Quantifying Bias in Library Classification Systems

Katie Warburton[1,2] (kwarburton@cs.toronto.edu)
Charles Kemp[1] (c.kemp@unimelb.edu.au)
Yang Xu[2,3] (yangxu@cs.toronto.edu)
Lea Frermann[4] (lea.frermann@unimelb.edu.au)

[1]School of Psychological Sciences, University of Melbourne
[2]Department of Computer Science, University of Toronto
[3]Cognitive Science Program, University of Toronto
[4]School of Computing and Information Systems, University of Melbourne

Abstract

Categorization is ubiquitous in human cognition and society, and impacts how we perceive and understand the world. In reflecting the needs and perspectives of their creators, no categorization system is entirely objective, and inbuilt biases can have harmful social consequences. Here, we propose methods for quantifying three kinds of category biases in hierarchical category systems. We present a study on two widely used library classification systems (the DDC and LCC) as large-scale examples of human categorization, and use our methods to quantify bias towards content associated with western (vs non-western) concepts in topic areas including history and religion. We find consistent evidence for western bias and show that the DDC tends to exhibit more western bias than the LCC. Our methods are general, and can be used to survey biases across topic areas, bias attributes, and hierarchical category systems.

Keywords: natural categories; bias; library classification

Introduction

Categories inevitably reflect the needs, perspectives and experiences of the people who create them (Bowker & Star, 2000). Consider, for example, Steinberg's famous depiction of the *View of the World from 9th Avenue*[1], which devotes half of the page to three New York city blocks but shows China, Russia, and Japan as small blobs on the horizon. A *View of the World from Tiananmen Square* would look rather different, and might include separate categories for Hebei and Hubei provinces while making no distinction between Washington State and Washington DC.

Although systems of categories are often subjective, the distinctions that they encode or fail to encode can have important social consequences. For example, Gould (1990) points out that the United State's categorization of drugs as legal or illegal results in some addictive drugs being advertised on TV, and others carrying life sentences. Categories can also lead to harmful stereotypes, especially when coarse categories are used for members of outgroups in contrast to the finer-grained categories used for members of one's ingroups (Park & Rothbart, 1982). Because category systems can encode and reinforce stereotypes, it is important to ensure that the biases encoded by a system are acknowledged as such instead of treated as ground truth.

Here our goal is to develop methods for quantifying biases built into the structure of category systems. We take library classification as a case study, and focus on quantifying and comparing biases built into the Library of Congress Classification (LCC) and the Dewey Decimal Classification (DDC), two US-based systems. We focus on library classifications for three main reasons. First, they are large-scale examples of human categorization that are directly accessible and much more amenable to computational analysis than the category systems that we all carry around in our heads. Studying formal systems like these is valuable in its own right and can also contribute to a better understanding of categorization in general (Glushko et al., 2008). Second, library classification systems are often perceived as neutral or objective, which makes it all the more important to acknowledge the biases that may be implicit in them. Third, focusing on library classification allows us to build on a large body of existing work in the library and information sciences devoted to uncovering and mitigating bias in the LCC (Knowlton, 2005; Kam, 2007; Angell & Price, 2012; Howard & Knowlton, 2018), the DDC (Kua, 2008; Higgins, 2016; Westenberg, 2022), or both (Mai, 2010; Zins & Santos, 2011).

Previous work on bias in the LCC has documented that both the language used to label categories and the location of topics in the structure of the classification scheme can encode harmful bias. For example, unglossed religious terms like "God," and "devotional literature" refer to these concepts only in the context of Christianity (Knowlton, 2005). The DDC has been shown to be biased in its categorization of non-western languages and literature (Kua, 2008) and Asian American materials (Higgins, 2016), and frequently assigns disproportionately small category divisions to non-western racial and religious groups (Westenberg, 2022). The coverage of human knowledge by the first two levels of the DDC and LCC is biased and unsystematic, especially in the coverage of religions other than Christianity (Zins & Santos, 2011). Bias is prevalent throughout both systems and one of the most well-documented biases is western bias, or bias in favour of western culture. We therefore develop methods to quantify this bias and examine how it varies across topic areas.

