racism – United States,” and “United States – Race relations.” Participant 15 was one exception to the theme of participants crediting partially enumerated vocabularies with helping them achieve specificity. They said that their tags for *Caste* were their most specific because they used minimal vocabulary control, and there was “nothing stopping me from using a specific tag.” However, Participant 9, who also tagged *Caste* with minimal vocabulary control felt that they had too much freedom to be specific and questioned their own tag, “Caste-freedom.” Overall, though, this study found that participants felt that forms of vocabulary control that listed authorized indexing terms with guidance about broader and narrower terms (and to some extent passive vocabulary control which suggests indexing terms), helped them to assign indexing terms at the right level of specificity. Additionally, this support for specificity was, at least in some case, connected with feeling in control. For example, Participant 11, who ranked flexible control as making them feel most in control cited the guidance provided by broader and narrower terms in the *Sears List*, saying that they “helped with specificity.”

A second theme that emerged was that, while participants were asked about their indexing terms being too specific, not specific enough, and just the right level of specificity, only a few participants cited indexing terms that they considered to be too specific. Instead, participants viewed more specificity as a positive, as a goal to reach toward in indexing and tagging. This viewpoint suggests that, as inexperienced indexers and taggers, the participants in the study found that achieving sufficient specificity was a challenge. It also helps to explain why participants preferred forms of vocabulary control that helped them find and assign more specific indexing terms. More experienced indexers and taggers might be less dependent on the support provided by some forms of vocabulary control when they aim to

assign appropriately specific indexing terms to documents.

Finally, another theme that emerged from participants' discussion of specificity was that their familiarity with and relationship to the content and positions of the book affected the specificity of their indexing terms. For example, Participant 7 felt that their indexing terms for *Mismatch* were not specific enough because they did not understand affirmative action enough. Participant 9 noted that her indexing terms for *The End of Racism* were less specific because she did not agree with the positions of the book and did not feel as motivated to represent it with specific terms. She ended up assigning broad subject headings that were featured in the book, including, "Racism," "Race," and "Society."

Exhaustivity

Exhaustivity in subject indexing and subject tagging is about assigning indexing terms that fully cover the topics in a document. This study hypothesized that indexers and taggers who felt more in control would be more likely to exhaustively index and tag documents, so this study aimed to understand what effected participants' ability to achieve exhaustivity by looking at the number of indexing tags that participants assigned for each form of vocabulary control and by thematically analyzing participants' responses to interview questions about the exhaustivity of their indexing terms. As with specificity and exhaustivity, this study did not attempt to evaluate the exhaustivity of participants' assigned indexing terms.

This study aimed to understand exhaustivity by measuring the number of indexing terms that participants assigned as a rough quantitative measure of exhaustivity. Lancaster who describes exhaustivity as assigning "enough terms to cover the subject matter of the document rather completely" also argues that exhaustivity "roughly corresponds to the number of terms assigned on average" [85]. Based on Lancaster's argument, this study analyzed the total number of tags assigned while participants used each of the forms of vocabulary control, as shown in [Table 5.8](#). Participants assigned the most indexing terms with passive vocabulary control (count=176), nearly as many with minimal vocabulary control (count=156), and fewer indexing terms with the two forms of vocabulary control associated with subject indexing: flexible control (count=122) and rigid control (count=114). A one-

way ANOVA test found that vocabulary control had a significant effect on the number of indexing terms assigned (F-value=5.8, p-value<0.05).

	Minimal	Passive	Flexible	Rigid
Total Indexing Terms	156	176	122	114

Table 5.8: The total number of indexing terms assigned by participants while using each of the four different forms of vocabulary control (Minimal, Passive, Flexible, and Rigid). Lancaster argues that exhaustivity “roughly corresponds to the number of terms assigned on average” [85], so the data in this table was treated as a rough quantitative measure of exhaustivity. A one-way ANOVA test found that vocabulary control had a significant effect on the number of indexing terms assigned (F-value=5.8, p-value<0.05).

Post hoc t-tests showed significant pair-wise differences between the following pairs, showing a clear split between forms of vocabulary associated with tagging (minimal and passive) and forms of vocabulary control associated with subject indexing (flexible and rigid):

- Minimal vocabulary control and rigid vocabulary control (p-value<0.05)
- Minimal vocabulary control and flexible vocabulary control (p-value<0.05)
- Passive vocabulary control and rigid vocabulary control (p-value<0.05)
- Passive vocabulary control and flexible vocabulary control (p-value<0.05)

While number of indexing terms is only a rough quantitative proxy for exhaustivity (especially for inexperienced indexers who showed they were more likely to use redundant indexing terms and to add extra indexing terms when they were uncertain), this quantitative finding was statistically significant and suggested that subject tagging has at least the potential to encourage greater exhaustivity in indexing terms because of the ease of assigning indexing terms with minimal and passive vocabulary control.

Although quantitative analysis found the vocabulary control had a significant effect on the number of indexing terms assigned and while Lancaster and others argue that number of terms assigned “roughly corresponds” to exhaustivity [85], this study investigated exhaustivity further by analyzing participants responses to interview questions about exhaustivity.

These interview questions aimed to understand when participants felt their indexing terms were exhaustive and not exhaustive, with the goal of better understanding the relationship between vocabulary control, exhaustivity, and feelings of control. First, participants' interview responses were analyzed to identify which forms of vocabulary control participants used for the books which they felt they indexed and tagged most and least exhaustively. These counts, which do not show the significant split between subject tagging and subject indexing that was seen in the preceding analysis of total indexing terms, are show in [Table 5.9](#).

	Minimal	Passive	Flexible	Rigid
Most Exhaustive	4	3	5	5
Least Exhaustive	5	2	7	3

Table 5.9: The number of times that a form of vocabulary control was used to tag or index a book that a participant cited as having their most exhaustive and least exhaustive set of indexing terms.

These data were inconclusive, in part because they measured the effect of at least two factors that were not isolated: form of vocabulary control (which this research aimed to study) and book-centric properties. Participants explained the exhaustivity of their indexing terms through both these lenses. For example, seven participants identified *Mismatch* as the book which they indexed or tagged most exhaustively, and many of them noted the detailed summary as a key aid in reaching exhaustivity. In contrast, other participants noted how vocabulary control affected their level of exhaustivity. For example, Participant 14 and Participant 15 noted that the autocomplete feature in the UI for passive vocabulary control showed them possibilities and helped them to expand the coverage of indexing terms. Participant 22 said that the extensive list of subject headings in the *LCSH* helped them “confirm their hunches” and arrive at an exhaustive set of indexing terms. Interestingly, minimal vocabulary control, which had a high rate of assigned indexing terms, was cited only once as the reason that a participant had exhaustive tags (participants cited book-centric factors the other three times when minimal control was associated with the most exhaustive

tags) and was frequently cited as a reason for less exhaustivity. For example, participant 14 said that they had nothing to refer to when assigning tags with minimal vocabulary control. Participant 15 contrasted minimal control with the other forms of vocabulary control, noting that those forms provided more direction and “filled in gaps,” especially the easy to access broader and narrower terms in the *Sears List*. While some participants noted that they assigned a less exhaustive set of indexing terms because of the restrictions of rigid vocabulary control, overall, participants’ responses seemed to indicate that they felt they could assign more exhaustive headings with forms of vocabulary control that provided suggested or authorized indexing terms and that provided some structure to help indexers and taggers find additional indexing terms. This finding from thematic analysis seemed to contrast with the quantitative analysis of the number of tags assigned, suggesting that for inexperienced indexers and taggers, the number of tags assigned might not be a good proxy for exhaustivity. The inexperienced indexers and taggers who participated in the study were more likely than experienced indexers to assign redundant and closely related indexing terms (i.e., “Race,” “Race relations,” “Racism,” “Race discrimination,” and “Antiracism”) instead of choosing one indexing term to cover a topic. This strategy was particularly common with subject tagging and in cases when participants were not confident about their knowledge of the book that they were indexing or tagging. For example, Participant 24 described their tags for *The End of Racism* as their most exhaustive set of indexing terms but also admitted that they were not sure what the book was about and said that they adopted a strategy of using all the reasonable suggestions in the autocomplete UI to ensure coverage.

Accuracy

This study did not try to evaluate the accuracy of participants’ indexing terms. Instead, questions in the retrospective interview asked participants to self-report when they felt that their indexing terms were more accurate or less accurate and to discuss what they felt affected the accuracy of their indexing terms. As a first point of analysis, the form of vocabulary control that was used to tag and the index books for which participants felt

they assigned the most and least accurate indexing terms were identified and summarized in [Table 5.10](#). These data include many responses that mentioned book-centric properties in addition to or instead of forms of vocabulary control, and these data also showed no clear quantitative relationship between accuracy and form of vocabulary control. Therefore, this study did not identify an effect of vocabulary control on accuracy based on quantitative data.

	Minimal	Passive	Flexible	Rigid
Most Accurate	3	6	5	5
Least Accurate	3	4	7	4

Table 5.10: The number of times that a form of vocabulary control was used to tag or index a book that a participant cited as having their most accurate and least accurate set of indexing terms. These counts showed no clear quantitative relationship between form of vocabulary control and participants' self-reported feelings about the accuracy of their indexing terms.

While this study did not find a quantitative relationship between form of vocabulary control and self-reported accuracy, some themes did emerge from qualitative analysis. First, participants noted book-centric properties more frequently than they cited the vocabulary control. Second, participants who cited the form of vocabulary control tended to cite passive, flexible, and rigid control as contributing to more accurate indexing terms. For example, Participant 16 felt that the indexing terms that she assigned while using passive control were most accurate because the autocomplete feature helped her understand what was appropriate. Participant 22 felt that the indexing terms that she assigned while using rigid control were most accurate because with the *LCSH* “the path was laid out before me.” Only two participants cited indexing terms that they assigned while using minimal control as their most accurate, and both other these participants noted book-centric properties as the reason for the accuracy. For example, Participant 21 cited the summary on the front flap of *Mismatch* as the main factor that contributed to her accurate headings. Participant 14 added to the feeling that minimal vocabulary control did not promote accuracy when they responded to working with minimal control by saying, “It’s gonna be real chaotic, then!”

Third, participants who cited a form of vocabulary control as the reason that they considered a set of indexing terms to be least accurate, consistently noted either confusion with how to use flexible and rigid control or frustration over vocabulary restrictions in flexible and rigid control. Passive control received the most consistently positive comments because participants liked its balance of vocabulary support without strict vocabulary restrictions.

Finally, while this study focused on the effect of vocabulary control on accuracy—and found no significant effect—it is noteworthy to consider the effect of books on accuracy, particularly *The End of Racism*. Zero participants cited *The End of Racism* as the book for which they assigned the most accurate subject headings or tags. *The End of Racism* was also the most cited book for least accurate headings. These data reflect that *The End of Racism* had a combination of being long and complex, having a less clear summary than other books according to participants, and, importantly, having a political perspective that participants generally disagreed with. Participants' reactions to *The End of Racism* are discussed further in a later section that explores vocabulary control, critical cataloging, and racial and social justice.

Summary of Findings for Specificity, Exhaustivity, Accuracy

This study examined specificity, exhaustivity, and accuracy because of their connection—or perceived connection—to control in subject indexing and subject tagging. The study assumed that the appropriate level of specificity, the appropriate level of exhaustivity, and accuracy in subject headings were goals of subject indexers and subject taggers, and that if they were able to achieve these goals, they were likely in control of the subject indexer or subject tagging process. These analyses of specificity, exhaustivity, and accuracy were included in the study, then, to better understand the research question about how vocabulary control affected the subject indexing and subject tagging processes and indexers and taggers feelings of control during those processes. However, the analyses in this section were largely inconclusive. Based on quantitative and qualitative analyses of specificity, exhaustivity, and accuracy, this study found: 1) that inexperienced indexers and taggers feel that partially enumerated lists of indexing terms help them achieve specificity by providing

suggestions and, in some cases, providing guidance to, more specific indexing terms, 2) that forms of vocabulary control associated with subject tagging produce significantly more indexing terms per book than forms of vocabulary control associated with subject indexing, but that indexers and taggers perceptions of the exhaustivity of their indexing terms do not necessarily correlate with the number of indexing terms, 3) that vocabulary control does not necessarily have an effect on the accuracy of indexing terms, but inexperienced indexers and taggers might prefer the balance of freedom and guidance provided by passive vocabulary control when trying to assign accurate indexing terms, 4) that specificity, exhaustivity, and accuracy all depend on an indexer or tagger feeling in control of the subject indexing and tagging process; when inexperienced indexers and taggers feel restricted or unsupported by a form of vocabulary control, they are less likely to feel satisfied with the specificity, exhaustivity, and accuracy of their indexing terms, and 5) that book-centric properties and indexer and tagger domain expertise likely have a significant effect on specificity, exhaustivity, and accuracy, as suggested by Bertrand and Cellier, and the influence of these factors warrant further study to better understand the subject indexing and subject tagging processes.

5.2 Coextensiveness of Aboutness Statements and Indexing Terms

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 the choices that indexers and taggers made when they translated their aboutness statement into indexing terms. Did indexers and taggers omit concepts because they do not know how to express them with a controlled vocabulary? Did they substitute the term that they used in their aboutness statement with a term from a controlled vocabulary? Did they add concepts that they encountered while working through the process of subject representation?

If the concepts and their expression in an aboutness statement completely match the concepts and their expression in a set of indexing terms, the aboutness statement and the indexing terms are completely coextensive. For example, an aboutness statement and indexing term could both be inaccurate—such as “Firefighters” for *The Fire Next Time* by

James Baldwin—but still be completely coextensive if the word or concept “Firefighters” was in both the aboutness statement and the set of tags or subject headings. However, in practice, aboutness statements and indexing terms are infrequently perfectly coextensive. Indexers and taggers make choices during translation to add or omit concepts, to use different vocabulary for concepts, and to combine or split apart concepts. For example, indexers often combine concepts by using adjective-noun subject headings like “Literary prizes,” subdivided headings like “Cloning—Ethical aspects,” and compound headings like “Architecture and people with disabilities.”

For example, suppose that an indexer or tagger did not have the concept of “Race discrimination” in their aboutness statement, but they encountered this indexing term as a suggested tag or as part of the *Sears* or the *LCSH* vocabulary, and they included it in their indexing terms. In this case, the aboutness statement and indexing terms would lack coextensiveness but not because the controlled vocabulary limited the indexer or tagger’s control. It would be just the opposite—the lack of coextensiveness in this case would indicate that the controlled vocabulary empowered and guided the indexer or tagger.

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, 2) using participant quotes and thematic analysis to expand upon that quantitative analysis, 3) exploring participants’ responses to interview questions about coextensiveness to understand how participants perceived coextensiveness, and 4) analyzing the how and when participants consulted with bibliographic apparatus and other resources while translating their aboutness statements.

5.2.1 Coextensiveness and Types of Transformations

This study collected an aboutness statement for each book that each participant indexed or tagged (4 aboutness statements per participant x 18 participants = 72 aboutness statements) and a set of subject headings or tags for each book that each participant indexed

or tagged. The aboutness statements were analyzed and 473 concepts were identified. For example, the aboutness statement, “*Mismatch* is about the failure of affirmative action in schools and how alternative policies could better help students” contained the following concepts: “affirmative action, school, students, policy.” The study also collected the subject headings and tags that participants assigned to books, a total of 568 indexing terms. Next, the study compared the concepts (and how they were represented) in the aboutness statement with the subject headings and tags that the participants ultimately used to represent the books. In the example above, the participant used minimal vocabulary control and provided the following tags: “Affirmative action, academia, race, education, mismatch, college, reform.” In this case, the concept “affirmative action” underwent no transformation during the translation process, “policy” was subtracted because it appeared in the aboutness statement but not the indexing terms, and “academia” was added because it did not appear in the aboutness statement but was used as a tag.

The example in the preceding paragraph is a simple one that discusses only two types of transformations (and one example of no change) during the translation process. Bertrand and Cellier [12] provide a more robust taxonomy of transformations that occur during the translation from an aboutness statement to indexing terms, particularly for transformations that are not simple additions or subtractions. The current study used the Bertrand and Cellier taxonomy to understand the role of vocabulary control in coextensiveness, as described below:

Subtractive Transformations

- **Suppression:** Suppression is the most common transformation observed in Bertrand and Cellier’s study, accounting for nearly half (46.9%) of all transformations. Suppression was determined in the current study 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 might express a concept like “Blackness” but their indexing terms do not include an equivalent term.
- **Simplification:** Simplification was rarely observed by Bertrand and Cellier (1.7% of

concepts) and was not observed in the current study. I looked for simplification 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 separate indexing term “Libraries” in the subject representation step.

- **Generalization:** Generalization was also somewhat rare in Bertrand and Cellier’s study (2.8% of concepts), but, unlike simplification, was observed in the current study. I looked for generalization 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.
- **Dissociation:** Like simplification, dissociation was rarely observed in the Bertrand and Cellier study (1.6% of transformations). Also like simplification, dissociation begins with a complex (multi-part) concept. For example, the concept “Art of Black people” might be represented by two dissociated indexing terms: “Art” and “Blacks.” I 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 [105, 47, 75]. However, dissociation was not observed in this study.

Additive Transformations

- **Addition:** Addition is the second-most common transformation observed in Bertrand and Cellier’s study, accounting for 31% of all transformations. Addition was determined in the current study by identifying an indexing term that was used in the final representation of the book but that was 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. Ad-

dition 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” additional indexing terms as they worked with the autocomplete UI or the *Sears List* or the *LCSH* apparatus.

- **Precision:** Bertrand and Cellier observed precision in 10.5% of transformations in their study. In the current study, precision was determined by identifying an indexing term that is 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.”

Other Transformations

- **Substitution:** Bertrand and Cellier observed substitution in 5.5% of transformations in their study. They also noted that substitution—using an indexing term that is different from but synonymous with a concept expressed in the aboutness statement—“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 vocabulary differences between “undocumented immigrants” and “Illegal aliens” in the Dartmouth Library case could be classified as substitution if one term appeared in an aboutness statement and the other emerged as an indexing term, but these similar terms are significantly different in their connotation and in how indexers, taggers, and information seekers experience feelings of control when interacting with them. In the current study, 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.”

5.2.2 Analysis and Discussion of Transformations from Aboutness Statements to Indexing Terms

The following section analyzes and discusses the degree of coextensiveness between aboutness statements and indexing terms and the types of transformations that occurred while participants worked with each form of vocabulary control.

The most basic view of coextensiveness focuses on cases of complete coextensiveness, in which a concept and its expression in an aboutness statement remain unchanged in an indexing term. [Table 5.11](#) summarizes the number and percentage of cases of complete coextensiveness for each of the four forms of vocabulary control that were studied during the subject indexing and subject tagging phase of the study. A one-way ANOVA test determined that vocabulary control had a significant effect on the number of concepts that remained unchanged during translation ($p\text{-value} < 0.05$). Minimal control had the highest percentage of concepts that remained unchanged during the translation (20.5%) and that rigid vocabulary control had the lowest percentage of unchanged concepts (10.7%). Taken with the percentages for passive vocabulary control (14.4%) and flexible vocabulary control (12.4%), this data suggested that forms of vocabulary control associated with subject tagging afforded participants more control than forms of vocabulary control associated with subject indexing because tagging more readily allows the participants' initial representation of a concept to appear in the indexing terms.

	Minimal	Passive	Flexible	Rigid
Total Concepts	200	215	186	196
Count Unchanged	41	31	23	21
Percent Unchanged	20.5%	14.4%	12.4%	10.7%

Table 5.11: The total concepts in aboutness statements and indexing terms and the count and percentage of these concepts that remained unchanged from aboutness statement to indexing terms for four different forms of vocabulary control (Minimal, Passive, Flexible, and Rigid). Minimal vocabulary control yielded the highest percentage of completely coextensiveness concepts (20.5%) while rigid vocabulary control yielded the lowest percentage of completely coextensive concepts (10.7%).

These findings were corroborated by further quantitative and qualitative data. On the quantitative side, post hoc t-tests showed significant pair-wise differences between minimal vocabulary control and flexible vocabulary control (p-value<0.05), minimal vocabulary control and rigid vocabulary control (p-value<0.05), and passive vocabulary control and rigid vocabulary control (p-value<0.05). On the qualitative side, participants reinforced this interpretation by noting that they could use any tags that they wanted while using minimal control. For example, Participant 7 noted that she felt most in control while using minimal control because her tags directly reflected what she thought the book was about. Participant 9 noted that she had “the most say” in the representation of a concept when she used minimal control. Other participants connected the lower percentage of complete coextensiveness in rigid and flexible vocabulary control to feelings of control. For example, Participant 20 noted when they used rigid or flexible vocabulary control that they could not always use a term that was “close to the idea in [their] head” and felt controlled when they felt that the indexing language required the use of “Blacks” to represent the concept of “Black Americans.” In some cases, the concept in the aboutness statement and the assigned indexing term remained unchanged because the participant used flexible vocabulary control. For example, Participant 13 searched the *Sears List* for “White people,” “Whites,” and “Caucasian” without success. They asked, incredulously: “Is there no white people in here?!” and eventually chose to add “White people” as a local subject heading.