Our work goes beyond previous studies of bias in library classification by identifying several kinds of category bias and systematically quantifying how category biases vary across topic areas and across the two systems that we consider. Quantifying bias is important because a quantitative measure can be used to identify the parts of a system that

[1]See https://saulsteinbergfoundation.org/essay/view-of-the-world-from-9th-avenue/

In M. Goldwater, F. K. Anggoro, B. K. Hayes, & D. C. Ong (Eds.), *Proceedings of the 45th Annual Conference of the Cognitive Science Society.*

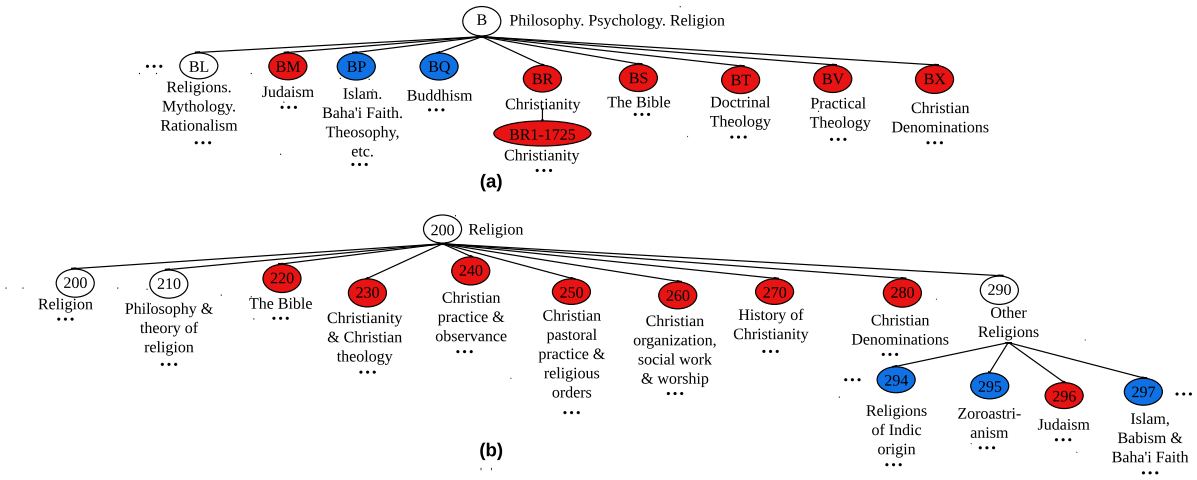


Figure 1: An example of how Religion is categorized in a) the LCC and b) The DDC. Red categories are tagged as western and blue as non-western. White categories cannot be tagged as just western or non-western.

show the strongest bias and are therefore most important to consider when proposing future improvements to the system. Although we focus on library classification, the methods we develop are broadly applicable to any hierarchical system of categories, and can potentially be applied to ontologies, biological classification systems, and hierarchical systems of natural categories (Mervis & Rosch, 1981).

We begin in the next section by providing background on the LCC and DDC. Next, we describe three kinds of bias that can be detected and measured in the structure of a hierarchical classification system. We then focus on western bias in particular, and present quantitative analyses that measure the extent to which the three kinds of category bias are present in the LCC and DDC. To preview our results, we find consistent evidence for western bias and show that the DDC tends to exhibit more western bias than does the LCC.

Library Classification Systems

The LCC and DDC share two key properties of library classifications. First, they were developed based on the principle of literary warrant. This means that their structures are derived from and justified by the body of literature that they classify (Svenonius, 2000). Second, they are hierarchical systems that contain a set of main classes, each corresponding to a different discipline. These main classes are recursively subdivided into increasingly more specific subcategories that classify smaller and smaller subsets of the literature.

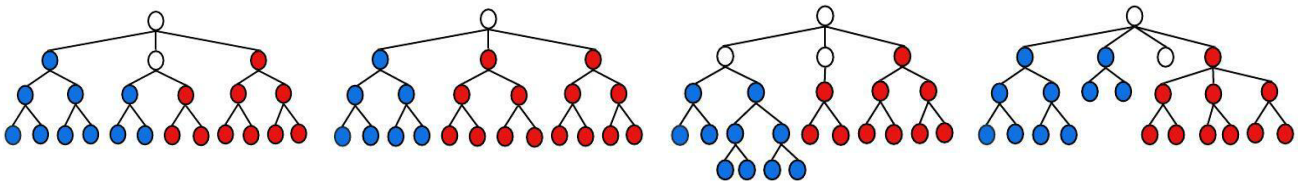
In the LCC there are 26 main classes and classification numbers are alphanumeric. There is no formal limit on the number of subcategories a category can have. The DDC has 10 main classes and classification numbers are entirely numeric. Each category can have a maximum of 10 children and the category hierarchy is represented by the position of the digit that differentiates a category. Figure 1 illustrates the classification of religious literature in both systems. The DDC enforces a strict upper limit on the number of subcategories (Svenonius, 2000). As a result, the LCC tends to be flatter and the DDC deeper.

Defining Category Bias

Concepts represented at the top level of a hierarchical classification are perceived as being more important than those at the bottom (Loehrlein, 2012). Similarly, the amount of detail and space afforded to a concept in a classification scheme likely affects how it is perceived. We thus define *category biases* as biases that are built into the structure of a hierarchical classification system. We defined three kinds of category biases: node count bias, level bias, and descendant bias, as illustrated in figure 2. Assume that blue and red represent distinct but comparable labels that can be applied to a set of categories in a classification system. For example, red categories could be those related to western topics and blue categories could be those related to non-western topics. For the examples in figure 2, we have assumed that an unbiased system treats red and blue categories identically.

Figure 2b illustrates node count bias, which occurs when there are more red than blue categories in a classification scheme. Level bias, in figure 2c, occurs when blue starting categories occur deeper in a classification structure than red starting categories. A *starting category* (or starting node) is the first category in a classification sub-tree that can be labelled as red or blue. Starting categories that are higher in the classification scheme are conceptualized as more general or important than those that are deeper. Level bias would present blue categories as more niche. Finally, figure 2d illustrates descendant bias, which occurs when red starting categories have more descendants than blue starting categories on average. Red categories are represented as having more fine-grained category divisions.

These biases may often be correlated in practice — for ex-



(a) Unbiased (b) Node Count Bias (c) Level Bias (d) Descendant Bias

Figure 2: An illustration of a) an unbiased classification b) a classification with node count bias, c) a classification with level bias, and d) a classification with descendant bias. The 3 category biases demonstrate bias against blue categories and in favour of red categories. White categories cannot be categorized as either red or blue.

ample, if there are more red categories (node count bias) it is likely that red starting categories will have more descendants (descendant bias). The biases, however, are conceptually distinct and can be separated in principle. For example, Figure 2d shows that even when node counts are held constant for red and blue it is possible to observe level bias (in favor of blue) and descendant bias (in favor of red). We therefore believe that considering the three biases individually is worthwhile as they highlight different aspects of category bias.

Methods

We used a tree structure representation to capture the hierarchical structure of the LCC and the DDC. Every node in the tree represents a different category in the classification, storing a category's name, label, and the books it classifies. The category label indicates the range of classification numbers that fall under it. The branches between a parent and its children represent the hierarchical relationship between a category and its subdivisions. We used the OhioLink Circulation Data as a large, publicly available data set of books classified under the LCC and DDC.[2] OhioLink contains 6.78 million MARC bibliographic records[3] representing books and manuscripts in the Ohio academic libraries (OhioLINK Collection Building Task Force, Gammon, & O'Neill, 2011). Only books that had both an LCC and a DDC classification were kept resulting in 3.31 million books. These books were placed into the DDC and LCC tree structures using their relevant classification numbers. For each book, we found the most specific category associated with its classification number, and then recursively added it to each parent category until the top of the tree was reached. This ensured that each parent category contained all the books of its subcategories.

To study western category bias, we selected main classes in the classification trees that related to history, religion, language, and literature as these categories tended to have category names that could be identified as western or non-western. Other topics such as philosophy, although having the potential to exhibit western biases, did not have categories that could clearly be identified as western or non-western. For the LCC these were categories in main classes B, D, E, F, and P which are named "Psychology, Philosophy, Religion," "World History," "History of the Americas," "History of the Americas (local)", and "Language & Literature" respectively. For main class B we only used the categories that classified religious literature. For the DDC these were main classes 2, 9, 4, and 8 which were named "Religion," "History," "Language", and "Literature." We treated language & literature as one topic because in the LCC many of the categories in main class P did not allow for easy separation of the two. Similarly, we grouped together all main classes related to history. We thus investigated western bias across three topics: religion, history, and language & literature.