While the preceding analysis, which focused on complete coextensiveness, was instructive and represented a significant finding of this study, it did not address many of the nuances involved in translating an aboutness statement into indexing terms. Bertrand and Cellier’s taxonomy of transformations addresses the many ways in which an initial concept might or might not be completely coextensive with that concept’s eventual representation in a subject heading or tag. For example, one might argue that substitution preserves much more of the original intent of a concept in an aboutness statement—and might even improve or refine the original expression of the concept—than suppression, which omits the concept completely.

The Sankey diagrams in [Figure 5.2](#) illustrate these nuances by visualizing all the transformations that occurred for each of the four forms of vocabulary control that participants

used in the study. Overall, the Sankey diagrams show that suppression was highest with forms of vocabulary control associated with indexing (flexible and rigid control) and that addition was highest with forms of vocabulary control associated with tagging (minimal and passive control). Forms of vocabulary control that provided more guidance through an autocomplete feature, thesaural relationships, and related search results (passive, flexible, and rigid control) showed a higher rate of participants adding to their original concepts to make them more precise. Each of the types of transformation are discussed in depth in the following sections.

Suppression Transformations

The type of transformations shown at the top of each diagram, suppression, proved to be more frequent for flexible and rigid control (count=55 and count=78, respectively) than for passive and minimal control (count=42 and count=43, respectively), suggesting that participants were more likely to abandon a concept from their aboutness statement if they could not find that concept or could not determine how to express that concept in a controlled vocabulary. However, a one-way ANOVA test indicated that vocabulary control did not have a significant effect on the number of suppression transformations (F-value=2.6, p-value=0.06). This quantitative analysis was somewhat surprising and motivated a qualitative analysis to understand why and when participants used suppression transformations.

In many cases, participants did suppress terms because of vocabulary control. For example, Participant 21 included the concept “public policy” in their aboutness statement for *The End of Racism*. When she translated her aboutness statement using the *Sears List of Subject Headings*, she looked for “Public policy” before looking for any other concepts, suggesting that she considered it a central concept. However, she did not find a “Public policy” subject heading and ended up suppressing that concept. In some instances, participants anticipated that the controlled vocabulary would not have the authorized heading that they wanted. For example, Participant 13 searched the *Sears List* for “Anti-blackness” while stating that it was “probably not there” and quickly moved on after remarking, “I was right.” Sometimes suppression occurred because the participant wanted to use a phrase to

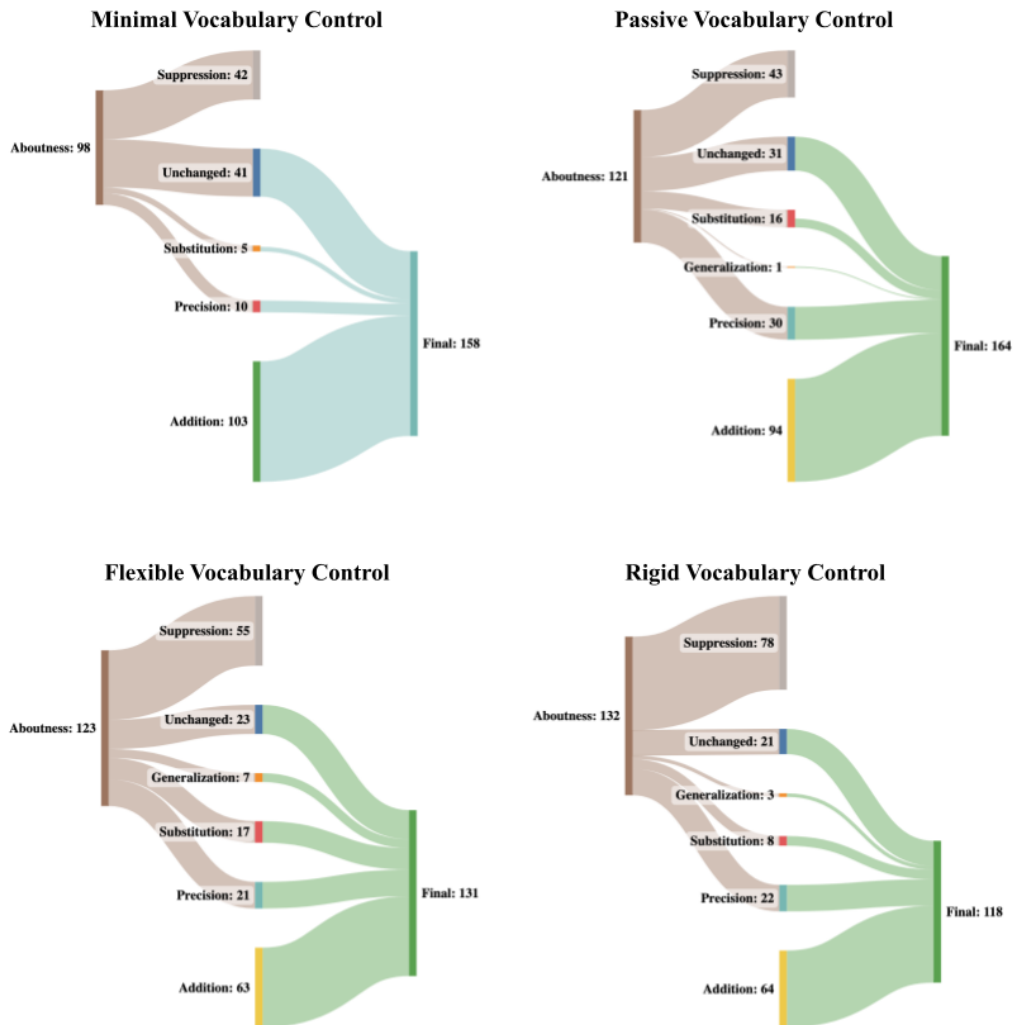


Figure 5.2: Sankey diagrams that show the types of transformations that all participants made they translated their aboutness statements into indexing terms while using four different forms of vocabulary control (minimal, passive, flexible, and rigid). The diagrams show that participants added many more new concepts during the subject representation phase when they used minimal and passive vocabulary control than when they used flexible and rigid vocabulary control. The diagrams also show that participants more often used precision transformations (adding detail through a subdivisions, adjectives, or other additions) while using passive, flexible, or rigid vocabulary control, all of which provide suggestions that include more precise indexing terms.

represent a concept and considered the available subject headings to be different in some important way. For example, Participant 22 wanted to use the phrase “White supremacy” to describe *White Fragility*. However, when they found only “White supremacy movements” in the *LCSH* authorized headings, they suppressed the concept because they wanted to represent an ideology and not movements. Participant 4 encountered a similar issue with passive vocabulary control: “Ugh! I was looking for ‘White Supremacy’ and it comes up with ‘White Supremacy Movements’ and that’s the only tag. That’s not really what this book is about.” In these examples, the participants wanted to express a concept in their subject headings or tags, but they felt controlled by the available vocabulary in the *Sears List*, the *LCSH*, and the autocomplete suggestions in the passive control UI.

However, not all suppression could be directly attributed to the form of vocabulary control. For example, Participant 16 suppressed most of the concepts from her aboutness statement for *The End of Racism* while working with rigid vocabulary control, including “America,” “History,” “Slavery,” “Civil rights,” and “Immigrant.” However, of these suppressed terms, she searched only for “Slavery” in the *LCSH* authorities, despite stating an intentional strategy to search for concepts in her aboutness statement. She did not, however, search for the other concepts that she suppressed. While the suppression of these other concepts might be attributable to fatigue from working with the controlled vocabulary, there was no direct evidence of that fatigue. Instead, many participants followed a pattern of suppressing terms from their aboutness statement for reasons other than the form of vocabulary control. In some cases, participants did not reference their aboutness statement after completing it, despite the research script encouraging them to translate their aboutness statement into subject headings or tags. This pattern accounted for much of the suppression that occurred with minimal vocabulary control, as participants who continued to reference their aboutness statement during the subject representation phase for minimal vocabulary control often directly typed concepts from the aboutness statement into the tagging UI. In other cases of suppression, participants simply changed their mind or questioned their initial analysis as they moved from analysis to representation, often consulting with the book that they were indexing or tagging to confirm their choices.

So far, this discussion of suppression transformations has focused on the lower rate of

suppression in tagging (minimal and passive vocabulary control) when compared to subject indexing (flexible and rigid vocabulary control). However, the difference between suppression with flexible vocabulary control (count=55, 29.6% of concepts) and rigid vocabulary control (count=78, 39.8% of concepts) was also noteworthy. Interestingly, the difference in suppression with these two controlled vocabularies might be explained by their main difference: The *Sears List* is not exhaustive, and thus recommends that indexers and subject headings “as needed.” Participants in this study added 20 such subject headings while working with the *Sears List of Subject Headings*, including “Dehumanization,” “Affirmative action,” “Law in education,” “Antiracism,” and “Social stratification.” While the alternative to adding a subject heading might not have always been suppression—participants might have continued to look for an alternative suitable subject heading—suppression is a likely path, as participants often abandoned concepts when they could not find a suitable *LCSH* subject heading. If the 20 subject headings that participants added while working with the *Sears List* had been suppressed, the suppression count for flexible control would have been 75 and the percent of suppressed concepts would have been 40.3%, which would align very closely with the total count and percentage of suppression for rigid control. It seems somewhat likely that the flexibility to add local subject headings with the *Sears List* reduces suppression in comparison to the rigid control of the *LCSH*. It is important to remember, though, that vocabulary control did not have a statistically significant effect on suppression transformations.

Based on this discussion of suppression, this study found: 1) that suppression might be caused by the form of vocabulary control and by factors outside of vocabulary control, including taggers’ and indexers’ use of their aboutness statement during the subject representation phase, 2) that the forms of vocabulary control associated with tagging—minimal and passive control—had the lowest rates of suppression but these rates were not significantly different from the rates for other forms of vocabulary control, and 3) that the flexibility to add local indexing terms in an otherwise controlled vocabulary like the *Sears List* seemed to reduce suppression in comparison to the more rigid approach of the *LCSH*—and affords indexers with more control to represent complex topics around racial and social issues.

Generalization Transformations

The second type of subtractive transformation is generalization, in which a specific concept is transformed into a less specific concept during translation. As shown in the Sankey diagrams, generalization was rare in the study, with a total of only 11 instances of generalization identified. Of those 11 instances, none occurred with minimal vocabulary control and only one occurred with passive vocabulary control. The other 10 instances occurred with forms of vocabulary control associated with subject indexing: seven instances for flexible vocabulary control and three instances for rigid vocabulary control. While the magnitude of these numbers was quite small, a one-way ANOVA test found that vocabulary control had a significant effect on generalization ($F\text{-value}=3.3$, $p\text{-value}<0.05$), and a post hoc t-test found a significant pair-wise difference between minimal vocabulary control and flexible vocabulary control ($p\text{-value}<0.05$).

Based on these quantitative findings, it is useful to characterize these generalization transformations and to try to understand how generalization relates to feelings of control. For example, one generalization transformation occurred when Participant 4 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, Participant 8 and Participant 23 transformed the specific concept of “Caste system” to the more general subject heading “Social classes” while working with the *Sears List*. As subtractive transformations, generalizations were not as dramatic as suppression transformations because part of the original concept survived the translation and because they occurred much less frequently in this study and in the Bertand and Cellier study. However, the nearly complete absence of generalization with minimal and passive vocabulary control suggests that this type of transformation is specific to the forms of vocabulary control associated with subject indexing. In contrast, taggers had little or no reason to generalize their original concept because they could use the original concept as a tag.

Based on this discussion of generalization, this study found: 1) that generalization is rare with all forms of vocabulary control for inexperienced indexers and taggers, 2) that when generalization does occur, it is more likely to occur with more rigid forms of vocabulary control that are associated with subject indexing—particularly flexible vocabulary control, and 3) that generalization can reflect an indexer or tagger feeling controlled by an indexing language.

Substitution Transformations

Bertrand and Cellier categorize substitution transformations as “neutral” because the concept in the initial aboutness statement remains present in the final subject headings or tags. For example, Participant 12 substituted “Universities” with the *Sears List* subject heading “Universities and colleges” without significantly changing the original concept. However, a substitution transformation is not always neutral, which, in some cases, can inform the relationship between coextensiveness and control. For example, Participant 11 and Participant 13 substituted “Affirmative action” with “Affirmative action programs,” which was nearly synonymous and preserved the original concept, but which held different meaning to them. In these cases, the addition of “programs” was not coded as a precision transformation because the participants indicated that they intended their original meaning but were substituting the authorized *Sears* heading. Relatedly, Participant 8 substituted her original concept of “White people” with the authorized *LCSH* heading “Whites,” and Participant 11 and Participant 24 substituted the authorized *Sears* 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.

The distribution of substitution transformations differed from the other transformations discussed so far. There was not a trend for more of less rigid forms of vocabulary control and there was not a split between forms of control associated with tagging and forms of control associated with indexing. As shown in [Figure 5.2](#), substitution was most common with

passive vocabulary control (count=16) and flexible vocabulary control (count=17). Substitution was less common with minimal vocabulary control (count=5) and rigid vocabulary control (count=8). A one-way ANOVA test found that vocabulary control had a significant effect on substitution transformations (F-value=3.5, p-value<0.05).

The similar counts for passive and flexible vocabulary control might seem anomalous at first, but passive control and flexible control share an important feature that might encourage substitution, especially for novice and advanced beginner indexers and taggers: Easy access to possible substitution through browsing activities. For passive control, many participants relied on the autocomplete feature to find related tags and tags with the same root. For example, Participant 24 typed “rac” into the user interface as an intentional strategy to see what was suggested, and then she quickly moved to using those suggestions. For flexible control, participants performed the print analog of this same behavior. For example, Participant 13 struggled at first to represent concepts related to racial and social issues, but she eventually found a set of relevant headings by scanning through the *Sears List* starting with “soci.” While this browsing strategy was used for both addition and substitution transformations, it is one likely reason that passive control and flexible control had the highest rates of substitution transformations. It also showed that novice and advanced beginner indexers and taggers would actively—even enthusiastically—seek guidance from an indexing language and did not express feeling of frustration or of being controlled in these situations. Another likely reason that passive control and flexible control had higher rates of substitution than rigid control is that the *LCSH* is a much more expansive indexing language, and thus more likely include subject headings that match the form of the topics and concepts in the aboutness statement.

Interestingly, substitution did not always involve the expected workflow of looking for term in an aboutness statement, not finding that term or finding that it was an entry term, and then using the authorized term instead. In many cases, participants arrived at a substitution transformation by other means. For example, Participant 8, who had “White people” in her aboutness statement and “Whites” in her subject headings, never tried to look up “White people” in the Library of Congress Authorities. Instead, she looked at the verso for *White Fragility* after completing her aboutness statement, found the subject heading

“Whites” and quickly chose to use that authorized heading instead of pursuing her original form of “White people.” Similarly, Participant 11 and Participant 24, who both had “Antiracism” in their aboutness statements and ultimately used the *Sears* authorized heading “Racism,” did not try to find “Antiracism” in the *Sears List*. Instead, they started searching the *Sears List* with “Race” and “Racism,” followed the browsing behavior described above, and felt that the concept of “Antiracism” from their aboutness statements was covered with the subject heading “Racism.” Interestingly, this “Antiracism” to “Racism” substitution transformation occurred in the opposite direction for two participants while they were working with the *LCSH*. Participant 14 and Participant 21 both had “Racism” in their aboutness statements and ultimately included the *LCSH* authorized heading “Antiracism” in their subject headings. Both Participant 14 and Participant 21 chose to search for “Antiracism” during the subject representation step even though they did not include the term in their aboutness statements.

When the transformation workflow *did* follow the expected pattern for substitution, participants did not express the type of frustration or lack of control that librarians expressed over the *LCSH* requirement to use authorized headings with the phrase “Illegal aliens.” For example, Participant 13 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.” Similarly, when Participant 5 entered “Universities” into the tagging UI with autocomplete enabled, she saw “College and universities” and quickly chose to select it. She was so satisfied with that substitution, that she also briefly considered an additional and more specific suggestion that came up: “Colleges and universities–Entrance exams.” Participant 14 had a similar experience with passive control. She had the concept “racial integration” in her aboutness statement and entered “Integration” in the tagging UI. When she found “School integration,” she quickly selected it and was satisfied with the substitution. Throughout the study, I did not find a case of substitution where the participant showed frustration or expressed that they felt controlled. Instead, they were generally satisfied with the authorized or suggested indexing term when they used substitution and were more likely to use suppression if they did not find an authorized heading that matched the concept in their aboutness statement or to use

their own phrasing if the form of vocabulary control permitted.

Based on this discussion of substitution, this study found: 1) that substitution occurs more frequently with passive control and flexible control than with other forms of vocabulary control, at least for novice and advanced beginner indexers and taggers, who are more likely to browse indexing languages for suggestions, 2) that substitution does not always involve the expected workflow of looking for a concept from the aboutness statement, not finding it, and replacing it with an authorized indexing term; instead substitution often occurs because the indexer or tagger finds additional information during the subject representation step or is satisfied with indexing terms that they find while looking up other concepts from their aboutness statement, and 3) that even when substitution does follow the expected workflow, novice and advanced beginner indexers and taggers do not necessarily feel controlled when they make substitution transformations; instead, they were satisfied with the substitution choices that they made or they used other strategies if they could not find a suitable substitution. This last finding did not match the feelings that librarians expressed around control and lack of control in the “Illegal aliens” case at Dartmouth, which is somewhat surprising. This difference can possibly be attributed to inexperienced indexers and taggers being more willing to accept authorized headings and suggested tags as useful guidance or could be attributed to participants’ use of strategies other than substitution when they were not satisfied with authorized headings and suggested tags. In the “Illegal aliens” case, librarians did not feel that those options were viable.

Precision Transformations

Precision is an additive transformation because it involves adding to an original concept from an aboutness statement via phrases, nouns, adjectives, and subdivisions in an indexing term. For example, Participant 14 performed a precision transformation when she translated “Racism” to “Racism against black people” while using rigid vocabulary control to index *White Fragility*. Similarly, Participant 11 performed a precision transformation when she translated “Affirmative action” to “Affirmative action—United States—public opinion” while using rigid vocabulary control to index *Mismatch*. While these two examples both

involved rigid vocabulary control, precision transformations occurred with all four forms of vocabulary control and were most prevalent with passive vocabulary control (count=30), flexible vocabulary control (count=21), and rigid vocabulary control (count=22). Minimal vocabulary control had only ten precision transformations. While these counts varied somewhat, a one-way ANOVA test found that vocabulary control did not have a significant effect on precision transformations (F-value=2.3, p-value=0.09).

Although the effect of vocabulary control on precision transformations was not statistically significant, this study used thematic analysis to better understand when and why participants made precision transformations. The lower rate of precision transformations with minimal vocabulary control, which not statistically significant, was noteworthy not just because of its lower frequency, but also because of the terms that were translated. All ten of the precision transformations with minimal vocabulary control fell into two groups: 1) adding the word “relations” to a concept (i.e., “Race relations” and “gender relations”) and combining “United States” with another topics (i.e., “United States History” and “Racism—United States”). Both these groups of precision transformations were likely influenced by participants’ experiences with pre-coordinate headings from earlier in the study or from before the study. For example, Participant 17 found the subject heading “Race relations” in the *Sears List* while indexing their first book, remembered it as a good subject heading or tag, and ended up applying it to all four books that she indexed or tagged.

Many other precision transformations were more closely connected with the form of vocabulary control. For example, Participant 23 looked up “Economics” in the *Sears List*, found the more specific terms “Economic conditions” and “Socioeconomics” listed as narrower terms. He expressed satisfaction with the more specific terms and felt in control of the choice to use them. Similarly, Participant 9 searched in the autocomplete UI for “Psychology,” found the more specific term “Social psychology” as a suggested tag, and chose to make a precision transformation and use the more specific term. Participant 15 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*. Participant 14 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. These positive feelings were somewhat unsurprising considering that these cases of precision transformations were “success stories” where the participant found a more specific term in a controlled vocabulary and chose to use that term instead of the less precise term in their aboutness statement. Still, this theme is important to note because it shows many instances of novice and advanced beginner indexers and taggers feeling supported—not controlled—by controlled vocabularies and improving their indexing terms by being more precise.