We manually tagged the categories selected for each topic as western or non-western, drawing on distinctions that have been previously suggested in the literature.[4] Still, the tagging process is inevitably subjective, and in cases where a label of western or non-western was unclear, we left the category untagged, aiming for precision over recall. This somewhat limits the results, as there might be cases where a country or language, etc. falls into a category with a clear label in the LCC but not the DDC or vice versa.

The classes related to history tended to be divided into categories based on divisions such as country or continent. We therefore used a list of western countries that were defined based on a cultural definition as opposed to a political, economic, or geographical definition of what is considered western (de Espinosa, 2017; Trubetskoy, 2017; Hall, 2018). For example, Australia is considered a western country despite not being geographically in the western hemisphere. 68 countries, about 35% of the world's countries, were included in the list of western countries and it was assumed that countries left off the list were non-western. For each history-focused main class, we considered all categories associated with a country and tagged them as western or non-western based on the list. The tagged category became a starting category. If a category represented a group of countries (i.e. a category for a continent or a region) and all the categories beneath it shared the same tag, then that broader category became the starting cat-

[2]OhioLINK Circulation Data is made available by OCLC Online Computer Library Center, Inc. and OhioLINK under the ODC Attribution License

[3]The MARC 21 Format for bibliographic data is a digital format used to describe items catalogued by libraries.

[4]See `https://github.com/katie-warburton/Quantifying-Bias-Library-Classification` for the list of category tags.

egory and inherited the tag. Similarly, every category under a starting category inherited the starting category's tag.

In the language and literature-related classes, some categories were related to regional divisions like the history-focused classes so we based our tagging on the list of western countries used previously. Examples of these categories include "German literature" and "Languages and literature of Eastern Asia, Africa, Oceania." Some categories were related to language families so we considered where these languages or language groups originated from to make the tagging choice. "Romance languages" is one example. The main deviation from the tagging method used for history was how we tagged Indigenous languages and literature from North America, South America, and Oceania. Consistent with our cultural definition of the western concept (Hall, 2018), we tagged them as non-western even if they originated from a country or region that is listed as western.

Finally, for the main classes covering religion, we mostly tagged Abrahamic religions as western and other religions as non-western. The few exceptions included tagging Scientology as western and Islam as non-western. Islam is an Abrahamic religion, but its status as western is unclear. We made the conservative decision to tag Islam as non-western, because tagging Islam as western would probably only increase any western bias that we might find. The categories Doctrinal Theology and Practical Theology were tagged as western as they have been identified as only classifying literature on Christianity (Zins & Santos, 2011).

Results

In total there were 3009 categories on the topics of religion, language & literature, and history in the LCC, and 13,537 in the DDC. [5] Based on the tagging method, 87.9% of categories could be tagged as either western or non-western in the LCC, and 91.4% in the DDC. We refer to tagged categories as "nodes" to be consistent with our use of a tree representation.

Node Count Bias

To compute node count bias we compared the percentage of nodes tagged as western to the percentage of nodes tagged as non-western in both classification systems. In the LCC 2598 nodes were tagged (62.24% western). In the DDC there were 12,370 tagged nodes (69.36% western).

The same analysis was repeated on each topic individually, see figure 3. In each topic, there is a higher percentage of nodes tagged as western than non-western. In the LCC, religion has the highest percentage of western nodes. In the DDC this was history, however, religion had an almost comparable amount. For all topics, the DDC had a higher percentage of western nodes than the LCC. To test the significance of this result we randomly assigned all nodes a western or non-western label with equal probability. We repeated the process

[5] In the LCC 571 of the categories were from the topic Religion, 920 from Language & Literature, and 1518 from History. In the DDC 2420, 7758, and 3359 categories were from Religion, Language & Literature, and History respectively.

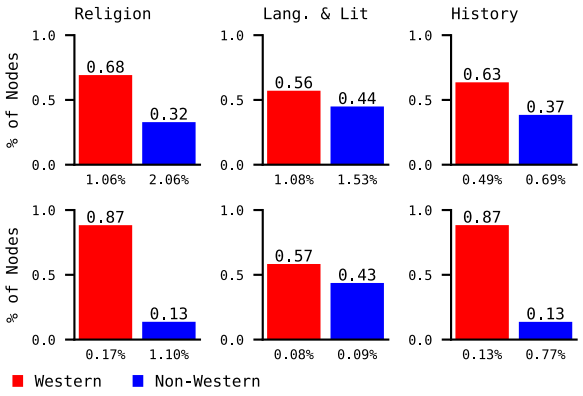


Figure 3: Node count bias for religion, language & literature, and history in the LCC (top) and the DDC (bottom). Each bar is annotated with the proportion of nodes it represents. The labels below the x-axis are the mean % of books per node.

10, 000 times, using the proportion of the times the absolute difference between western and non-western node counts was greater than or equal to the observed absolute difference as the p-value. For all topics in the DDC, and religion and history in the LCC, $p < 0.001$. For language & literature in the LCC, $p = 0.003$. All node count biases were significant.

We have conservatively assumed that an unbiased system has an equal number of western and non-western nodes, but this assumption could be adjusted using statistics such as population sizes or the percentage of western countries. If anything, these statistics tend to suggest that an unbiased system should devote more space to non-western than to western nodes. For example, Africa and Asia accounted for 75% of the population in 2022 (United Nations, DESA, Population Division, 2022). Western node count bias is substantial relative to a conservative 50-50 baseline, and would be even stronger relative to a a baseline favouring non-western nodes.

Library classification systems follow the principle of literary warrant thus it could be argued that there are more western categories because there are more western books. To test this, we calculated the mean rate of books per western node and non-western node in each system. These rates are reported as labels on the x-axis of figure 3. An unbiased system might be expected to have relatively equal rates of books per node. We found that language & literature in the DDC and history in the LCC have relatively equal rates of books per node for western and non-western nodes. Otherwise, there tend to be more books per non-western node than per western node. The difference in rates is most pronounced for religion (0.17% vs. 1.18%) and history (0.13% vs. 0.77%) in the DDC. This suggests that in some cases, especially in the DDC, the higher node count cannot entirely be accounted for by literary warrant. It is important to note that the rates of books per node come from books classified by the Ohio Academic libraries and thus the discussion of literary warrant is limited to how well the LCC and DDC fit these specific libraries.

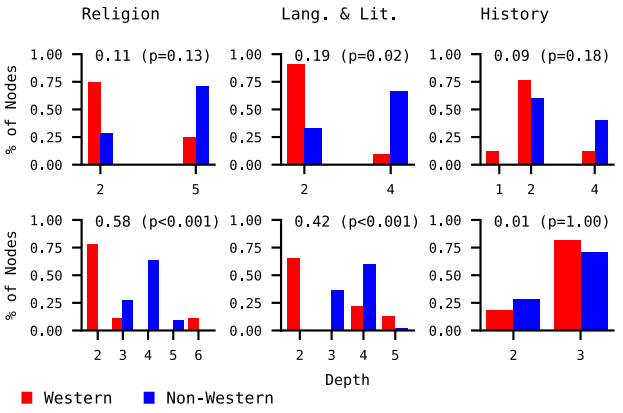


Figure 4: Level bias results for the topics religion, language & literature, and history in the LCC (top) and the DDC (bottom). Each plot is annotated with the JSD of the western and non-western distributions and the significance of the JSD in brackets. Significance was calculated with a permutation test.

Level Bias

To measure level bias we first compared the average depth of western starting nodes (starting categories) to the average depth of non-western starting nodes. If a group of nodes has a greater average starting depth then this implies that, on average, the nodes occur deeper in the tree than a group with a lesser average starting depth. To compute this statistic, the depths of the western and non-western starting nodes were collected and averaged. The LCC has an average western starting node depth of 2.28 and an average non-western starting node depth of 3.35. In the DDC these are 2.77 and 3.57.