Based on this discussion of precision transformations, this study found:1) that vocabulary control does not have a significant effect on precision transformations; however, precision transformations are somewhat less frequent with minimal vocabulary control, which does not offer any suggestions for more specific terms and 2) that novice and advanced beginner indexers and taggers felt in control when they searched or navigated controlled vocabularies and they found more specific terms, suggesting that controlled vocabularies—especially those with enumerated terms—provide useful support to inexperienced indexers and taggers and help them achieve the goal of specificity.

Addition Transformations

Addition transformations, which are featured at the bottom of the Sankey diagrams in [Figure 5.2](#), involve the indexer or tagger assigning an indexing term that was not denoted as a concept in the aboutness statement. Addition transformations were the most frequently observed type of transformation in the study, with participants making a total of 234 addition transformations. In fact, more than half of the 200 total concepts (the union of concepts in the aboutness statements and indexing terms) for books that were tagged using minimal vocabulary control were the result of addition transformations (count=103). Passive vocabulary control also had a high rate of addition transformations (count=94), while the forms of vocabulary control most associated with subject indexing had lower rates of addition transformations (flexible control count=63, rigid control count=64). A one-way ANOVA

test found that vocabulary control had a significant effect on addition transformations (F -value=3.9, p -value<0.05).

While the high rate of addition transformations in the study was noteworthy in its own right, the split between forms of vocabulary associated with tagging (minimal count=103, passive count=94) and forms of vocabulary control associated with subject indexing (flexible count=63, rigid count=64) was even more interesting because of what it said about control and ease of use in the subject indexing and subject tagging processes. This difference could be attributed, in part, to the ease of use of the forms of vocabulary control associated with tagging. Participants ranked these forms of vocabulary control as significantly easier to use, often citing the ease of assigning tags without checking to see if they were authorized—they could just type them in. In the case of addition transformations, that meant “just typing in” many indexing terms that they thought of during subject representation. When participants quickly performed addition transformations, especially with minimal control, they expressed satisfaction with completing the task and having “enough” indexing terms. It is important to note, though, that some participants felt that addition was more difficult with minimal control because it offered no suggestions or guidance. For example, Participant 13, while working with minimal vocabulary control, said “I can’t think of it off the top of my head.” Similarly, Participant 11 stated: “I don’t know what words to describe that. I feel like they’re not coming to me. I think that’s part of why when I looked at [*LCSH* terms in front of the book I was tagging], I think they gave me language. I have ideas and I want to put them into words, but I don’t know how to do that.” Overall, though, participants used addition transformations significantly more with minimal control and passive control.

Another explanation for the significant split in addition transformations was the co-occurrence of the codes “feeling uncertain about an indexing term choice” and “Aiming for exhaustivity.” When these two codes were identified together for participants working with minimal vocabulary control and passive vocabulary control, participants added many tags to ensure that at least one tag (or the combination of tags) fully represented a concept or concepts. For example, Participant 22 expressed uncertainty about how to represent race and racism while tagging *The End of Racism*. They chose to assign all three of “Racism,” “Race relations,” and “Race relations—United States” to ensure that they

had covered the concept instead of choosing just one tag to represent the concept. Participant 13, while working with passive vocabulary control and tagging *Caste*, assigned an even larger group of tags, “riffing” on the concept of discrimination to ensure exhaustivity: “Race discrimination,” “Discrimination,” “Discrimination against people with disabilities,” “Discrimination in education”, “Discrimination in employment,” “Discrimination in housing,” “Discrimination in public accommodations,” “Discrimination,” “Race discrimination,” and “Sex discrimination.” As novice and advanced beginner indexers and taggers, participants also assigned redundant indexing terms while using flexible control and rigid control, but not nearly at the rate that they did with minimal control and passive control. Instead, with flexible control and rigid control, the code “feeling uncertain about an indexing term choice” was more often associated with suppression. Participant 17 partially explained this difference while describing why they did not assign more subject headings for *Caste* while working with flexible vocabulary control: “Because I am unsure about this book, I am going to play it as safe as possible. I’m uncomfortable to add any more.” Participants indicated two reasons for this difference: 1) they ranked minimal control and passive control as significantly easier to use than flexible control and rigid control and 2) they expressed a different perception of tags, with multiple participants referring to them as “just tags,” and other participants noting that they could use as many tags as they wanted but that the number of subject headings were limited by cataloging standards and best practices.

Another theme that emerged with addition transformations—that was also seen in substitution transformations—was a browsing behavior wherein participants searched for a term or the root of a term in the autocomplete UI or the *Sears List* and then browsed for suggested or nearby terms. For example, while working with passive vocabulary control, Participant 14 stated: “So, I typed in education and now I am going to browse my many options. ‘Higher education,’ definitely. Let’s see what else. Being able to type quickly and browse and click and add it all on the same screen definitely makes me feel that I am going to add more tags to this one because it’s easier to just go: ‘This is relevant, this is relevant, this is relevant.’” In this quote, Participant 14 described an intentional strategy of addition via browsing, knowing that her aboutness statement lacked exhaustivity and that the autocomplete vocabulary could help her fill in the gaps and find additional

terms. She, like many participants who used a browsing strategy, did not seem to worry about potentially redundant indexing terms—assigning both “Higher Education” and “Colleges and universities”—or general and specific terms that, when assigned together, broke the principle of specific entry—assigning “Education,” “Blacks—Education,” and “African Americans—Education”. Instead, these participants enthusiastically added as many indexing terms as they could find, which suggested that they valued exhaustivity even at the expense of other indexing best practices and that they felt in control and supported by the form of vocabulary control that they were using while they made these addition transformations. It is important to remember, though, that while browsing was a noteworthy behavior of inexperienced indexers and taggers observed in this study, that browsing for indexing terms did not result in significantly more addition transformations. In fact, minimal vocabulary control—which has no browsing option—had the most addition transformations.

Finally, another participant behavior that was associated with addition transformations was their use of bibliographic apparatus during subject representation, which is discussed more thoroughly below. In short, some participants used the bibliographic apparatus in the book that they were indexing or tagging instead of or in addition to using their aboutness statements, and these participants tended to make more addition transformations. These were almost a different class of addition transformation, as they were not particularly related to participants’ interaction with vocabulary control, but instead about their approach to using bibliographic apparatus.

Based on this discussion of addition transformations, this study found: 1) that addition transformations are used by inexperienced indexers and taggers more than any other form of transformation, including suppression, 2) that inexperienced indexers and taggers use addition transformations significantly more while working with forms of vocabulary control typically associated with tagging (minimal control and passive control) than forms of vocabulary control associated with subject indexing, 3) that ease of use—being able to type in tags as they think of them—is a primary driver of addition transformations for inexperienced indexers and taggers, 4) that, when working with forms of vocabulary control associated with tagging, inexperienced indexers and taggers respond to uncertainty by using addition transformations to increase exhaustivity, even at the cost of redundancy, and 5) that, while

vocabulary control has a significant effect on addition transformations, indexers' and taggers' approaches to using bibliographic apparatus instead of or in addition to aboutness statements during subject representation also plays a significant role in the rate of addition transformations.

Summary of Transformations

The preceding sections discussed, in depth, each type of transformation observed during the study and found that there was not a simple relationship between vocabulary control and coextensiveness. Instead, analysis of each type of transformation revealed complex dynamics that depended on participants' approaches to indexing and tagging. Some large trends did emerge, though. First, as shown in [Table 5.12](#), 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. Second, those high-level trends and the detailed analysis of each form of transformation suggested a relationship between coextensiveness and control. This relationship was not, however, defined by feeling in control leading to high coextensiveness. While the forms of vocabulary control that participants ranked as making them feel more in control tended to have more terms unchanged during translation, they also had more addition transformations and fewer suppression and generalization transformations. Finally, this analysis revealed a noteworthy behavior of novice and advanced beginner indexers and taggers: a browsing technique—either by typing the roots of terms (i.e., “rac...”) in an autocomplete UI 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.

5.2.3 Analysis and Discussion of Participants' Self-Reported Degree of Coextensiveness

While the previous section took a primarily quantitative approach to understanding the coextensiveness of the aboutness statements and indexing terms that participants produced while indexing and tagging, this section takes a more qualitative approach and considers how participants reflected on coextensiveness through a series of interview questions that

	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 5.12: Comparison of the frequency of different types of additive, subtractive, and substitutive transformations performed by participants when they translated their aboutness statements to indexing terms while using four different forms of vocabulary control (Minimal, Passive, Flexible, and Rigid). The taxonomy of transformations is based on a study by Bertrand and Cellier [12]; however, two transformations in that taxonomy—simplification and dissociation—were not observed in the current study and were not included in the table. “Unchanged” is not a form of transformation described by Bertrand and Cellier but represents complete coextensiveness between a concept in the aboutness statement and that concept’s representation in an indexing term. Note that suppression and generalization increased as vocabulary control became more rigid and that addition and “unchanged” increased as vocabulary control became less rigid.

they were asked after completing indexing and tagging tasks. By analyzing participants’ perceptions of the coextensiveness of their aboutness statements and their explanations for why their aboutness statements and indexing terms were coextensive or not, this section aims to get at participants’ experiences with the indexing and tagging process and how different forms of vocabulary control affected those perceptions.

The data in this section came primarily from the series of retrospective interview questions that asked participants to reflect on how coextensive their aboutness statements and indexing terms were, which were particularly similar and why, which were particularly different and why, and what most influenced the coextensiveness of their aboutness statements and indexing terms. Participants’ responses were recorded, transcribed, and coded. The codebook was built inductively and is available in [Appendix K](#). Of the 18 total participants in the study, only 6 participants self-reported the coextensiveness of their aboutness statements and indexing terms as “Strong coextensiveness” (count=3) or “Somewhat strong

coextensiveness” (count=2), suggesting that overall participants felt that their indexing and tagging work lacked coextensiveness. Those who self-reported “Strong coextensiveness” or “Somewhat strong coextensiveness” had a higher overall rate of concepts that remained unchanged during translation (0.24 unchanged concepts for each transformation) than participants as a whole (0.17 unchanged concepts for each transformation), suggesting that, as a whole, participants’ perceptions of coextensiveness matched the measured coextensiveness; however, not all participants’ self-reported coextensiveness matched the quantitative data. For example, Participant 9’s response was coded as “Somewhat strong coextensiveness,” but she had only one unchanged concept across the four books that she tagged and indexed (0.02 unchanged concepts for each transformation). Similarly, Participant 7’s response was coded as “Strong coextensiveness,” but she had a relatively low rate of unchanged concepts (0.10 unchanged concepts for each transformation). These data suggested that there was some connection between coextensiveness as measured quantitatively in the previous section and coextensiveness as experienced and self-reported by participants, but that qualitative analysis of self-reported experiences was important to the study because it revealed participant experiences that coding and counting transformations did not necessarily reveal. For example, Participant 4’s aboutness statement included the phrase “system of racial castes globally and within the United States.” Because their subject headings did not include the concept of “global” or “globally,” the omission of that concept in indexing terms was coded as a suppression transformation. However, Participant 4 reported being satisfied with the coextensiveness of their aboutness statement and subject headings in this case because they felt that “Caste—India” was a close match to their original intent. This section explores these types of self-reported data not to contradict the previous section but to provide an additional perspective and to better understand participants’ experiences with the process of translation. Participants provided a wide range of responses to the open-ended questions that were analyzed for this section, but some themes emerged during coding, and these themes are discussed in the subsections below.

Connection and Lack of Connection Between Subject Analysis and Subject Representation

Participants diverged dramatically in how closely they viewed the connection between subject analysis and subject representation in the subject indexing and subject tagging processes. Some participants used the output of the subject analysis step—their aboutness statement—as a direct input to the subject representation step, moving through the words and phrases in their aboutness statements as candidate subject headings and tags. This phenomenon, coded as “Direct use of aboutness statement,” was identified eight times by participants. For example, Participant 24 stated, “Overall, [my aboutness statements] pretty much matched what I ended up entering [as indexing terms] because I looked back to them a lot to see what terms I wrote down, to see if anything popped up to me that I could type in and search.” Similarly, Participant 14 stated, “I usually started with the things I had written down, so I started with those terms.” When participants perceived a strong process connection between subject analysis and subject representation, that perception tended to be consistent across different forms of vocabulary control. For example, Participant 5 described their process of translation and representation for three types of vocabulary control: “For autocomplete, I just typed in the terms in my aboutness statement and most of them were there, so I just used them. For the *Sears List*, I went down the list in the aboutness statement and flipped through the book and looked for terms. For free tagging, just entered the words from aboutness statement.”

Other participants did not perceive a strong process connection between subject analysis and subject representation, despite the study script explicitly telling them to “translat[e] your aboutness statements into indexing terms” during the subject representation step. Many participants did not even look at their aboutness statements after writing it, with one participant—perceiving no further use for their aboutness statement—trying to hand the aboutness statement to the researcher after writing it. Another participant produced aboutness statements within a few minutes of receiving a book, only to ignore the aboutness statement while closely analyzing the book during the subject representation step. This phenomena of using the book to be indexed instead of the aboutness statement, coded as “Bibliographic apparatus,” was identified seven times. Participant 9 was particularly

aware that they had disconnected their aboutness statement from the subject representation step: “[My main influence was the summaries in the books. I didn’t really use the aboutness statements in my aboutness statement. I didn’t really use the extra step that I had done.” However, this process disconnect between subject analysis and subject representation did not always correlate with participants feeling that their aboutness statement and indexing terms lacked coextensiveness. For example, Participant 9, who self-reported somewhat strong coextensiveness between aboutness statement and indexing terms, said, “Interestingly, I didn’t really look at my aboutness statements while I was tagging it. I just turned right back to the text. Because the text is still going to be more informative than my aboutness statement, which has only two sentences. It was easier to go back to the source because it had more.”

One reason that participants expressed a disconnect between subject analysis and subject representation was participants’ perception that aboutness statements and indexing purposes had different purposes or that they were different writing genres. This perception was captured in the code “Aboutness different from representation,” which was identified 10 times. For example, Participant 14 stated, “I saw aboutness statements as a micro-outline that included things like ‘proposal for the future’ which I didn’t use as a tag. The aboutness statement was for identifying things that are not subjects but more what it is.” Participant 21 contrasted their goals in writing an aboutness statement and choosing indexing terms, noting, “I tried too hard to make the aboutness statements sound good or be a blurb.” Similarly, Participant 23 noted, “[The aboutness statement] was more prose than subject headings to represent the book.” In contrast, some participants treated the aboutness statement more as a draft to gather ideas, including ones that they did not plan to use directly, as described by Participant 15: “I took notes not to transcribe but to generate ideas and help my brain. I wrote down ‘ego,’ and I’m not going to tag ‘ego.’ But seeing it next to other concepts helped me think about the bigger picture.”

These two different approaches to the connection between subject analysis and subject representation—directly using terms from the aboutness statement during translation and aboutness and disconnecting the aboutness statement from indexing terms— were observed with roughly the same frequency (coded nine times and seven time, respectively) and there

was no obvious pattern, demographic or otherwise, for who used one strategy or the other. This suggests that while some novice and advanced beginner indexers and taggers follow Langridge's and Lancaster's call to keep the steps sequential and discrete when guided to do so, many others follow the pattern that Šaupel and others observed in experienced indexers of blurring the lines between subject analysis and subject representation. This finding warrants further investigation, as co-mingling analysis and representation might not be dependent on knowledge of the indexing language.

Effect of Different Forms of Vocabulary Control on Coextensiveness

Quantitative analysis of the types of transformations that occurred when participants translated aboutness statements to indexing terms showed the complex relationship between forms of vocabulary control and coextensiveness. This section further investigates that relationship by analyzing participants' reflections on the coextensiveness of their aboutness statements and the subject headings and tags that they ultimately used to represent the books in the study. 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). In contrast, other codes about influences on coextensiveness were identified much less frequently: "Influence of order (learning effect)" (count=4), "Influence of lack of domain knowledge" (count=3), "Influence of personal bias" (count=1), "Influence of user focus" (count=1). The "Influence of vocabulary control" and "Vocabulary control effect on term choice" codes co-occurred in some cases, but were tracked separately to reflect, respectively, general statements about the influence of vocabulary control and specific statements about choosing, transforming, or suppressing a term because of a form of vocabulary control. For example, Participant 22 made the following general statement about vocabulary control: "The way in which I was locating or deriving terms from the *Sears List* or from *LCSH* or autocomplete was definitely a big influence.

It was steering.” And Participant 17 made a specific statement about a concept that they transformed during translation: “I matched ‘white racial equilibrium’ with, I think I related it back to ‘racism’ or ‘equality.’”

Participant 9 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 Library of Congress and it wouldn’t exist, and I was just like, okay, I guess I don’t have that tag. 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, Participant 9 noted that minimal vocabulary control allowed them the freedom to choose their 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. For example, Participant 24 reported a similar experience of not

finding a term in the *LCSH* that she wanted to use, but she did not express feelings of being controlled, saying in an upbeat tone, “I couldn’t find something and I said ‘okay.’” Finally, in the extended quote above, Participant 9 described the browsing technique that was associated with substitution and addition transformations for passive control and flexible control in the quantitative analysis of coextensiveness. Participant 9 spoke specifically about browsing the *Sears List* and finding narrower terms that resulted in addition and precision transformations. This quote illustrated and corroborated the quantitative finding that passive control and flexible control were more likely to lead to addition and precision transformations by supporting participants’ browsing processes.

Other Influences on Coextensiveness

While participants’ perception of the relationship between subject analysis and subject representation and the influence of different forms of vocabulary control were the most frequently identified codes related to coextensiveness, participants identified two other influences: “Influence of order (learning effect)” (count=4) and “Influence of lack of domain knowledge” (count=3) that warrant discussion.

Participants in the study were novice and advanced beginner indexers and taggers, and many participants reported that the study was their first hands-on experience with subject indexing. Therefore, this study, unlike studies that observed more experienced indexers, had an opportunity to observe and hear reflections about participants’ earliest experiences with the subject indexing and subject tagging processes—and the rapid learning that some participants experienced between indexing or tagging their first book to indexing or tagging their fourth book. Participants self-reported two main types of order effects through their responses to interview questions: 1) learning to anticipate subject representation while working on subject analysis to make the overall process more efficient, and 2) identifying and remembering key indexing terms within a domain and assigning those indexing terms for other documents and while using other forms of vocabulary control. Specific to the first point, Participant 22, who described their previous indexing and tagging experience as “0,” noted that for the first three books that they “didn’t have [their] system in place yet,” but

that they “got wise” by the last book that they tagged and indexed, and that they were better at “noting potential keywords to go in the aboutness statement.” This reflection suggests that, at least for some participants, practicing the full processes of subject indexing and subject tagging even just a few times quickly led them to more efficient processes, including anticipating indexing terms, which is a behavior more associated with experienced indexers [116]. However, that anticipation strategy was not necessarily for a specific controlled vocabulary, but more generally for the types of words that might make good candidate indexing terms. Other participants adapted their process even more, moving from paragraph-length prose aboutness statements in their first round or two of subject analysis to streamlined lists of keywords for subsequent rounds. For example, [Figure 5.3](#) shows Participant 5’s aboutness statement for the third book that they tagged or indexed, *Mismatch*. The short list of six concepts contrasts with the full page of notes and paragraph-length prose that Participant 5 wrote for the first book that they indexed.

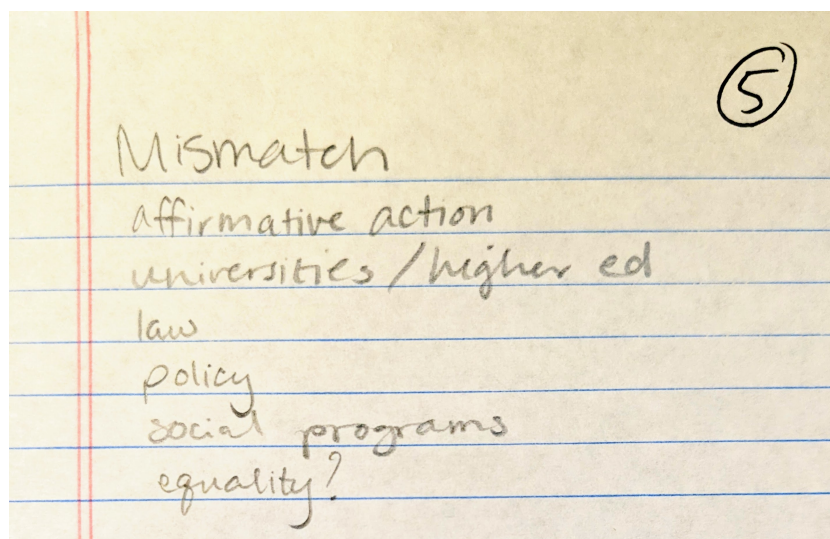


Figure 5.3: Participant 5’s aboutness statement for the third book that they tagged or indexed, *Mismatch*. This streamlined aboutness statement, with a short list of six concepts instead of paragraph-length prose illustrates a pattern that some participants exhibited of learning to write aboutness statements in anticipation of assigning subject headings and tags.

Specific to the second point, in some cases, learning to anticipate was not just about

a general strategy but was about specific terms. For example, Participant 11 noted that her experience with not finding the subject heading “inequity” in the *Sears List* for her first book and using “inequality” instead shifted her thinking for indexing and tagging subsequent books: “When I was doing my analysis, I was anticipating what the vocabulary may or may not have in it, which shifted some of the terms that I chose in the first place.” While Participant 11 indicated that their learned anticipation resulted in substitution and subtraction across different forms of vocabulary control, other participants applied with they learned in additive ways. For example, Participant 17 found the subject heading “Race relations” in the *Sears List* while indexing their first book, remembered it as a good subject heading or tag, and ended up applying it to all four books that she indexed or tagged.

The second influence on coextensiveness discussed in this section was represented with the code “Influence of lack of domain knowledge.” Domain knowledge has been identified in theory [67, 68] and in empirical studies [12] as an important factor in how people index and tag documents. This study collected data about participants’ domain knowledge by asking participants to indicate how familiar they were with a sampling of books about racial and social justice. While that data showed no particular pattern, interestingly, the participants who raised the code “Influence of lack of domain knowledge” included the participant with the highest score and the participant with the lowest score for familiarity with racial and social justice literature, suggesting that indexers and taggers with all levels of domain knowledge can be self-critical and question their domain expertise. Participant 7, who was the participant least familiar with racial and social justice literature, described how not knowing the book’s domain or perspective well—or disagreeing with the book’s perspective—can affect translation and coextensiveness:

For *The End of Racism* and *Mismatch*, I can understand what the book is about but to connect them to subjects that I understand was very difficult because I don’t understand the concepts in the way that they are written about in the books. So, to translate that to subject headings was very difficult to me because the translation doesn’t feel as direct... *The End of Racism* is about racism, but

what its saying about racism doesn't ring true to what I think of racism.

This quote showed that domain expertise was not important just for subject analysis but was important specifically for translation and for an indexer or tagger to be able to communicate their intended meaning through an indexing term. Participant 17, who was the participant most familiar with racial and social justice literature, also indicted that her lack of domain knowledge affected coextensiveness but was more willing to trust authors' vocabulary choices over her own because they know the domain better: "I am not familiar with the genre, but I am assuming that the authors are familiar with it so I am assuming that they will use their subject vocabularies to their advantage. So, if I kept seeing words pop up [in a book], I am assuming that's the subject vocabulary." While exhibited different approaches to subject representation, Participant 7 and Participant 17 expressed some lack of confidence and lack of control in the translation process because of their perceived lack of domain knowledge.

5.2.4 Analysis and Discussion of Bibliographic Apparatus and Coextensiveness

One theme that emerged during the study was that participants treated the relationship between subject analysis and subject representation differently. Some participants treated subject analysis and subject representation sequentially, and they used their aboutness statements as the primary or exclusive input to the subject representation step. Other participants relied more heavily on the bibliographic apparatus (table of contents, summaries, etc.) of the book that they were indexing or tagging instead of or in addition to their aboutness statements during subject representation. These different strategies affected the coextensiveness of aboutness statements and indexing terms, and this section investigates that difference by analyzing data collected about the bibliographic apparatus that participants used during the study, including when they used it and for how long.

During the study, 600 total events in which participants looked at bibliographic apparatus were logged and coded. Of the 600 total bibliographic apparatus events, 510 occurred during subject analysis (481 before participants starting writing aboutness statements and 139 while participants wrote aboutness statement) and 80 occurred during subject representation. This section, which focuses primarily on coextensiveness, investigates those 80

events that occurred during subject representation in detail, but first it considers the overall use of bibliographic apparatus.

Participants' Use of Bibliographic Apparatus

As participants indexed and tagged books, I logged their actions, tracking which bibliographic apparatus they looked at, how long they looked at that bibliographic apparatus (brief, extended, full read), and at which stage of the subject indexing or subject tagging process they looked at the bibliographic apparatus (initial subject analysis, while writing aboutness statement, subject representation). The full codebook for bibliographic apparatus is listed in [Appendix K](#). I coded a total of 562 instances of participants looking at bibliographic apparatus, which meant that, on average, participants looked at bibliographic apparatus 7.81 times for each book that they indexed or tagged. The minimum rate for these bibliographic apparatus events for a participant was 4.5 events per book, for participant 9. The maximum rate of bibliographic apparatus events for a participant was 13.75 events per book, for participant 20. If a participant looked at a bibliographic apparatus for a specific book more than once, I logged each instance in most cases. For example, participant 23 looked at the back cover summary of *White Fragility* early in their analysis process and then, after looking at other parts of the book, returned to the back cover summary while writing their aboutness statement. In this case, the code “Back cover summary” was logged twice. [Table 5.13](#) shows the 10 most used bibliographic apparatus in the study. Overall, the novice and advanced beginner indexers and taggers in the study focused on the bibliographic apparatus that indexing literature and standards recommend. Note that the combined total of logged codes for back cover summaries and front flap summaries represented the most used form of bibliographic apparatus, with 111 total logged codes. Participants looked at summaries, on average, 1.54 times for each book that they tagged or indexed. When they looked at summaries, participants tended to look at them for a significant amounts of time. Except for 4 cases, every time a participant indexed or tagged a book, they registered at least one summary code with “Extended” or “Full read” indicated for the time. Most of the “Brief” instances for summaries were when participants returned to reference the summary

after having read it a first time.

Code	Count
Table of Contents	94
Front cover	76
Front flap summary	59
Back cover summary	52
Front matter (general)	48
Praise (back cover or front matter)	38
Preface or forward	33
Biographical info about author	32
Chapters in body of book	25
Index at back of book	24

Table 5.13: The 10 types of bibliographic apparatus that were most used by participants throughout the study. Note that some counts are higher than the total number of indexing and tagging tasks in the study (72) because participants used the same bibliographic apparatus multiple times during the same task. Note also that summaries, both front flap and back cover, were used 111 times total. Overall, participants focused on the bibliographic apparatus that indexing literature and standards recommend.

The Use of Bibliographic Apparatus During Subject Representation

While participants focused on the bibliographic apparatus on which they were expected to focus, *when* they focused on these apparatus was somewhat unexpected. The experiment was designed around a sequential process. Participants would perform subject analysis, produce an aboutness statement, and then translate that aboutness statement into a subject representation. However, many participants relied heavily on bibliographic apparatus in addition to or instead of the aboutness statement during subject representation. This study coded 80 instances of participants looking at bibliographic apparatus during subject representation. This section analyzes those instances to understand the strategies that participants used and to determine the effect of those strategies on coextensiveness and

control.

Participants' use of bibliographic apparatus during subject representation fell into three groups: 1) no use of bibliographic apparatus during subject representation, 2) limited use of bibliographic apparatus for reference and confirmation during subject representation, and 3) extensive use of bibliographic apparatus during subject representation. For example, Participant 13 used bibliographic apparatus during subject representation more than any other participant, reading extensively during representation for each book that she indexed and logging 17 bibliographic apparatus events during representation. [Table 5.14](#) shows how many participants fell into each of group for use of bibliographic apparatus during subject representation.

Group	Number of Participants
No use of bibliographic apparatus during subject representation	3
Limited use of bibliographic apparatus for reference during subject representation	10
Extensive reading and use of bibliographic apparatus during subject representation	5

Table 5.14: Counts of three groups of participants based on their use of bibliographic apparatus during subject representation.

The five participants who used bibliographic apparatus extensively during the subject representation step overall had a much lower degree of coextensiveness (9.5%)than other participants (17.6%). Two types of transformations were particularly prevalent for these five participants: suppression (count=65) and addition (count=165). The rate of suppression for the five participants who used bibliographic apparatus extensively during the subject representation step was 22.2% across all forms of vocabulary control, which was, surprisingly, lower than the 30.3% rate of suppression for other participants. However, the rate of addition for the five participants who used bibliographic apparatus extensively during the subject representation step was 56.5%, which was markedly higher than the 31.5% rate of addition for other participants. The high rate of addition makes sense because this group

of participants was consulting the book directly and not necessarily working with their aboutness statements. Interestingly, these higher rates scaled up relatively equally across all four forms of vocabulary control, with minimal and passive vocabulary control having higher rates of addition than flexible and rigid vocabulary control—just like in the study as a whole.

Summary of Findings Related to Bibliographic Apparatus

Based on these analyses of participants' use of bibliographic apparatus during subject indexing and subject tagging, this study found: 1) that inexperienced indexers and taggers use a variety of bibliographic apparatus throughout documents, but, despite their lack of experience, they focus on the bibliographic apparatus that are most recommended by subject indexing literature and standards, 2) that some inexperienced indexers and taggers prefer to use bibliographic apparatus extensively during subject representation, even in an experimental setting designed to make the aboutness statement the primary input for subject the subject representation step, suggesting that even inexperienced indexers and taggers are inclined to break the sequential workflow of the subject indexing and subject tagging processes championed by Langridge and Lancaster, even when they cannot anticipate using a specific vocabulary or form of vocabulary control, and 3) that indexers and taggers who use bibliographic apparatus extensively during subject representation, somewhat unsurprisingly, have a lower rate of coextensiveness (mostly due to addition transformation). However, those differences in coextensiveness were largely independent of the effect of different forms of vocabulary control.

5.3 Vocabulary Control, Critical Cataloging, and Racial and Social Justice

This study was inspired by critical catalogers like Sanford Berman, who pushed for changes to problematic headings in the *LCSH* while at the same time acknowledging that indexing languages like the *LCSH* are unlikely to be superseded—despite catalogers' frustrations with their shortcomings—because they are deeply embedded in our information ecosystem [10]. Many of those shortcomings and problematic headings are related to racial and social justice, for example, the “Illegal aliens” and “Indians of North America” headings in the

LCSH which have been discussed throughout the study. These problematic headings as well as headings that are arguably missing from controlled vocabularies, structural biases in indexing languages [105], a lack of support around indexing documents about racial and social issues, and the promise, demonstrated by Melissa Alder [1] and others, for different forms of vocabulary control to authentically represent marginalized groups, all led this study to ask research sub-question C: How do indexers' and taggers' experiences while working with different forms of vocabulary control inform criticisms of controlled vocabularies, particularly for documents about 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 a controlled vocabulary and how respond in those situations, participants' reflections on how their identities (including racial identities) informed their indexing and tagging processes, participants' challenges in indexing and tagging books with viewpoints on racial and social justice that they disagree with, and the effect of different forms of vocabulary on participants' ability to navigate complex racial and social issues in indexing and tagging.

5.3.1 Working with and around Problematic and Missing Indexing Terms

Problematic authorized headings in controlled vocabularies, like "Illegal aliens" and "Indians of North America" in the *LCSH*, were a primary motivator for this study because they situate subject indexers at the center of a difficult situation: Use the authorized term that supports cataloging objectives but causes harm? Or use an unauthorized term that shows more empathy, but which can be difficult implement in a catalog and might made the resource difficult to find? Similarly, if a concept is not represented in a vocabulary such as the *LCSH*, indexers have little recourse if they need to adhere to authorized headings in their catalogs. As previously quoted in this study, Participant 5 dramatically communicated this theme of feeling controlled by a controlled vocabulary in these difficult situations: "Oh, man, with the *LCSH*, you gotta use some terms that hurt your soul sometimes." While

the *LCSH* has introduced a process for suggesting changes to headings and the addition of new headings via SACO, that process is relatively slow. Other forms of vocabulary control support easier and faster methods for using non-problematic indexing terms, although some, like the flexible control of the *Sears List*, still require non-trivial effort by the indexer or tagger to assign the non-problematic term. This study was designed to compare these different forms of vocabulary control, and this section considers specific cases when participants encountered a problematic heading or missing heading, their responses to the situation, and how vocabulary control informed their course of action.

Many of the problematic or missing indexing terms that participants encountered during the study were problematic or less than ideal for subtle reasons. For example, Participant 11 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.” (The distinction between equity and equality is a relatively recent rhetorical development [101].) Because she was using the *Sears List*, she could have added a local subject heading of “Equity,” but she was not confident about adding local indexing terms—and, if she had been, she might have found it difficult to disambiguate from the financial meaning of *equity*. The end result in this case, was that the resource was not exhaustively described by indexing terms and, as discussed previously, Participant 11 was disinclined to use the indexing term “Inequity” or “Equity” with other forms of vocabulary control. 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 Participant 11 to assign “Inequity” and convey the meaning that they intended instead of suppressing the concept. Adler recommends an approach that could help to fill term gaps in controlled vocabularies: “On a practical level, libraries of all types and sizes should support and encourage metadata librarians and catalogers to augment the catalog with local data, create local and subject-specific classifications and subject access

tools, encourage participatory and social cataloging, and invent alternative ways to map knowledge in the library” [2].

In another case, Participant 25 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 study did not try to evaluate whether “Multiculturalism” or “Pluralism (Social sciences)” was a better choice of indexing term, instead focusing on the *process* that the participant followed. In this case, that process showed the importance of scope notes in guiding indexers and taggers while they work with complex social and racial issues. The inexperienced 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 their scope notes along with 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 was not particularly relevant to the books in the study, but which Participant 4 encountered in the *LCSH* while trying to assign subject headings related to India for the book *Caste*. The following extended quote comes from Participant 4’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 subdivided 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 an inexperienced indexer was quick to note its presence and flag it as problematic. Second, problematic headings can exercise control over indexers and taggers, even when the indexer or tagger does not need to use the heading or a related term. In this case, Participant 4 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.”

The indexing terms that study participants cited most often as problematic were “Whites” and “Blacks.” Because these single-word designators emphasize skin-color and erase humanity, they have largely been replaced in contemporary usage by phrases such as “White people,” “Black people,” and “People of color.” However, “Whites” and “Blacks” remain authorized headings in the *Sears List* and the *LCSH*, and that dated language stood out to study participants as problematic and offensive. Participant described their reaction to these terms in context with her understanding of how they are used (or not used) in contemporary language: “I balked at the term ‘Whites’ because I know there is also a term ‘Blacks’ in *LCSH* and that ‘Blacks’ shouldn’t be used.” Participant 20 felt that they had to use these terms, though, saying that they felt “stuck” with “Blacks” and “Whites” in the *LCSH*. They elaborated: “I feel like ‘Blacks’ (plural) has a negative or pejorative connotation, but

I don't really have much evidence to support that claim. Just a feeling of discomfort...I think what made me uncomfortable about the 'Blacks' subject heading and this 'Whites' one too is that's just not how we really talk about groups. We'd say 'white people' or 'black people' or 'white community' or 'black community,' so it is outdated language." Participant 13 searched the *Sears List* for "White people," "Whites," and "Caucasian" without success. They asked, incredulously: "Is there no white people in here?!" and eventually chose to add "White people" as a local subject heading. Similarly, Participant 5 noted the "Whites" subject heading in the *LCSH* but quickly chose to use "White people" while working with minimal vocabulary control "because I can do that." In contrast, other participants were more willing to conform to a controlled vocabulary. For example, Participant 8 had "White people" in her aboutness statement but found the *LCSH* subject heading "Whites" and quickly chose to use that authorized heading instead of pursuing her original form of "White people." She did not appear upset to make the transformation.

Finally, it is worth noting that while none of these cases involved minimal vocabulary control, minimal control is not a solution to problematic indexing terms. While minimal vocabulary control provides no authorized or suggested indexing terms to indexers and taggers, taggers can and do assign problematic terms, either intentionally or through implicit bias. Problematic tags were not an issue in this controlled experiment, but they are rampant on web platforms and often require extensive moderation. Perhaps the best approach, then is for editors of indexing languages to avoid being overly cautious about updating authorized and suggested indexing terms, even at the cost of frequent catalog revisions.

Based on this discussion 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 assigned if controlled vocabularies were not overly cautious about updating indexing terms related to racial and social issues, 5) that eliminating vocabulary control 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.

5.3.2 Interrogating Whiteness in Subject Indexing and Subject Tagging

One of the main themes of critical cataloging is that white perspectives (along with heterosexual, male, cisgender, and Christian perspectives) are privileged in KOS and controlled vocabularies, and that these biases are systemic and harmful to those not in the privileged groups. While these biases have long histories, they also have current realities. One of those realities is that librarians and catalogers in the United States are overwhelmingly white. According to the 2017 ALA Demographic Study [112], 86.7% of members identify as white. This study chose to recruit participants who roughly reflected this racial demographic, so that it could interrogate how white racial identities shaped indexers' and taggers' processes as they worked with books that addressed racial and social issues. The participants in this study were less white than the ALA, but with 66.7% of participants identifying as "White" or "European American," the study population was still majority white and somewhat representative of the racial demographics of librarians and catalogers in the United States. Participants were also asked about their familiarity with representative books about racial and social justice to understand their familiarity the literature, but no clear trends emerged with respect to participants' familiarity with books about racial and social justice or how that familiarity might affect their indexing and tagging processes. 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. Themes that emerged from this analysis included: "Feeling personal identity (including racial identity) controlled" and "Feeling unqualified or unsure about how to represent a difficult topic (like racial or social justice)." 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 for participants who identified as people of color and participants who identified as white. While participants might have been motivated to assign accurate and equitable indexing terms in general or because they were being observed in a study, some participants specifically identified the subjects of the books as a reason to take particular care to do excellent work. For example, Participant 5, 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 so they could index and tag in a way that showed care for people and works that had been harmed in the past. These sentiments were expressed broadly but were also addressed to specific issues, such as how to work with and around the problematic and missing indexing terms discussed above.

Another theme that emerged—that 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. Participant 16, 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 did assign very broad terms, such “Race,” “Racism,”

and “Historical.” In contrast, Participant 16 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. The tags that she assigned, including “Race awareness,” “Race relations,” “Whiteness,” and “Social justice,” were arguably more specific and, by her self-report, more authentic.

Participant 14 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—Participant 14 distanced herself from the Black identity and lived experiences of Isabel Wilkerson, which were 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. Participant 17, 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.” Similar to Participant 14, Participant 17 also assigned a sparse set of tags that lacked specificity.

While this discussion has so far focused on *Caste*, it is important to note that *The End of Racism* which also addresses the history of race and racism in the United States and also was written by a person of color; However, it did not elicit the same emotions as *Caste*, likely because it questions the existence of racism in the United States and is unsympathetic to movements for racial and social justice. Study participants, both white and people of color, responded negatively to the views expressed in *The End of Racism* and, while they stated that they did not feel confident about representing the contents of the book, it was because they disagreed with the views expressed in the book—not because they might fail to

represent the lived experience of Dinesh D'Souza. The dynamics of indexing and tagging *The End of Racism* are discussed in depth in the following section. Finally, multiple participants noted that the authors of *Mismatch* were two white men and questioned their authority to make claims about what was best for students of color. The reaction to this book was not as strong as the reaction to *The End of Racism*, but it showed that white participants were worried about perpetuating harmful viewpoints presented by authors who had not lived the experiences that they were documenting in their book. This reaction contrasted with participants' reactions to *White Fragility*, which was generally viewed favorably and seen as a familiar book written by a white author that participants could identify with. Participants wanted to index and tag *White Fragility* in a way that would increase the book's visibility because, unlike *Mismatch*, they could identify with the book and the author without so much skepticism.

Based on this discussion 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 inexperienced indexers—at least those studying in a library science program—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, 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 4) white indexers and taggers are skeptical of white authors' views on race unless they are familiar with the white author's viewpoint and confident that they can authentically represent that viewpoint.

5.3.3 *Indexing and Tagging Books that are Critical of the Conversations around Racial and Social Justice*

Although this study asked questions about the effect of vocabulary control on the subject indexing and subject, a key theme that emerged that centered instead on the views expressed in books that were critical of racial and social justice, participants' reactions to those views,

and how those reactions affected the subject indexing and subject tagging process. This study chose books with different critical perspectives to explore the breadth of indexing and tagging work done in libraries, on the web, and in other collections—and the breadth of challenges faced by indexers and taggers. One of the most challenging critical perspectives that participants encountered in this study came in the form of conservative firebrand Dinesh D’Souza’s intentionally provocative book, *The End of Racism*. *The End of Racism* is a “case against liberal race policies” [37] and challenges many widely recognized facts about race in the United States. When participants engaged with the book, they were incredulous to find, for example, a chapter titled, “An American Dilemma: Was Slavery a Racist Institution?” This section examines how participants reacted to *The End of Racism* and other books in the study, how the indexed and tagged *The End of Racism* and other books after their initial reactions, and how vocabulary control affected their indexing and tagging processes for books that challenged their view on racial and social justice.

Many participants had a strong initial reaction to *The End of Racism*. For example, when Participant 7 was handed the book and saw D’Souza’s name on the cover, she immediately noted that she disagreed with the author and said, “I’m going to be biased about this.” Other participants read the summary on the back cover and were taken aback by the provocative questions posed about race and racism in the United States, some reading the summary out loud in disbelief and others expressing anger. Despite these initial reactions, participants aimed to take a neutral tone while writing aboutness statements. Some editorializing came through in subtle ways, however. Participants used quotation marks around phrases like “black failure”, “liberal crusade,” and “deficient IQ” in their aboutness statements to indicate that those phrases were not their own, but that they came directly from the book. Similarly, participants used hedging language to indicate that they found the books arguments unconvincing. For example, Participant 9 described the book as an “attempt to combat the ‘liberal crusade’ against racism” and Participant 21 used the phrase “hopes to expose the way that attitudes on race make it hard to address real issues.” The use of “attempts” and “hopes” suggested that the participants disagreed with the arguments but still felt that they should include those arguments in the aboutness statement so they could reflect what the book was about.

As participants moved to the working with forms of vocabulary control during subject representation, they continued trying to repress their most negative feelings toward *The End of Racism* and tried to assign neutral and descriptive indexing terms. No participants openly editorialized with their indexing terms or used indexing terms to express their distaste for the positions expressed in *The End of Racism*. This was somewhat surprising because Munk and Mørk's taxonomy of tagging strategies lists "categorization through value judgment" as a common form of tagging; however, even with minimal and passive vocabulary control, participants tried to avoid expressing value judgments. They did, however, reflect on how that was challenging and how that challenge affected their subject indexing and subject tagging processes. Participant 16 was aware of her "negative personal opinion" of *The End of Racism* but focused on choosing *LC* subject headings to help information seekers find the book. Those terms turned out to be quite cautious and generic, though: "Race," "Racism," "Prejudice," and "Nonfiction." Similarly, Participant 9 stated that they found *The End of Racism* to be the most difficult book to tag or index because they did not care about the book as much as other books in the study—they were annoyed by its argument and felt themselves questioning whether it was worth the effort to represent the book with indexing terms. They assigned more subject headings than Participant 7, but they were similarly generic. Their headings included: "Racism," "Public policy," "Society," and "Race." Participant 17 felt that their indexing terms for *The End of Racism* were her least accurate because they tried to contain their personal bias and not offend potential information seekers. Similarly, Participant 11 felt that their representation of *The End of Racism* was their least accurate because their personal biases meant that they could not reflect the book's conservative political leaning—they "lacked words for that." For example, Participant 7, who said, "I'm going to be biased about this" when they first received *The End of Racism*, consciously tried to not show her bias and to be "truthful" with the tags that she assigned. She used the suggested tags provided by passive control's autocomplete feature to assign relatively neutral (but still descriptive) tags like "African Americans—Civil rights," "Racial discrimination," and "Social policy." Participant 14, by considering the consequences of assigning indexing terms, provided a deep insight into the mixed motivations and mixed emotions involved in indexing or tagging a document that you do not sympathize with: "It would be hard for me

to give a fair shake to opinions that I don't agree with, so I am kinda struggling with that right now. I don't want to make this book sound interesting because I don't want anyone to read it. And that sounds horrible because that's technically not what I'm supposed to be doing, so that's a weird feeling to have."

Some participants had a similar, although less dramatic, experiences with trying to index and tag the book *Mismatch*, which controversially argues that "affirmative action hurts students it's intended to help" [114]. When she was handed *Mismatch*, Participant 21 reacted right away to the title and subtitle: "I'm immediately on edge on a book about affirmative action. I want to see who the authors are." Those feelings of being on edge or questioning the authors' viewpoint continued throughout the subject indexing and subject tagging process for some participants. Participant 20 felt that their subject headings for *Mismatch* were insufficiently specific because they "got hung up on evaluating the author's argument instead of identifying the subject." In these cases—and even more so in the many cases involving *The End of Racism*—participants' reactions to the views expressed in the book were the primary drivers of their indexing and tagging processes. This dynamic of the views of the book driving the indexing and tagging process also applied, although to a lesser extent, when participants agreed with the views of the book or found them sympathetic. In this study, participants tended to be more sympathetic to *White Fragility* and *Caste*, and they expressed responsibility and pressure to honor the books with which they agreed. As opposed to admitting, "I don't want to make this book sound interesting because I don't want anyone to read it," participants adopted a strong user-orientation for books that they agreed with so that those books would be more visible in a catalog or folksonomy.

Based on this discussion of themes related to books that indexers and taggers disagree with, this study found: 1) that the views expressed in a book, particularly if they are controversial or do not agree with the indexer's or tagger's view, influence the entire subject indexing and subject tagging process from subject analysis to subject representation, 2) that inexperienced indexers and taggers consciously balance their negative feelings about these books with their responsibility to assign specific, exhaustive, and accurate indexing terms for the book, but feel that they often fail to index and tag these books impartially, 3) that inexperienced indexers and taggers feel additional responsibility to assign specific,

exhaustive, and accurate indexing terms to books that they agree with and that they think should get more visibility, and 4) that, while participants were not particularly satisfied with the indexing terms that they assigned to books that they disagreed with, they were, in many cases, relieved to find terms that someone else had authorized and that they considered neutral to assign to challenging books.

Chapter 6

CONCLUSION

Motivated by criticisms of controlled vocabularies and the emergence of new forms of vocabulary control, this study observed 18 participants as they indexed and tagged books using four different forms of vocabulary control, and then asked them to reflect on their experiences. Their aboutness statements, indexing terms, and reflections on the subject indexing and subject tagging process were analyzed to explore an overarching research question about the effect of vocabulary control on the subject indexing and subject tagging processes and to answer three research sub-questions:

- How do different forms of vocabulary control affect indexers' and taggers' *experiences* as they index and tag documents? When do they feel controlled? In control? Out of control?
- How do different forms of vocabulary control affect the coextensiveness of aboutness statements and indexing terms?
- How do indexers' and taggers' experiences while working with different forms of vocabulary control inform criticisms of controlled vocabularies, particularly for documents about racial and social justice?

The following sections summarize this study's findings for each of these research sub-questions. After addressing these research sub-questions, this conclusion notes some limitations of the study's findings, looks to the future by considering the role of human indexers and taggers in light of Artificial Intelligence (AI) developments, and closes by considering how this study might inform the information seeking scenarios that were described in the introduction of this study and by sharing some hopes for the impact that this study might have.

6.1 *Forms of Vocabulary Control and Feelings of Control*

This study analyzed the concept of *control* in subject indexing and subject tagging and found that different forms of vocabulary control affected indexers' and taggers' experiences, processes, and feelings of control.

When analyzing participants feelings of being in control and being controlled while subject indexing and subject tagging, this study found: 1) that inexperienced 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 inexperienced indexers and taggers consider post hoc control more similar to flexible and rigid control than other forms of vocabulary control associated with tagging because post hoc control has an agent that can control them, 3) 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, with terms missing from a vocabulary, and with someone or something being able to exert control over the indexer or tagger.

When analyzing a second axes for control, from being in control to being out of control, this study found: 1) that inexperienced 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, 2) that flexible control and rigid control are more likely to make inexperienced indexers and taggers feel out of control if they equate being out of control with restrictions and that minimal control is more likely make inexperienced 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 inexperienced participants. This study suggests additional research to understand the relationship between control, autonomy, and minimal guidance for more experienced indexers and taggers.

When analyzing the ease of use of different forms of vocabulary control and the relationship between ease of use and control, this study found: 1) that vocabulary control has a

significant effect on how easy or difficult inexperienced indexers and taggers find the translation and subject representation processes, how long those processes take, and how many indexing terms indexers and taggers assign to documents 2) that the forms of vocabulary control associated with tagging (minimal and passive) are perceived as easier to use, take less time to use, and result in more indexing terms being assigned than forms of vocabulary control associated with subject indexing (flexible and rigid), 3) that inexperienced indexers describe ease of use based on how quickly they can understand and use a form of vocabulary control, how much the form of vocabulary control includes or allows the terms that they want to use, how much feedback and support the form of vocabulary control provides, how complex the form of vocabulary control is, and how constrained or restricted the form of vocabulary control made them feel. These findings showed that while ease of use is not the same as control and ease of use is more dependent on the specific manifestation of a form of vocabulary control (i.e., a print copy of the *Sears List*), ease of use is still a useful metric for understanding how and when indexers and taggers feel in control, out of control, and controlled while doing subject indexing and subject tagging work.

When analyzing participants' reflections on the specificity, exhaustivity, and accuracy of their indexing terms, this study found: 1) that inexperienced indexers and taggers feel that partially enumerated lists of indexing terms help them achieve specificity by providing suggestions and, in some cases, providing guidance to, more specific indexing terms, 2) that forms of vocabulary control associated with subject tagging produce significantly more indexing terms per book than forms of vocabulary control associated with subject indexing, but that indexers and taggers perceptions of the exhaustivity of their indexing terms do not necessarily correlate with the number of indexing terms, 3) that vocabulary control does not necessarily have an effect on the accuracy of indexing terms, but inexperienced indexers and taggers might prefer the balance of freedom and guidance provided by passive vocabulary control when trying to assign accurate indexing terms, 4) that specificity, exhaustivity, and accuracy all depend on an indexer or tagger feeling in control of the subject indexing and tagging process; when inexperienced indexers and taggers feel restricted or unsupported by a form of vocabulary control, they are less likely to feel satisfied with the specificity, exhaustivity, and accuracy of their indexing terms, and 5) that book-centric properties and indexer

and tagger domain expertise likely have a significant effect on specificity, exhaustivity, and accuracy, as suggested by Bertrand and Cellier, and the influence of these factors warrant further study to better understand the subject indexing and subject tagging processes.

Overall, this study found that forms of vocabulary control had a significant effect on inexperienced indexers' and taggers' processes and feelings of control during those processes. That effect manifested in many ways, showing that the relationship between vocabulary control and indexers' and taggers' feelings of control is complex and suggesting that one form of vocabulary control is not best for all situations or all values. For example, when participants were asked which forms of vocabulary control made them feel most in control and most controlled, they reported that minimal control made them feel most in control. But minimal control was also likely to make the inexperienced indexers and taggers in the study feel out of control. Based on these findings, this study does not recommend a specific form of vocabulary control to increase indexers' and taggers' feelings of control, but it does suggest that passive vocabulary control and post hoc vocabulary control show promise as easy to use forms of vocabulary control that balance support and freedom for indexers and taggers.

6.2 *Forms of Vocabulary Control and Coextensiveness*

This study defined coextensiveness 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 analyzed participants' aboutness statements and indexing terms, participants' reflections of the coextensiveness of their aboutness statements and indexing terms, and participants' use of bibliographic apparatus during translation and subject representation to better understand how vocabulary control affected—or controlled—the choices that indexers and taggers made when they translated their aboutness statement into indexing terms.

Based on analysis of suppression transformations, this study found: 1) that suppression might be caused by the form of vocabulary control and by factors outside of vocabulary control, including taggers' and indexers' use of their aboutness statement during the subject representation phase), 2) that the forms of vocabulary control associated with tagging—

minimal and passive control—had the lowest rates of suppression but these rates were not significantly different from the rates for other forms of vocabulary control, and 3) that the flexibility to add local indexing terms in an otherwise controlled vocabulary like the *Sears List* seemed to reduce suppression in comparison to the more rigid approach of the *LCSH*—and affords indexers with more control to represent complex topics around racial and social issues.

Based on analysis of generalization transformations, this study found: 1) that generalization is rare with all forms of vocabulary control for inexperienced indexers and taggers, 2) that when generalization does occur, it is more likely to occur with more rigid forms of vocabulary control that are associated with subject indexing—particularly flexible vocabulary control, and 3) that generalization can reflect an indexer or tagger feeling controlled by an indexing language.

Based on analysis of substitution transformations, this study found: 1) that substitution occurs more frequently with passive control and flexible control than with other forms of vocabulary control, at least for novice and advanced beginner indexers and taggers, who are more likely to browse indexing languages for suggestions, 2) that substitution does not always involve the expected workflow of looking for a concept from the aboutness statement, not finding it, and replacing it with an authorized indexing term; instead substitution often occurs because the indexer or tagger finds additional information during the subject representation step or is satisfied with indexing terms that they find while looking up other concepts from their aboutness statement, and 3) that even when substitution does follow the expected workflow, novice and advanced beginner indexers and taggers do not necessarily feel controlled when they make substitution transformations; instead, they were satisfied with the substitution choices that they made or they used other strategies if they could not find a suitable substitution. This last finding did not match the feelings that librarians expressed around control and lack of control in the “Illegal aliens” case at Dartmouth, which is somewhat surprising. This difference can possibly be attributed to inexperienced indexers and taggers being more willing to accept authorized headings and suggested tags as useful guidance or could be attributed to participants’ use of strategies other than substitution when they were not satisfied with authorized headings and suggested tags. In the “Illegal

aliens” case, librarians did not feel that those options were viable.

Based on analysis of precision transformations, this study found: 1) that vocabulary control does not have a significant effect on precision transformations; however, precision transformations are somewhat less frequent with minimal vocabulary control, which does not offer any suggestions for more specific terms and 2) that novice and advanced beginner indexers and taggers felt in control when they searched or navigated controlled vocabularies and they found more specific terms, suggesting that controlled vocabularies—especially those with enumerated terms—provide useful support to inexperienced indexers and taggers and help them achieve the goal of specificity.

Based on analysis of addition transformations, this study found: 1) that addition transformations are used by inexperienced indexers and taggers more than any other form of transformation, including suppression, 2) that inexperienced indexers and taggers use addition transformations significantly more while working with forms of vocabulary control typically associated with tagging (minimal control and passive control) than forms of vocabulary control associated with subject indexing, 3) that ease of use—being able to type in tags as they think of them—is a primary driver of addition transformations for inexperienced indexers and taggers, 4) that, when working with forms of vocabulary control associated with tagging, inexperienced indexers and taggers respond to uncertainty by using addition transformations to increase exhaustivity, even at the cost of redundancy, and 5) that, while vocabulary control has a significant effect on addition transformations, indexers’ and taggers’ approaches to using bibliographic apparatus instead of or in addition to aboutness statements during subject representation also plays a significant role in the rate of addition transformations.

Based on these analyses of participants’ use of bibliographic apparatus during subject indexing and subject tagging, this study found: 1) that inexperienced indexers and taggers use a variety of bibliographic apparatus throughout documents, but, despite their lack of experience, they focus on the bibliographic apparatus that are most recommended by subject indexing literature and standards, 2) that some inexperienced indexers and taggers prefer to use bibliographic apparatus extensively during subject representation, even in an experimental setting designed to make the aboutness statement the primary input for subject the

subject representation step, suggesting that even inexperienced indexers and taggers are inclined to break the sequential workflow of the subject indexing and subject tagging processes championed by Langridge and Lancaster, even when they cannot anticipate using a specific vocabulary or form of vocabulary control, and 3) that indexers and taggers who use bibliographic apparatus extensively during subject representation, somewhat unsurprisingly, have a lower rate of coextensiveness (mostly due to addition transformation). However, those differences in coextensiveness were largely independent of the effect of different forms of vocabulary control.

Overall, this study found: 1) that, for inexperienced indexers and taggers, suppression and generalization transformations increase as vocabulary control become more rigid, and addition transformations and “unchanged” (no transformation) concepts increase as vocabulary control became less rigid, 2) that, while there is a relationship between participants’ feelings of control and coextensiveness, feelings of control did not necessarily lead to high coextensiveness. While the forms of vocabulary control that participants ranked as making them feel more in control tended to have more terms unchanged during translation, they also had more addition transformations and fewer suppression and generalization transformations, 3) that novice and advanced beginner indexers and taggers use a browsing technique—either by typing the roots of terms (i.e., “rac...”) in an autocomplete UI 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, and this browsing strategy makes them feel more in control, 4) that even when presented with a study design that intentionally separated subject analysis and subject representation, inexperienced indexers and taggers are likely to co-mingle these processes by continuing doing subject analysis work during the subject representation step. These overall findings underline the complexity of the subject indexing and subject tagging processes, show that vocabulary control can affect many different types of transformations when indexers and taggers translate concepts into indexing terms, and show that a lack of coextensiveness between aboutness statement and indexing terms does not necessarily mean that the indexer or tagger lacks control.

6.3 Racial and Social Justice in Subject Indexing and Subject Tagging

Inspired by critical cataloging literature and catalogers' frustrations with inflexible rules around problematic indexing terms like "Illegal aliens" and "Indians of North America," this study aimed to better understand the relationship between vocabulary control, indexers' and taggers' feelings of control, and racial and social justice in subject indexing and subject tagging.

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 assigned if controlled vocabularies were not overly cautious about updating indexing terms related to racial and social issues, 5) that eliminating vocabulary control 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.

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 inexperienced indexers—at least those studying in a library science program—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, 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 4) white indexers and taggers are skeptical of white authors' views on race unless they are familiar with the white author's viewpoint and confident that they can authentically represent that viewpoint.

Based on analysis of themes related to books that indexers and taggers disagree with, this study found: 1) that the views expressed in a book, particularly if they are controversial or do not agree with the indexer's or tagger's view, influence the entire subject index and subject tagging process from subject analysis to subject representation, 2) that inexperienced indexers and taggers consciously balance their negative feelings about these books with their responsibility to assign specific, exhaustive, and accurate indexing terms for the book, but feel that they often fail to index and tag these books impartially, 3) that inexperienced indexers and taggers feel additional responsibility to assign specific, exhaustive, and accurate indexing terms to books that they agree with and that they think should get more visibility, and 4) that, while participants were not particularly satisfied with the indexing terms that they assigned to books that they disagreed with, they were, in many cases, relieved to find terms that someone else had authorized and that they considered neutral to assign to challenging books.

6.4 Validity and Limitations

While this study reported many findings related to the effects of vocabulary control, it is important to remember the limitations of this study. One of the biggest limitations of this study was the experience—or inexperience—of the participants. For example, many participants were interacting with the *Sears List* or the *LCSH* for the first time during the study and many were unfamiliar with how to build synthetic subject headings. When compared to professional indexers, participants likely responded differently to forms of vocabulary control, experienced feelings of control and lack of control differently, and learned more from the first book that indexed or tagged. Therefore, these findings must be qualified with phrases such as “this study found that inexperienced indexers and taggers...”

Another major limitation was that the study was conducted as a controlled experiment and did not observe indexers and taggers in their natural work environment. That research design choice was good for controlling variables and conducting consistent study sessions, but it did create an artificial experience in a lab setting with a think-aloud protocol. Therefore, this study was careful about making claims about how the subject indexing and subject tagging processes work in other environments.

Finally, while the books that participants indexed and tagged in the study presented diverse viewpoints, they were all in the same domain of race, racism, and racial and social justice in the United States. Indexers and taggers would likely encounter different challenges while indexing and tagging books in other domains, so this study does not make claims that these findings can be applied across all domains.

6.5 The Role of Humans in Knowledge Organization

This study intentionally focused on *human* experiences with subject indexing and subject tagging, asking human participants to do the work and trying to understand their emotions, reactions, and struggles as they worked with different forms of vocabulary control. That choice came from a personal belief that subject indexing and subject tagging are socio-cultural practices; that they are about humans communicating with each other; that they are about subject indexers and subject taggers assigning subject headings and tags that help other humans find and understand information. This study argued in [Chapter 3](#) that because of systemic bias in algorithm-based knowledge organization and retrieval and because many machine-based indexing systems use training data provided by human indexers and taggers, that understanding how humans do subject indexing and subject tagging work, how they deal with biases and issues of control, how they advocate for (or do not advocate for or feel they cannot advocate for) for racial and social justice, is even more important with the emergence of machine indexing and algorithm-based KOS.

While this study was being conducted, AI and Large Language Models (LLM) publicly made a major step forward with the release of ChatGPT [\[106\]](#). ChatGPT is not a subject indexing and subject tagging tool, but it is a general purpose chatbot that can answer questions and respond to writing prompts—including prompts asking it to perform subject indexing and subject tagging tasks. In many cases, the answers and replies that ChatGPT returns are factual, relevant, and well-written; however, ChatGPT and other LLM chatbots tend to “hallucinate” [\[58\]](#) or confidently return incorrect information to users.

Interested in seeing how these strengths and weaknesses applied to subject indexing and subject tagging, this study asked ChatGPT to write aboutness statements and assign subject headings and tags to the books used in the study. Some of the responses are available

in [Appendix L](#). Overall, ChatGPT performed reasonably well, following rules for authorized headings and for the most part assigning reasonable subject headings and tags. ChatGPT did, however, make some choices that student participants found inaccurate (i.e., “White supremacist movements”) and problematic (“Whites” and “Multiculturalism”). And, while ChatGPT happily returned *Sears List* headings when it was tested in March 2023, a test in May 2023 met this response: “I’m sorry, but as an AI language model, I do not have access to the *Sears List of Subject Headings* or any other proprietary library cataloging system. However, I can suggest some general subject headings that could be applicable to the book” [\[106\]](#). While this response showed that ChatGPT is not a viable replacement for human indexers and taggers at this point, its “general subject headings” might serve as a good starting point or useful aid for an inexperienced indexer or tagger. That said, it is important to note that as ChatGPT improves (and it inevitably will) that ChatGPT does not attempt to index and tag books in the way that humans index and tag books. Instead, ChatGPT looks at training data—including the words that humans have used to describe the book—and returns some of those words that humans have used in a nicely formatted response. And, before the recent update that discontinued the use of *Sears* or *LCSH* headings, it also looked at indexing languages that were developed by humans. Humans, then, with all their biases, feelings of control, and complex relationships with racial and social justice, are still essential to thoughtful, accurate, and user-focused subject indexing and subject tagging. In the future, though, humans might need to think of themselves as both indexers and taggers *and* as trainers for AI indexing and training models. When they do, being in control of their indexing and tagging processes and avoiding biased, problematic, or inaccurate indexing terms will be even more important because their work will likely be copied and amplified by AI algorithms.

6.6 Closing Words

The introduction to this study began by comparing two scenarios—searching for a book about long-distance running in a public library and searching for blog posts about science fiction on a community website—and established that subject indexing and subject tagging are different approaches to knowledge organization. These scenarios also introduced some

of the problems with subject indexing and subject tagging, such as biased indexing terms and ambiguous tags. This study aimed to better understand those—and other—problems by introducing a model with five forms of vocabulary control and by studying, from the perspective of indexers and taggers, the effect that these forms of vocabulary control had on the subject indexing and subject tagging processes. The results of this study suggest some targeted changes that would be relevant to the scenarios. For example, *LCSH* editors could listen to concerns about “Indians of North America” being a problematic heading, or libraries could use a more flexible indexing language (which happens to not use “Indians of North America” but also afford more flexibility to work around headings) or grant local catalogers the time and resources to create and assign their own headings. *MetaFilter* could introduce passive vocabulary control or post hoc vocabulary control to ensure—or at least encourage—uniform and unique tags. These changes would reflect the nuanced understanding of control and vocabulary control presented in this study, but, of course, they would also require investing time and effort into making the changes.

One hope for this study is that it shows the value to racial and social justice of making these types of changes to indexing languages and to indexing and tagging process, and that it provides concrete data to guide those changes in a way that supports indexers and taggers—or at least inexperienced indexers and taggers like those who participated in this study. A second hope is that presenting vocabulary control as a spectrum and detailing the differences in how inexperienced indexers and taggers work with these systems will inspire people to consider different approaches to vocabulary control, especially forms that balance indexer and tagger control with vocabulary control, such as passive control, post hoc control, and flexible control. Another hope is that this study’s choice to analyze vocabulary control from the perspective of indexers and taggers makes those critical roles a focal point for how people think about designing indexing languages and indexing and tagging processes, because building systems that prioritize indexers’ and taggers’ ability to thoughtfully exercise control should benefit all users of those systems. Another smaller hope is that this study answers not just questions about vocabulary control, feelings of control, and racial and social justice, but also makes a meaningful contribution to the small but vital body of research that has observed the subject indexing process—and that it

represents a meaningful starting point for observing the tagging process. Finally, I hope that this study adds to the critical cataloging conversation and inspires indexers and taggers, information seekers, designers of indexing languages, and anyone who works with Knowledge Organization Systems to be more aware of the control dynamics involved in these systems and how control, vocabulary control, and racial and social justice are intertwined in these complex human systems.

REFERENCES

- [1] ADLER, M. 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, 4 (2009), 309–331.
- [2] ADLER, M. Classification along the color line: Excavating racism in the stacks. *Journal of critical library and information studies* 1, 1 (2017).
- [3] ADLER, M., AND TENNIS, J. Toward a taxonomy of harm in knowledge organization systems. *Knowledge Organization* 40, 4 (2013), 266–272.
- [4] ALLMARK, P. Should research samples reflect the diversity of the population? *Journal of medical ethics* 30, 2 (2004), 185–189.
- [5] AMERICAN LIBRARY ASSOCIATION. Committee on equity, diversity, inclusion and social justice. <http://www.ala.org/pla/about/people/committees/pla-tfedi>. Accessed: 2022-10-17.
- [6] BARON, J. E., AND GROSS, T. Sorrow, fury, helplessness, and cynicism: An account of the library of congress subject heading” illegal aliens”.
- [7] BATES, J., AND ROWLEY, J. Social reproduction and exclusion in subject indexing: A comparison of public library opacs and librarything folksonomy. *Journal of Documentation* (2011).
- [8] BATES, M. J. Indexing and access for digital libraries and the internet: Human, database, and domain factors. *Journal of the American Society for information science* 49, 13 (1998), 1185–1205.
- [9] BEGHTOL, C. Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents. *Journal of documentation* (1986).
- [10] BERMAN, S. *Prejudices and Antipathies: A Tract on the LC Subject Heads Concerning People*. The Scarecrow Press, 1971.
- [11] BERMAN, S., AND GROSS, T. Expand, humanize, simplify: An interview with sandy berman. *Cataloging & Classification Quarterly* 55, 6 (2017), 347–360.
- [12] BERTRAND, A., AND CELLIER, J.-M. Psychological approach to indexing: effects of the operator’s expertise upon indexing behaviour. *Journal of information science* 21, 6 (1995), 459–472.

- [13] BONE, C., AND LOUGHEED, B. Library of congress subject headings related to indigenous peoples: Changing LCSH for use in a canadian archival context. *Cataloging & Classification Quarterly* 56, 1 (2017), 83–95.
- [14] BOWKER, G. C., AND STAR, S. L. *Sorting things out: Classification and its consequences*. MIT press, 2000.
- [15] BRAINWANE. Hugo award finalists include a story in tweeted images. <https://www.metafilter.com/194942/Hugo-Award-finalists-include-a-story-in-tweeted-images>. Accessed: 2022-05-6.
- [16] BRANCHE, C. L. Diversity in librarianship: Is there a color line? *The 21st-century Black librarian in America* (2012), 203–206.
- [17] BRAUN, V., AND CLARKE, V. Can i use ta? should i use ta? should i not use ta? comparing reflexive thematic analysis and other pattern-based qualitative analytic approaches. *Counselling and Psychotherapy Research* 21, 1 (2021), 37–47.
- [18] BRAUN, V., AND CLARKE, V. To Saturate or not to Saturate? Questioning Data Saturation as a Useful Concept for Thematic Analysis and Sample-size Rationales. *Qualitative research in sport, exercise and health* 13, 2 (2021), 201–216.
- [19] BROWN, P., HIDDENLEY, R., GRIFFIN, H., AND ROLLASON, S. The democratic indexing of images. *New Review of Hypermedia and Multimedia* 2, 1 (1996), 107–120.
- [20] BULLARD, J. Curated folksonomies: three implementations of structure through human judgment. *KO Knowledge Organization* 45, 8 (2019), 643–652.
- [21] CHARTERS, E. The use of think-aloud methods in qualitative research an introduction to think-aloud methods. *Brock Education Journal* 12, 2 (2003).
- [22] CHU, C. M. Transformative information services: Uprooting race politics. In *Proceedings of the Black Caucus of the American Library Association Conference, Las Vegas, NV* (1999).
- [23] CHU, C. M., AND O’BIEN, A. Subject analysis: the critical first stage in indexing. *Journal of Information Science* 19, 6 (1993), 439–454.
- [24] CIFOR, M., AND RAWSON, K. Mediating queer and trans pasts: The Homosaurus as queer information activism. *Information, Communication & Society* (2022), 1–18.
- [25] CLEMENTS, L., AND LIEW, C. L. Talking about tags: An exploratory study of librarians’ perception and use of social tagging in a public library. *The Electronic Library* (2016).
- [26] COCQ, C. Indigenous voices on the web: Folksonomies and endangered languages. *Journal of American Folklore* 128, 509 (2015), 273–285.

- [27] CORBIE-SMITH, G., THOMAS, S. B., AND GEORGE, D. M. M. S. Distrust, race, and research. *Archives of internal medicine* 162, 21 (2002), 2458–2463.
- [28] CRIST, E. A., CLARK, K., ET AL. A critical phenomenology of whiteness in academic libraries. *The Journal of Academic Librarianship* 48, 4 (2022), 102557.
- [29] CUTTER, C. A., AND FORD, W. C. *Rules for a dictionary catalog*, vol. 10. US Government Printing Office, 1904.
- [30] DEWEY, J. Propositions, warranted assertibility and truth. In *The Later Works of John Dewey, 1925-1953*, J. Boydson, Ed. Southern Illinois University Press, Carbondale, 2008, pp. 168–188.
- [31] DEWEY, J. Reconstruction in philosophy. In *The Middle Works of John Dewey, 1899-1924*, J. Boydson and G. Murphy, Eds. Southern Illinois University Press, Carbondale, 2008, pp. 77–202.
- [32] DIANGELO, R. *White fragility: Why it's so hard to talk to White people about racism*. Beacon Press Boston, 2018.
- [33] DIGITAL TRANSGENDER ARCHIVE. Homosaurus Vocabulary Site. <https://homosaurus.org/about>, 2022. Accessed: 2022-07-31.
- [34] DOUSA, T. M. Classical pragmatism and its varieties: on a pluriform metatheoretical perspective for knowledge organization. *KO KNOWLEDGE ORGANIZATION* 37, 1 (2010), 65–71.
- [35] DOYLE, A. M., LAWSON, K., AND DUPONT, S. Indigenization of Knowledge Organization at the Xwi7xwa Library.
- [36] DREYFUS, H., DREYFUS, S. E., AND ATHANASIOU, T. *Mind over machine*. Simon and Schuster, 2000.
- [37] D'SOUZA, D. *The end of racism: Principles for a multiracial society*. Free Press New York, 1995.
- [38] DUARTE, M. E., AND BELARDE-LEWIS, M. Imagining: Creating spaces for indigenious ontologies. *Cataloging & Classification Quarterly* 53, 5-6 (2015), 677–702.
- [39] ERICSSON, K. A., AND SIMON, H. A. *Protocol analysis: Verbal reports as data*. the MIT Press, 1984.
- [40] ESPINAL, I. A new vocabulary for inclusive librarianship: Applying whiteness theory to our profession. In *The Power of Language = el poder de la palabra: Selected Papers from the Second REFORMA National Conference*, L. Castillo-Speed, Ed. Libraries Unlimited, 2001, pp. 131–149.
- [41] FARROW, J. F. A cognitive process model of document indexing. *Journal of documentation* (1991).

- [42] FEINBERG, M. An examination of authority in social classification systems. *Advances in Classification Research Online* 17, 1 (2006), 1–11.
- [43] FIDEL, R. Who needs controlled vocabulary? *Special libraries* 83, 1 (1992), 1–9.
- [44] FIDEL, R. User-centered indexing. *Journal of the American Society for Information Science* 45, 8 (1994), 572–576.
- [45] FOSKETT, A. C. Dead heads? *Library Journal* 94, 8 (1969), 1559.
- [46] FOSKETT, A. C. *The subject Approach to Information*. Facet Publishing, 1996.
- [47] FOX, M. J. “Priorities of arrangement” or a “Hierarchy of oppressions?”: Perspectives on intersectionality in knowledge organization. *KO KNOWLEDGE ORGANIZATION* 43, 5 (2016), 373–383.
- [48] FROHMAN, B. Rules of indexing: A critique of mentalism in information retrieval theory. *Journal of Documentation* 46, 2 (1990), 365–386.
- [49] FURNER, J. Dewey deracialized: A critical race-theoretic perspective. *KO Knowledge Organization* 34, 3 (2007), 144–168.
- [50] GALVAN, A. Soliciting performance, hiding bias: Whiteness and librarianship. *the Library with the Lead Pipe* (2015), 135–160.
- [51] GEROLIMOS, M. Tagging for libraries: a review of the effectiveness of tagging systems for library catalogs. *Journal of Library Metadata* 13, 1 (2013), 36–58.
- [52] GLASER, B. G., AND STRAUSS, A. L. *The discovery of grounded theory: Strategies for qualitative research*. Routledge, 2017.
- [53] GOLDBERGER, S., AND HUBERMAN, B. A. The structure of collaborative tagging systems. *arXiv preprint cs/0508082* (2005).
- [54] GOOD, B. M., AND TENNIS, J. T. Term based comparison metrics for controlled and uncontrolled indexing languages. *Information Research: An International Electronic Journal* 14, 1 (2009).
- [55] GRADY, M. P. *Qualitative and action research: A practitioner handbook*. Phi Delta Kappa International, 1998.
- [56] GREENAWAY, S., THELWALL, M., AND DING, Y. Tagging YouTube: A classification of tagging practice on YouTube. In *Proceedings of the 12th International Conference on Scientometrics and Infometrics* (2009), pp. 660–664.
- [57] GROSS, T., TAYLOR, A. G., AND JOUDREY, D. N. Still a lot to lose: the role of controlled vocabulary in keyword searching. *Cataloging & classification quarterly* 53, 1 (2015), 1–39.

- [58] GUERREIRO, N. M., ALVES, D., WALDENDORF, J., HADDOW, B., BIRCH, A., COLOMBO, P., AND MARTINS, A. F. Hallucinations in large multilingual translation models. *arXiv preprint arXiv:2303.16104* (2023).
- [59] GUPTA, M., LI, R., YIN, Z., AND HAN, J. Survey on social tagging techniques. *ACM Sigkdd Explorations Newsletter* 12, 1 (2010), 58–72.
- [60] HAACK, S. Pragmatism. In *The Blackwell companion in philosophy*, N. Bunnin and E. P. Tsui-James, Eds. Malden, MA, 2003, pp. 774–789.
- [61] HALPIN, H., ROBU, V., AND SHEPHERD, H. The complex dynamics of collaborative tagging. In *Proceedings of the 16th International Conference on World Wide Web (WWW '07)* (New York, New York, 2007), ACM, p. 211–220.
- [62] HARPALANI, V. ‘safe Spaces’ and the Educational Benefits of Diversity.
- [63] HARPER, F. M., AND KONSTAN, J. A. The Movielens Datasets: History and Context. *Acm transactions on interactive intelligent systems (tiis)* 5, 4 (2015), 1–19.
- [64] HATHCOCK, A. White librarianship in blackface: Diversity initiatives in lis. *the Library with the Lead Pipe* 7 (2015).
- [65] HIDDENLEY, R., AND RAFFERTY, P. Democratic Indexing: an approach to the retrieval of fiction. *Information Services & Use* 17, 2-3 (1997), 101–109.
- [66] HJØRLAND, B. Towards a theory of aboutness, subject, topicality, theme, domain, field, content... and relevance. *Journal of the American Society for Information Science and Technology* 52, 9 (2001), 774–778.
- [67] HJØRLAND, B. Domain analysis in information science: eleven approaches—traditional as well as innovative. *Journal of documentation* (2002).
- [68] HJØRLAND, B. Domain analysis. *KO Knowledge Organization* 44, 6 (2017), 436–464.
- [69] HOLSTROM, C. Social tagging: Organic and retroactive folksonomies. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries* (2018), pp. 179–182.
- [70] HOLSTROM, C. Local Authorial Voice and Global Authorial Voice in Community-Author Knowledge Organization Systems. *Advances in Classification Research Online* 29, 1 (2019), 6–8.
- [71] HOLSTROM, C. The effects of suggested tags and autocomplete features on social tagging behaviors. *Proceedings of the Association for Information Science and Technology* 57, 1 (2020), e263.
- [72] HOLSTROM, C. The Sears List of Subject Headings: Social and cultural dimensions in historical, theoretical, and design contexts. *NASKO* (2021), 1–11.

- [73] HOLSTROM, C. Analyzing the Structure and Dynamics of Control Relationships in the Case of “Illegal Aliens” in the Library of Congress Subject Headings. In *Advances in Knowledge Organization: Knowledge Organization Across Disciplines, Domains, Services and Technologies. Proceedings of the Seventeenth International ISKO Conference*. (Aalborg, Denmark, July 2022), vol. 19, pp. 133–146.
- [74] HONMA, T. Trippin’over the color line: The invisibility of race in library and information studies. *InterActions: UCLA journal of education and information studies* 1, 2 (2005).
- [75] HOWARD, S. A., AND KNOWLTON, S. A. Browsing through bias: The library of congress classification and subject headings for african american studies and lgbtqia studies. *Library Trends* 67, 1 (2018), 74–88.
- [76] HUTCHINS, W. J. The concept of ‘aboutness’ in subject indexing. In *Aslib proceedings* (1978), vol. 30, MCB UP Ltd, pp. 172–181.
- [77] ISO. Documentation—methods for examining documents, determining their subjects and selecting indexing terms. *International Organization for Standardization, Geneva 5963* (1985).
- [78] JACKSON, C., CROWSTON, K., ØSTERLUND, C., AND HARANDI, M. Folksonomies to support coordination and coordination of folksonomies. *Computer Supported Cooperative Work (CSCW)* 27, 3 (2018), 647–678.
- [79] JOUDREY, D. N. *Building puzzles and growing pearls: A qualitative exploration of determining aboutness*. University of Pittsburgh, 2005.
- [80] KIPP, M. Controlled vocabularies and tags: An analysis of research methods. *NASKO* 3, 1 (2011), 23–32.
- [81] KIPP, M., AND CAMPBELL, D. G. Patterns and inconsistencies in collaborative tagging systems: An examination of tagging practices. *Proceedings of the American Society for Information Science and Technology* 43, 1 (2006), 1–18.
- [82] KROSKI, E. The hive mind: Folksonomies and user-based tagging. *InfoTangle Blog* (2005).
- [83] KUMASI, K. D., JIMES, C., GODWIN, A. E., PETRIDES, L. A., AND KARAGLANI, A. A preliminary study interrogating the cataloging and classification schemes of a k-12 book discovery platform through a critical race theory lens. *Open Information Science* 4, 1 (2020), 106–121.
- [84] LANCASTER, F. W. *Vocabulary control for information retrieval*. Washington, Information Resources Press, 1972.
- [85] LANCASTER, F. W. *Indexing and Abstracting in Theory and Practice, Third edition*. University of Illinois, Champaign, Illinois, 2003.
- [86] LANGRIDGE, D. W. *Subject analysis: principles and procedures*. London; New York: Bowker-Saur, 1989.

- [87] LEE, C.-J. G. Reconsidering constructivism in qualitative research. *Educational Philosophy and theory* 44, 4 (2012), 403–412.
- [88] LIBRARY OF CONGRESS. Subject heading manual: H0180 assigning and constructing subject headings. <https://www.loc.gov/aba/publications/FreeSHM/H0180.pdf>, 2016.
- [89] LITTLETREE, S., AND METOYER, C. A. Knowledge organization from an indigenous perspective: The Mashantucket Pequot Thesaurus of American Indian Terminology project. *Cataloging & Classification Quarterly* 53, 5-6 (2015), 640–657.
- [90] LO, G. “Aliens” vs. catalogers: Bias in the library of congress subject heading. *Legal Reference Services Quarterly* 38, 4 (2019), 170–196.
- [91] LU, K., AND KIPP, M. Understanding the retrieval effectiveness of collaborative tags and author keywords in different retrieval environments: An experimental study on medical collections. *Journal of the Association for Information Science and Technology* 65, 3 (2014), 483–500.
- [92] MAI, J.-E. Deconstructing the indexing process. *Advances in Librarianship* 23 (1999), 269–298.
- [93] MAI, J.-E. *The subject indexing process: an investigation of problems in knowledge representation*. The University of Texas at Austin, 2000.
- [94] MAI, J.-E. Folksonomies and the new order: authority in the digital disorder. *KO Knowledge Organization* 38, 2 (2011), 114–122.
- [95] MARABLE, M. Transforming ethnic studies: Theorizing multiculturalism, diversity, and power. *Souls* 3, 3 (2001), 6–15.
- [96] MASON, M. Sample size and saturation in phd studies using qualitative interviews. *Forum: qualitative social research* 11, 3 (2010).
- [97] MATHES, A. Folksonomies-cooperative classification and communication through shared metadata, 2004.
- [98] MIHAS, P., AND INSTITUTE, O. *Learn to build a codebook for a generic qualitative study*. SAGE Publications, Limited London, UK, 2019.
- [99] MIKSA, F. L. *The subject in the dictionary catalog from Cutter to the present*, 1983.
- [100] MILLS, C. W. *Sociology and Pragmatism: The Higher Learning in America*. Oxford University Press, 1964.
- [101] MINOW, M. EQUALITY VS. EQUITY. *American Journal of Law and Equality* 1 (09 2021), 167–193.
- [102] MORGAN, D. L. Pragmatism as a paradigm for social research. *Qualitative inquiry* 20, 8 (2014), 1045–1053.

- [103] MUNK, T. B., AND MØRK, K. Folksonomies, tagging communities, and tagging strategies: an empirical study. *Knowledge organization* 34, 3 (2007), 115–127.
- [104] NOBLE, S. U. Algorithms of oppression. In *Algorithms of Oppression*. New York University Press, 2018.
- [105] OLSON, H. A. The power to name: Representation in library catalogs. *Signs: journal of women in culture and society* 26, 3 (2001), 639–668.
- [106] OPENAI. ChatGPT. <https://chat.openai.com/chat>. Accessed: 2023-02-15 and 2023-05-08.
- [107] OVERFLOW, S. Tag synonyms. <https://stackoverflow.com/tags/synonyms>. Accessed: 2023-04-30.
- [108] PETERS, I., AND WELLER, K. Tag gardening for folksonomy enrichment and maintenance. *Webology* 5, 3 (2008), 1–18.
- [109] QUIGLEY, K. S., LINDQUIST, K. A., AND BARRETT, L. F. Inducing and measuring emotion and affect: Tips, tricks, and secrets.
- [110] QUNITARELLI, E. Power to the people. In *ISKO Italy-UniMIB meeting, Milan, June 24, 2005* (2005).
- [111] RANGANATHAN, S. *Prolegomena to Library Classification*. Madras Library Association, 1937.
- [112] ROSA, K., AND HENKE, K. 2017 ALA Demographic Study. Tech. rep., American Library Association, 2017.
- [113] ROWLEY, J. The controlled versus natural indexing languages debate revisited: a perspective on information retrieval practice and research. *Journal of information science* 20, 2 (1994), 108–118.
- [114] SANDER, R., AND TAYLOR JR, S. *Mismatch: How Affirmative Action Hurts Students It S Intended to Help, and Why Universities Won't Admit It*. Basic Books, 2012.
- [115] SAUNDERS, B., SIM, J., KINGSTONE, T., BAKER, S., WATERFIELD, J., BARTLAM, B., BURROUGHS, H., AND JINKS, C. Saturation in qualitative research: exploring its conceptualization and operationalization. *Quality & quantity* 52, 4 (2018), 1893–1907.
- [116] ŠAUPERL, A. *Subject determination during the cataloging process*. The University of North Carolina at Chapel Hill, 1999.
- [117] ŠAUPERL, A. Subject cataloging process of slovenian and american catalogers. *Journal of documentation* (2005).
- [118] SEIGFRIED, C. H. *Feminist Interpretations of John Dewey*. Penn State Press, 2010.

- [119] SHIRKY, C. Ontology is overrated: Categories, links, and tags. <http://www.shirky.com/writings/ontology-overrated.htm> (2005).
- [120] SILVA LUNA, D., AND BERING, J. M. Varieties of awe in science communication: Reflexive thematic analysis of practitioners' experiences and uses of this emotion. *Science Communication* (2022), 10755470221098100.
- [121] SMIRAGLIA, R. P. Subject access to archival materials using lcsb. In *Describing Archival Materials*. Routledge, 2014, pp. 63–90.
- [122] SNOW, K., AND DUNBAR, A. W. Advancing the relationship between critical cataloging and critical race theory. *Cataloging & Classification Quarterly* 60, 6-7 (2022), 646–674.
- [123] SOERGEL, D. *Organizing information: Principles of data base and retrieval systems*. Elsevier, 1985.
- [124] SOERGEL, D. Indexing and retrieval performance: The logical evidence. *Journal of the American Society for Information Science* 45, 8 (1994), 589–599.
- [125] STRÜBING, J. Research as pragmatic problem-solving: The pragmatist roots of empirically-grounded theorizing. *The Sage handbook of grounded theory* (2007), 580–602.
- [126] SVENONIUS, E. *The intellectual foundation of information organization*. MIT press, 2000.
- [127] TENNIS, J. T. *Conceptions of subject analysis: a metatheoretical investigation*. University of Washington, 2005.
- [128] TENNIS, J. T. Social tagging and the next steps for indexing. *Advances in classification research* 17 (2006).
- [129] TENNIS, J. T. 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, 7 (2012), 1350–1359.
- [130] THE UNIVERSITY OF BRITISH COLUMBIA. Indigenous Knowledge Organization. <https://xwi7xwa.library.ubc.ca/collections/indigenous-knowledge-organization/>, 2023. Accessed: 2023-04-30.
- [131] UNITED STATES NATIONAL COMMISSION FOR THE PROTECTION OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH. *The Belmont report: ethical principles and guidelines for the protection of human subjects of research*, vol. 2. Department of Health, Education, and Welfare, National Commission for the . . . , 1978.
- [132] UNIVERSITY OF WASHINGTON INFORMATION SCHOOL. MLIS Statistics — Information School — University of Washington. <https://ischool.uw.edu/programs/mlis/statistics/>, 2022. Accessed: 2022-07-15.

- [133] URQUHART, C. *Grounded theory for qualitative research: A practical guide*. Sage, 2012.
- [134] WARNER, J. N. Moving beyond whiteness in north american academic libraries.
- [135] WHIPPS, J., AND LAKE, D. Pragmatist Feminism. In *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed., Winter 2020 ed. Metaphysics Research Lab, Stanford University, 2020.
- [136] WILKERSON, I. *Caste: The origins of our discontents*. Random House, 2020.
- [137] WILSON, H. W. *Sears List of Subject Headings*, 22 ed. H.W. Wilson Company, 2018.
- [138] WILSON, P., AND KOEPP, D. W. *Two Kinds of Power: An Essay on Bibliographical Control*. Univ of California Press, 1968.
- [139] Y2KARL. Can science fiction wake us up to our climate reality? <https://www.metafilter.com/194120/Can-Science-Fiction-Wake-Us-Up-to-Our-Climate-Reality>. Accessed: 2022-05-6.

Appendix A
RECRUITING EMAILS

I sent two recruiting emails to the MLIS student mailing list at the University of Washington. First I sent the following initial recruiting email:

Subject: Participate in a Research Study on Subject Indexing, Tagging, and Vocabulary Control

Hi UW MLIS Students!

Are you interested in subject indexing? Social tagging? Vocabulary control? Do you care about how all of these topics are related to racial and social justice? Do you want to learn more about these topics and get a \$20 gift card?

I am a PhD student in the iSchool, and I am conducting a research study to understand the effect of different types of vocabulary control on the subject indexing and subject tagging processes. I want you to participate in the study! You can have any level of experience with subject indexing, tagging, and vocabulary control to participate. Study sessions will take place in-person on the Seattle campus from January 9 to February 28. Study sessions will last about an hour and a half.

If you're interested, please sign up for a spot on the Subject Indexing and Tagging Scheduler. After you sign up, I'll send you a survey and consent form and will confirm the time of your session. If you can't find a time that works for you or if you have any questions, please email me to get scheduled or to get answers.

Thanks!

Chris Holstrom

iSchool PhD Candidate

cholstro@uw.edu
chrisholstrom.com

A month later, I sent a follow-up recruiting email to the same UW MLIS mailing list:

Subject: You Still Have a Chance to Participate in a Research Study on Subject Indexing, Tagging, and Vocabulary Control

Hi UW MLIS Students!

At the beginning of the quarter, I sent you a recruiting email for a research study about subject indexing, social tagging, vocabulary control, and racial and social justice. Many of you have already participated, which has been great! I am writing again because I have a handful of available time slots remaining in February, and I'd love for you to participate if you're available.

Here are some questions and answers about the study:

What is it?

A research study conducted by an iSchool PhD student where you'll tag and index four books and answer interview questions about your experience.

Who can sign up?

Anyone in the UW MLIS program. You do not need any experience with subject indexing or tagging.

What's in it for me?

You'll get some hands-on experience with subject analysis, subject indexing, vocabulary control, and more. You'll also get a \$20 gift card to the University Book Store. Finally, the participants who have done the study so far have described it as a good experience and as fun.

How long does it take?

About an hour and a half.

Where is it?

Research sessions take place in-person in Mary Gates Hall.

When are the sessions?

I have open slots on February 7, 13, 14, 17, and 27.

How do I sign up?

If you're interested, please sign up for a time slot on the Subject Indexing and Tagging Scheduler. After you sign up, I'll send you a survey and consent form and will confirm the time of your session. If you can't find a time that works for you or if you have any questions, please email me to get scheduled or to get answers.

Thanks!

Chris Holstrom

iSchool PhD Candidate

cholstro@uw.edu

chrisholstrom.com

Appendix B

DEMOGRAPHIC SURVEY

All participants in the study filled out the following demographic survey before arriving at their study session—or, in a few cases, at the beginning of their session. I provided each participant with a Participant ID, so that their name was not connected with their responses, and assured participants that their demographic data will be kept private and will not be associated with their name.

The questions in the demographic survey are listed below, and the actual survey is available via Google Forms at the following URL: <https://forms.gle/JfN3n8nVhRAx97cy8>.

Survey:

1. Age:

- 20 and younger
- 21-24
- 25-29
- 30-34
- 35-39
- 40-44
- 45-49
- 50-54
- 55-59
- 60 and older

2. Gender: Enter any gender or genders that you identify with.

3. **Race and Ethnicity:** Which of the following categories best describes you? Check any and all categories that you identify with and feel free to enter multiple responses in the Other option.

- American Indian or Alaska Native
- Asian
- Black or African American
- Middle Eastern or North Africa
- Native Hawaiian or Pacific Islander
- White
- Other ...

4. **Education:** List any post-secondary degrees or certificates that you have completed and any academic programs in which you are currently enrolled.

5. **LIS Courses:** Check any LIS courses that you have completed or are currently taking. The courses listed below are UW courses. If you took an equivalent course at another institution, check the UW course here. Use “Other” for courses about subject indexing, tagging, or vocabulary control that are not listed.

- LIS 501: History and Foundations of Libraries and Librarianship
- LIS 510: Information Behavior
- LIS 512: Community Analysis
- LIS 520: Concepts, Services, and Issues for Information Professionals
- LIS 521: Principles of Information Services
- LIS 522: Collection Development
- LIS 530: Organization of Information and Resources
- LIS 531: Catalogs, Cataloging, and Classification
- LIS 534: Indigenous Systems of Knowledge
- LIS 537: Construction of Indexing Languages
- LIS 539: Metadata Design
- LIS 544: Information Retrieval Systems
- LIS 553: Information and Social Justice

- Other ...

6. **Indexing, Tagging, and Cataloging Experience:** Indicate your years of experience and briefly describe your experience with indexing, tagging, and cataloging, including work experience and any academic experience beyond small class activities or homework. If you have previously worked with the *LCSH* or the *Sears List*, make sure to note that experience.

7. **Authors and Books** Indicate which of the following authors and books you have read, have heard of, or have not heard of. *[Note: The Google Form will show these books to participants in random order.]*

- The Fire Next Time - James Baldwin
- The Emperor of All Maladies - Siddhartha Mukherjee
- The Sympathizer - Viet Thanh Nguyen
- White Fragility - Robin DiAngelo
- The New Jim Crow - Michelle Alexander
- Guns, Germs, and Steel - Jared Diamond
- So You Want to Talk About Race - Ijeoma Oluo
- The Invention of Nature - Andrea Wulf
- How to Be An Antiracist - Ibram X. Kendi
- Between the World and Me - Ta-Nehisi Coates
- Piranesi - Susanna Clarke
- Just Mercy - Bryan Stevenson
- Caste: The Origins of Our Discontents - Isabel Wilkerson
- The Handmaid's Tale - Margaret Atwood
- March - John Lewis, Andrew Aydin, and Nate Powell
- Where We Stand - bell hooks
- The Nickel Boys - Colson Whitehead
- The Immortal Life of Henrietta Lacks - Rebecca Skloot
- Mismatch - Richard Sander and Stuart Taylor, Jr.
- The End of Racism - Dinesh D'Souza

Appendix C

SCRIPT TO INTRODUCE THE STUDY TO PARTICIPANTS

When participants arrived, I greeted them and introduced the study. The goals of the introduction were to give the participants a high-level overview of what they would be doing, to explain the goals of the study, and to make the participants feel comfortable.

Script:

Welcome! I'm really glad you're here. My name is Chris and I'll be the researcher working with you today. Your study session has three parts:

1. You've already filled out a short demographic survey.
2. Next, you'll index or tag four books. During that part of the study I'll ask you to "think aloud"—to verbalize your thoughts.
3. Third, I'll ask you some questions about your experience.

The whole session should take about an hour and a half. You can ask to take a break or opt out at any time.

You might be wondering what this study is about beyond the short email blurb about subject indexing, tagging, vocabulary control, and racial and social justice. I'll expand on that a little bit. The study aims to understand how different forms of vocabulary control affect the subject indexing or tagging process. Tagging and subject indexing both involve assigning keywords or phrases that describe a document. You'll index or tag four books, each with a different form of vocabulary control:

- Free tagging
- Tagging with autocomplete suggestions
- Subject indexing with the *Sears List of Subject Headings*

- Subject indexing with the *Library of Congress Subject Headings*

You'll know which form of vocabulary control you are using, but they won't necessarily come in the order that I just said them. While you index and tag, you'll "think aloud" so I can know what you are thinking or wondering. After you tag and index all four books, I'll ask you some interview questions to understand even more about what you're thinking and feeling as you do this work.

It is important for you to know that I'm not testing you, I'm testing the different forms of vocabulary control. If you ever get stuck or frustrated, you're not "doing it wrong." You're actually providing me with really interesting data, and the more that you verbalize your thoughts in the think-aloud protocol and the interview, the better.

It is also important to know that the books that I will ask you to index and tag are about topics related to race and racism. The subject matter of the books might make participating in this study stressful for you. If you ever want to skip a book or stop participating in the study, just say so. If sharing your emotions helps you work through them, please know that everything that you say and that happens in your study session will be anonymized. I want this to be a safe space for you, so please let me know anything that I can do to help you feel safe.

Do you have any questions? *[Answer questions.]*

Okay, let's get started. Here is a gift card and a consent form. The gift card is for you to keep for showing up today, and you get to keep it even if you opt out of parts or all of the study. The consent form is the same one that I sent you by email to review. If you're comfortable with what it says, go ahead and sign it. I can give you a copy for your records.

[Pause for participant to fill out consent form.]

Thank you for filling that out. Now we can get started with subject indexing and tagging!

You'll be tagging or indexing four books, one at a time. For each book, I'll first ask you to do some subject analysis work and come up with an aboutness statement that describes the book. I'll ask you to write down your aboutness statement on this notepad. The aboutness

statement can be a sentence, a paragraph, or a list of words or phrases. Whatever works best for you to describe what the book is about. When you feel reasonably comfortable with your aboutness statement, I'll introduce the form of vocabulary control that you'll be using and give you instructions. Okay, here's your first book. Remember to verbalize your thoughts—think aloud!—and let me know if you have any questions. And remember that you can opt out of any part of the study or the study as a whole at any time. Okay, go for it!

[Remember to start audio recording.]

Appendix D
CONSENT FORM

Each participant read and signed the following consent form before beginning the study. The consent form that participants were given was formatted to fit on one page.

Form:

Consent for Participation: Subject Indexing, Tagging, and Controlled Vocabularies Study

I volunteer to participate in this research study conducted by Chris Holstrom (cholstro@uw.edu), PhD candidate from the University of Washington Information School. I understand that the study is designed to gather information about the subject indexing process, tagging, and the effects of different forms of vocabulary control. I also understand that the study will investigate racial and social justice and will involve analyzing and working with books that cover these difficult topics. I will be one of approximately 20 participants in this research.

1. My participation in this project is voluntary. I understand that I will not be paid for my participation beyond a \$20 gift card to the University Book Store. I may withdraw and discontinue participation at any time without penalty, and I will receive the gift card regardless of whether I complete the study. If I decline to participate or withdraw from the study, no one on my campus will be told.
2. I understand that most participants will find the study activities and the interview that follows to be interesting and thought-provoking. If, however, I feel uncomfortable in any way during the study session, I have the right to decline to answer any question, to skip any portion of the study, or to end the study.

3. Participation in the study involves three parts: 1) a demographic survey, 2) a series of four subject indexing and tagging activities where I will be asked to “think aloud” or verbalize my thoughts, and 3) an interview that will ask me to reflect on the activities in the study. The study session will last approximately 90 minutes. Notes will be written and audio will be recorded during the study session. If I do not want to be audio recorded, I will not be able to participate in the study.
4. I understand that the researcher will not identify me by name in any reports using information obtained from this interview, and that my confidentiality as a participant in this study will remain secure. Subsequent uses of records and data will be subject to standard data use policies which protect the anonymity of individuals and institutions.
5. I understand that I will not be judged based on my participation in this study or based on my responses to the activities or the interview in the study. My participation in this study will not affect my grades in UW MLIS courses or my standing in the program.
6. I understand that this research study has been reviewed and approved by the Institutional Review Board (IRB) in the Human Subjects Division of the Office of Research at the University of Washington. To report research problems or ask questions regarding the treatment of participants in this study, the IRB may be contacted through the contact information listed here: <https://www.washington.edu/research/hsd/hsd-contact/>.
7. I have read and understood the explanation provided to me. I have had all my questions answered to my satisfaction, and I voluntarily agree to participate in this study.
8. I have been given a copy of this consent form.

My Signature

Date

My Print Name

Researcher's Signature

Appendix E

SCRIPT TO INTRODUCE MINIMAL CONTROL (FREE TAGGING)

I used the following script to help guide the participant from creating an aboutness statement to tagging the book in a free tagging (or minimal vocabulary control) user interface.

Script:

Thank you for that aboutness statement!

Next, we're going to move to the second stage for this book, translating your aboutness statement into indexing terms. For this book, I'm going to ask you to use this user interface to enter a tag or tags that describe what the book is about. *[Show the UI on the computer.]* You can enter any tags that you want and as many or as few tags as you like. Just type them into this field and click **Add Tag** or the **Enter** key. When you have entered all of the tags that you want to enter, click **Submit**. Remember that there are not right and wrong answers or a right or wrong way to do this. I am interested in observing your process and hearing your thoughts as you work. Remember to keep thinking out loud.

Let me know if you have any questions and feel free to get started.

Appendix F

SCRIPT TO INTRODUCE PASSIVE CONTROL (AUTOCOMPLETE TAGS ENABLED)

I used the following script to help guide the participant from creating an aboutness statement to tagging the book in a tagging user interface with autocomplete enabled. The autocomplete feature passively suggests tags from a controlled vocabulary.

Script:

Thank you for that aboutness statement!

Next, we're going to move to the second stage for this book, translating your aboutness statement into indexing terms. For this book, I'm going to ask you to use this user interface to enter a tag or tags that describe what the book is about. *[Show the UI on the computer.]* You can enter any tags that you want and as many or as few tags as you like. Just type them into this field and click **Add Tag** or the **Enter** key. You'll notice that as you type, the user interface suggests tags for you from a controlled vocabulary. You can use this autocomplete feature if you like. When you have entered all of the tags that you want to enter, click **Submit**. Remember that there are not right and wrong answers or a right or wrong way to do this. I am interested in observing your process and hearing your thoughts as you work. Remember to keep thinking out loud.

Appendix G

SCRIPT TO INTRODUCE FLEXIBLE CONTROL (*SEARS LIST*)

I used the following script to help guide the participant from creating an aboutness statement to choosing subject headings from the *Sears List of Subject Headings*.

Script:

Thank you for that aboutness statement!

Next, we're going to move to the second stage for this book, translating your aboutness statement into indexing terms. For this book, I'm going to ask you to use the *Sears List of Subject Headings* to describe what the book is about. [*Hand copy of the Sears List to the participant.*] The *Sears List* provides you with a list of authorized subject headings and with instructions for creating your own subject headings. You can also create Pre-coordinated subject headings by combining multiple terms to make a more specific heading, like "Music—History and criticism."

You can choose or create as many subject headings as you like. Just type them into this field and click **Add Subject Heading** or the **Enter** key. When you have entered all of the subject headings that you want to enter, click **Submit**. Remember that there are not right and wrong answers or a right or wrong way to do this. I am interested in observing your process and hearing your thoughts as you work. Remember to keep thinking out loud.

Appendix H

SCRIPT TO INTRODUCE RIGID CONTROL (*LCSH*)

I used the following script to help guide the participant from creating an aboutness statement to choosing subject headings from the *LCSH*.

Script:

Thank you for that aboutness statement!

Next, we're going to move to the second stage for this book, translating your aboutness statement into indexing terms. For this book, I'm going to ask you to use the *Library of Congress Subject Headings* to describe what the book is about. [*Show the participant the Library of Congress Authorities at <https://authorities.loc.gov/> and show them how to search the Subject Authority Headings.*] The *LCSH* provides you with a list of authorized subject headings and with instructions for creating pre-coordinated headings. Pre-coordinated subject headings combine multiple terms to make a more specific heading, like "Music—History and criticism."

You can choose as many subject headings as you like; however, all of your subject headings—whether they use subdivisions or not—must be authorized headings. You cannot create your own subject headings. For each subject heading that you choose, just type them into this field and click **Add Subject Heading** or the **Enter** key. When you have entered all of the subject headings that you want to enter, click **Submit**. Remember that there are not right and wrong answers or a right or wrong way to do this. I am interested in observing your process and hearing your thoughts as you work. Remember to keep thinking out loud.

Appendix I

**SCRIPT FOR WHEN PARTICIPANTS FINISHED INDEXING OR
TAGGING A BOOK**

I used this script each time that a participant finishes tagging or indexing a book and is about to start working on another book.

Script:

Awesome! Thank you! Now we're going to repeat that process with a different book and a different form of vocabulary control. Like before, I'll ask you to first do some subject analysis work to write an aboutness statement that describes the book. The aboutness statement can be a sentence, a paragraph, or a list of words or phrases. Whatever works best for you. When you feel reasonably comfortable with your aboutness statement, I'll introduce the form of vocabulary control that you'll be using and give you instructions. Okay, here's your next book. Remember to verbalize your thoughts—think aloud!—and let me know if you have any questions. Go for it!

Appendix J

INTERVIEW QUESTIONS

After they completed all four subject indexing and subject tagging activities, I interviewed the participants. The goals of the interview were to further understand the participants' responses to the subject indexing and subject tagging activities, to facilitate direct comparisons of the different forms of vocabulary control, to better understand participants' perceptions of control, and to follow up on any interesting events or statements made during the subject indexing and subject tagging activities.

The questions listed here represent the core or base interview questions. All participants were asked these questions, and many participants were asked follow-up questions specific to their answers or to their subject indexing and subject tagging activities.

Core Interview Questions:

1. Which book did you find easiest to index or tag? Why?
2. Which book did you find most difficult to index or tag? Why?

[Provide a short explanation of the concepts of faithfulness of representation, exhaustivity, and specificity in the context of subject indexing and tagging.]

3. For which book did your subject headings or tags most faithfully—most accurately—represent the contents of the book? Why?
4. For which book did your subject headings or tags least faithfully—least accurately—represent the contents of the book? Why?
5. For which book did your subject headings or tags most exhaustively—most completely—represent the contents of the book? Why?

6. For which book did your subject headings or tags least exhaustively—least completely—represent the contents of the book? Why?
7. Did any of the subject headings or tags that you used seem too specific? Why?
8. Did any of the subject headings or tags that you used seem not specific enough? Why?
9. 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?
 - (a) What are some examples that you think were particularly close matches?
 - (b) Examples that were particularly different?
 - (c) What influenced your choices of tags or subject headings in relation to your aboutness statement?

Note: Questions 10, 11, and 12 rotated positions to avoid ordering effects.

10. During this study you experienced four different forms of vocabulary control: 1) free tagging, 2) tagging with autocomplete, 3) indexing with the *Sears List*, and 4) indexing with *LCSH*. Can you put these forms of vocabulary control in order from the one where you felt most in control to where you felt most controlled?
 - (a) What did “in control” mean to you in this context?
 - (b) What did “controlled” mean to you in this context?
 - (c) Why was *[fill-in-the-blank]* the one where you felt most in control?
 - (d) Why was *[fill-in-the-blank]* the one where you felt most controlled?

[Make sure to pick up cards to avoid anchoring effect with similar questions to follow.]

11. Now, can you put the four different forms of vocabulary control in order from easiest to use to most difficult to use?
 - (a) What did “easy to use” mean to you in this context?

- (b) What did “difficult to use” mean to you in this context?
- (c) Why was *[fill-in-the-blank]* the easiest to use?
- (d) Why was *[fill-in-the-blank]* the most difficult to use?

[Make sure to pick up cards to avoid anchoring effect with similar questions to follow.]

12. Now, can you put the four different forms of vocabulary control in order from the one where you felt most in control to where you felt most out of control?

- (a) What did “in control” mean to you in this context?
- (b) What did “out of control” mean to you in this context?
- (c) Why was *[fill-in-the-blank]* the one where you felt most in control?
- (d) Why was *[fill-in-the-blank]* the one where you felt most out of control?

13. Some tagging systems allow editors to change tags that you have entered after the fact, so that they can standardize language. For example, an editor might change your tag “Spectacles” to “Eyeglasses” because they settled on “Eyeglasses” as the standard term? How might you react if you knew editors might change your tags like that?

- (a) Would you feel in control, out of control, or controlled? Why?
- (b) Would you be more or less likely to reconsider your tags before submitting them, knowing that an editor might change them? Why?

14. Now, I am going to ask you to revisit the three sets of rankings that you just did with the note cards, but this time I want you to add “Tagging with Editors” (what we just discussed where editors can change tags after they are submitted) as a 5th type of vocabulary control.

[Note: For each of the following sub-questions, present the note cards back to the participant in the order that they had for each ranking.]

- (a) If you added this type of tagging to your ranking of most-in-control-to-most-controlled, where would you add it?

- (b) If you added this type of tagging to your ranking of most-in-control-to-most-out-of-control, where would you add it?
 - (c) If you added this type of tagging to your ranking of easiest-to-use-to-most-difficult-to-use, where would you add it?
15. Do you want to share anything else about this experience or how you felt as you tagged and indexed these books?

Appendix K

CODEBOOK

The codebook for this study was used for both the think-aloud protocol and the interview, and comprised three parts: 1) Participant responses to interview questions about coextensiveness 2) Bibliographic apparatus that participants used to find aboutness information, and 3) Reactions, responses, and behaviors of participants. The third part of the codebook grew organically from a seed list during the study and features a wide-ranging list of codes related to participants' feelings and reactions observed during the think-aloud protocol and the retrospective interview, with a particular focus on reactions related to racial and social justice.

K.1 Codebook for Participant Responses to Interview Questions About Coextensiveness

During the retrospective interview, participants were asked a series of questions about the coextensiveness of their aboutness statements and their indexing terms. The codes for thematic analysis of their responses are grouped and listed below.

Self-reported degree of coextensiveness:

- *Strong coextensiveness*: Use when the participant indicates that they felt their aboutness statement and indexing terms were highly coextensive.
- *Somewhat strong coextensiveness*: Use when the participant indicates that they felt their aboutness statement and indexing terms were coextensive. To differentiate from “strong coextensiveness,” look for moderating words like “pretty” or “fairly.”
- *Somewhat weak coextensiveness*: Use when the participant indicates that they felt their aboutness statement and indexing terms were not particularly coextensive but had some relationship to each other.

- *Weak coextensiveness*: Use when the participant indicates that they felt their aboutness statement and indexing terms were not at all coextensive.

Examples of strong and weak coextensiveness:

- *Close match example*: Use for specific examples of strong coextensiveness between aboutness statement and indexing terms.
- *Poor match example*: Use for specific examples of weak coextensiveness between aboutness statement and indexing terms.

Participant strategies and behaviors:

- *Aboutness different from representation*: Use when participants indicated that they had different goals with aboutness statements and indexing terms or that they saw aboutness statements and indexing terms as different genres of writing.
- *Direct use of aboutness statement*: Use when the participant indicated that they systematically worked through the keywords in their aboutness statement, like a list of candidate indexing terms.
- *Bibliographic apparatus, not using aboutness statement*: Use when the participant indicated that they referred to portions of the books that they were indexing or tagging during the subject representation step instead of using their aboutness statement.

Influences on coextensiveness:

- *Influence of vocabulary control*: Use for general instances of participants citing the form of vocabulary as having an effect on translation or subject representation.
- *Vocabulary control effect on term choice*: Use for specific examples of the form of vocabulary control influencing a transformation, including suppression.
- *Influence of lack of domain knowledge*: Use when the participant indicated that they felt that their lack of domain knowledge affected the coextensiveness of their aboutness statement and their indexing terms.

- *Influence Personal bias*: Use when the participant indicates that their personal bias (or their attempt to minimize their personal bias) influenced the coextensiveness of their aboutness statement and indexing terms.
- *Influence of order (learning effects)*: Use when the participant indicated that the order of tasks influenced coextensiveness. For example, if the participant learned about the process or a specific indexing term from previous tasks.
- *Influence of user focus*: Used when the participant cited anticipation of index users' needs or search strategies as an influence on coextensiveness.

K.2 Codebook for Bibliographic Apparatus that Participants Used to Find Aboutness Information

The codebook for bibliographic apparatus was adapted from *ISO Standard 5963: Documentation—Methods for examining documents, determining their subjects, and selecting indexing terms* [77], with input from my experience teaching, observing, and practicing indexing and tagging, and from observed behavior during the study. The codebook is organized by each code's location in the books that were used in the study. Codes such as "Front matter (general)" were used when participants looked quickly at that part of the book without focusing on a specific part indicated in one of the other codes.

Book cover and dust jacket:

- *Front cover*
- *Front flap summary*
- *Back cover summary*
- *Praise (back cover or front matter)*
- *Biographical info about author*
- *BISAC subject codes on back cover*
- *Back cover (general)*

Front matter:

- *Title page*

- *Copyright page and copyright date*
- *Suggested LCSH subject headings*
- *Preface or foreword*
- *Author's note*
- *Acknowledgments*
- *Table of contents or chapter headings*
- *List of figures*
- *Quotations*
- *Front matter (general)*

Body of the book:

- *Introduction or opening chapters*
- *Chapters in body of book*
- *Section headings*
- *Illustrations or figures*
- *Tables and captions*
- *Underlined or emphasized words or phrases*
- *Body of the book (general)*

Back matter:

- *Conclusion, epilogue, or final chapter*
- *Index at back of book*
- *Glossary of terms*
- *Citations, bibliography, or notes*
- *Recommended resources and further reading*

K.3 Codebook for Participant Reactions, Responses, and Behaviors

The codebook for participant reactions, responses, and behaviors was seeded with an initial list of codes and was updated as part of thematic analysis conducted throughout the study. This codebook is wide-ranging and covers many types of topics and themes, including participant reactions to forms of vocabulary control and material and apparatus; social and racial justice; and concepts in indexing and tagging like accuracy, exhaustivity, and specificity. Many codes come in pairs that indicate opposites. For example, “Expressing familiarity with the form of vocabulary control” and “Expressing lack of familiarity with the form of vocabulary control.”

Subject analysis:

- *Feeling uncertain about the subject(s) of the book*
- *Feeling clarity about the subject(s) of the book*
- *Feeling that they failed to express the aboutness of the book*
- *Feeling that they expressed the aboutness of the book*
- *Questioning or revising wording of the aboutness statement*
- *Writing aboutness statement in list form*
- *Writing aboutness statement in paragraph form*

Vocabulary Control and Indexing Terms:

- *Feeling confusion with rules of the indexing language*
- *Feeling confident with rules of the indexing language*
- *Frustration over vocabulary control limiting choices*
- *Satisfaction with vocabulary control providing clarity*
- *Feeling uncertain about an indexing term choice*
- *Feeling clarity about an indexing term choice*
- *Feeling intimidated by the form of vocabulary control*
- *Feeling confused by the form of vocabulary control*
- *Feeling supported by the form of vocabulary control*

- *Expressing familiarity with the form of vocabulary control*
- *Expressing lack of familiarity with the form of vocabulary control*
- *Browsing the vocabulary for additional ideas*
- *Not able to think of a term*
- *Aiming for exhaustivity*
- *Aiming for specificity*
- *Aiming for accuracy*
- *User-focused indexing, thinking of expected user search behaviors*
- *Document-focused indexing, trying to honor the perspective or vocabulary in the book*

Materials and Apparatus:

- *Feeling confusion with the user interface*
- *Feeling satisfaction with the user interface*
- *Participant error with the user interface*
- *Feeling intimidated by materials or apparatus*
- *Feeling supported by materials or apparatus*

Feelings of Control and Completion:

- *Feeling controlled*
- *Feeling in control*
- *Feeling out of control*
- *“Satisficing”*
- *Giving up (Finishing task without feeling completeness)*
- *Feeling of satisfaction with task completion*
- *Happiness with the subject indexing or tagging process*
- *Frustration with the subject indexing or tagging process*
- *Feeling pressed for time, hurried*
- *Circumventing controls*
- *Choosing not to circumvent controls*
- *In control meaning freedom to do what the participant wanted to do*

- *In control meaning that the vocabulary control supported the participant and helped them find additional indexing terms*
- *In control meaning that that participant felt they do a good job indexing or tagging*
- *In control meaning that the indexing term directly reflected the participant's thoughts*
- *Controlled meaning constrained or restricted*
- *Controlled meaning that another agent has control*
- *Controlled meaning the participant felt controlled by rules and processes*
- *Controlled meaning that the vocabulary did not have enough terms or did not have the right terms*
- *Out of control meaning that the vocabulary and its rules are restrictive*
- *Out of control meaning that the form of vocabulary control lacks guard rails*
- *Easy to use meaning vocabulary matches or allows participant's terminology*
- *Easy to use meaning easy to understand, lower effort required*
- *Easy to use meaning the form of vocabulary control provided useful feedback or guidance*
- *Difficult meaning complex or overwhelming*
- *Difficult meaning constrained or restricted*
- *Difficult meaning the form of vocabulary control provided insufficient feedback or guidance*

Racial and Social Justice:

- *Feeling negative about the content or perspective of the book*
- *Feeling positive about or empathizing with the content or perspective of the book*
- *Wanting to promote a book that they agree with or consider important*
- *Wanting to "hide" a book that they view as problematic by under-indexing or other strategy*
- *Aiming to not reveal personal racial or social bias in indexing terms*
- *Feeling disconnect between author's use of racial or social term and perceived meaning of related indexing term*
- *Feeling alignment of author's use of racial or social term and perceived meaning of related indexing term*

- *Feeling personal identity (including racial identity) controlled*
- *Feeling empowered to express personal identity (including racial identity)*
- *Feeling unqualified or unsure about how to represent a difficult topic (like racial or social justice)*
- *Feeling reassured about how to represent a difficult topic (like racial or social justice)*
- *Finding an authorized or suggested indexing term (related to racial or social justice) problematic*
- *Feeling responsibility to represent racial and social issues equitably*

Appendix L

CHATGPT RESPONSES TO SUBJECT INDEXING AND SUBJECT TAGGING PROMPTS

The following screenshots represent a selection of ChatGPT's response to prompts based on the tasks that participants completed in the study.

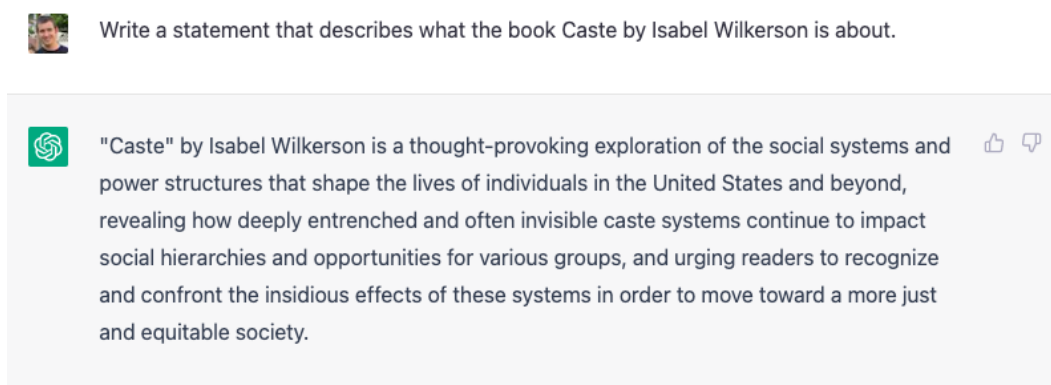




Figure L.1: ChatGPT response to the prompt "Write a statement that described what the book *Caste* by Isabel Wilkerson is about." The response was accurate but read more like a high-level summary than an aboutness statement.



What Library of Congress Subject Headings would you use to index the book *Caste* by Isabel Wilkerson?



Here are some possible Library of Congress Subject Headings (LCSH) that could be used to index the book "Caste" by Isabel Wilkerson:  

- Caste -- United States
- Social stratification -- United States
- Race relations -- United States
- Racism -- United States
- Social classes -- United States
- Social structure -- United States
- Social status -- United States
- Discrimination -- United States
- Group identity -- United States
- Prejudices -- United States
- Equality -- United States
- Social justice -- United States

Please note that this is a non-exhaustive list and other LCSH terms may be used as well.

Figure L.2: ChatGPT response to the prompt “What Library of Congress Subject Headings would you use to index the book *Caste* by Isabel Wilkerson?” The response seemed reasonable and authorized headings but was possibly overenthusiastic about geographical subdivision.



What Library of Congress Subject Headings would you use to index the book *Caste* by Isabel Wilkerson?



Here are some possible Library of Congress Subject Headings (LCSH) that could be used to index the book "*Caste*" by Isabel Wilkerson: 👍 🗨️

- Caste -- United States
- Social stratification -- United States
- Race relations -- United States
- Racism -- United States
- Social classes -- United States
- Social structure -- United States
- Social status -- United States
- Discrimination -- United States
- Group identity -- United States
- Prejudices -- United States
- Equality -- United States
- Social justice -- United States

Please note that this is a non-exhaustive list and other LCSH terms may be used as well.

Figure L.3: ChatGPT response to the prompt “What Library of Congress Subject Headings would you use to index the book *Caste* by Isabel Wilkerson?” The response seemed reasonable and used authorized headings but was possibly overenthusiastic about geographical subdivision.



What Sears List of Subject Headings headings would you use to index the book *White Fragility* by Robin DiAngelo?



Here are some possible Sears List of Subject Headings (SLSH) that could be used to index the book "*White Fragility*" by Robin DiAngelo:



- Racism -- United States
- Antiracism -- United States
- White supremacy movements -- United States
- Whites -- Race identity
- Prejudices
- Discrimination
- Multiculturalism
- Social justice
- Equality
- Social awareness
- Social problems
- United States -- Race relations
- Race discrimination -- United States
- Racial awareness

Please note that this is a non-exhaustive list and other SLSH terms may be used as well.

Figure L.4: ChatGPT response to the prompt “What Sears List of Subject Headings headings would you use to index the book *White Fragility* by Robin DiAngelo?” The response seemed reasonable, used authorized headings, and was fairly exhaustive. Notably, though, the response used two main headings that study participants flagged as inaccurate (“White supremacy movements”) or problematic (“Whites”).



What keywords would you use to describe what the book *Mismatch* by Richard Sander and Stuart Taylor Jr. is about?




"Mismatch" by Richard Sander and Stuart Taylor Jr. is a book about the effects of affirmative action policies in higher education and the unintended consequences that can arise from them. Here are some possible keywords that could describe the book:






- Affirmative action
- Higher education
- Admissions policies
- Diversity
- Racial preferences
- College success
- Law schools
- STEM fields
- Meritocracy
- Standardized testing
- Academic achievement
- Equal opportunity
- Social justice
- Race relations
- Educational inequality

Please note that these are speculative and may not necessarily reflect the entirety of the

Figure L.5: ChatGPT response to the prompt “What keywords would you use to describe what the book *Mismatch* by Richard Sander and Stuart Taylor Jr. is about?” The response seemed reasonable but did not cover topics such as the legal battles over affirmative action or California Proposition 209, which multiple participants noted was an important topic. Also, the tag “Meritocracy” could be considered problematic when presented without context, as it is here.

 What keywords would you use to describe what the book *The End of Racism* by Dinesh D'souza is about?

 "The End of Racism" by Dinesh D'Souza is a book about race relations and politics in the United States. Here are some possible keywords that could describe the book:  

- Racism
- Race relations
- Identity politics
- Multiculturalism
- Affirmative action
- Civil rights
- Black culture
- White guilt
- Ethnicity
- Immigration
- Political correctness
- Social justice
- Conservative politics
- Liberalism
- American history

Please note that these are speculative and may not necessarily reflect the entirety of the book, but they represent common themes and topics that are likely to be discussed.

Figure L.6: ChatGPT response to the prompt “What keywords would you use to describe what the book *The End of Racism* by Dinesh D’Souza is about?” The response seemed reasonable, but the tags “Multiculturalism” and “Black culture” could be considered problematic when presented without context, as they are here, because of the way that the book frames these topics.

VITA

Chris Holstrom is a PhD candidate in the Information School at the University of Washington. His research focuses on subject indexing and subject tagging. Chris holds degrees in Physics, English, Technical Communication, and Information Science. He invites you to read more about his research and teaching at <http://chrisholstrom.com>, and he welcomes your comments at cholstro@uw.edu.