For each topic, we plotted the distributions of starting nodes over the classification tree depths in figure 4. We then computed the Jensen Shannon Divergence (JSD) between the distribution of western and non-western starting depths to determine how divergent the two depth distributions were. To test the significance of the results we performed permutation tests. For each topic, the depth labels were shuffled among the western and non-western nodes to create randomized depth distributions. This was done 10,000 times and the proportion of times the JSD between the randomized western and non-western depth distributions was greater than or equal to the actual JSD was used as the p-value.

For all topics in the LCC, the average depth of non-western starting nodes is greater than the average depth of western starting nodes. This is also the case for religion and language & literature in the DDC. The one exception is history in the DDC, where the average western depth is slightly greater than the average non-western depth. When comparing the distribution of western starting nodes across the classification depths to the distribution of non-western starting nodes, language & literature has the most divergent distribution in the LCC, and religion has the most divergent distribution in the DDC. In both classification systems, history is the least divergent and not significant as $p = 0.18$ and $p = 1.00$ in the LCC and DDC respectively. In the LCC, only language & literature had a significantly divergent depth distribution ($p = 0.02$). In the DDC, this was religion ($p < 0.001$) and language & literature ($p < 0.001$). Religion did not demonstrate significant level bias in the LCC ($p = 0.13$).

	Prob. NW depth > W depth	
	LCC	**DDC**
Overall	0.87	0.82
Religion	0.89	0.89
Lang. & Lit.	0.95	0.76
History	0.68	0.36

Table 1: The probability that a randomly selected non-western starting node (NW) is deeper in a classification tree than a randomly selected western starting node (W), given that the two nodes are not equal.

The DDC tended to have a larger set of starting node depths than the LCC (i.e., the starting nodes for religion are spread across 5 different tree depths in the DDC versus just 2 in the LCC). This made it difficult to directly compare the LCC to the DDC. Therefore, as an alternative statistic for level bias, we calculated the probability that a randomly selected non-western starting node is deeper in the tree than a randomly selected western starting node. A western and a non-western starting node were randomly sampled 10,000 times. The depths of the two nodes were compared to determine the number of times the non-western one was deeper than the western one and vice versa. The number of times the two starting nodes were equal in depth was discarded. The resulting statistic measured the probability that a non-western starting node would be deeper in a tree than a western starting node, given they were not of the same depth. The results are shown in table 1. For every topic except history in the DDC, it is more likely that a randomly selected non-western starting node is deeper in the tree than a randomly selected western node. Western nodes for history in the DDC have a higher chance of starting deeper in the tree. For this statistic, the LCC displays a slightly stronger level bias than the LCC.

Descendant Bias

We measure descendant bias by comparing the mean number of descendants (direct or indirect) per western starting node to the mean number of descendants per non-western starting node. We also recorded the number of starting nodes and the mean percentage of books per starting node. This was done for the LCC and DDC overall, as well as separately for religion, language & literature, and history. The results are shown in table 2. To test for significance, the western and non-western tags were randomly shuffled among the starting nodes and the absolute difference in western and non-western descendant means was recomputed. This was repeated 10,000 times, using the proportion of times the recomputed difference in means was greater than the absolute value of the observed difference in means as the p-value.

		# of Start Nodes		% of Books / Start Node		Avg. # of Desc. / Start Node	
		W	NW	W	NW	W	NW
O	**LCC**	36	26	2.8	3.9	43.9	36.7
	DDC	43	56	2.3	1.8	198.5	66.7
R	**LCC**	8	7	12.5	14.3	44.5	23.3
	DDC	9	11	11.1	9.1	196.7	22.6
LL	**LCC**	11	9	9.1	11.1	29.2	27.9
	DDC	23	38	4.4	2.6	183.7	82.2
H	**LCC**	17	10	5.9	10.0	53.2	54.1
	DDC	11	7	9.1	14.3	231.0	51.9

Table 2: Summary of descendant bias for the LCC and DDC overall (O), and for topics religion (R), language & literature (LL), and history (H). Descendants is abbreviated to *desc*.

There is evidence of descendant bias in the DDC but not in the LCC. In the LCC, the difference in western and non-western descendant means was minimal and not significant as $p = 0.42$ for religion, 0.92 for language & literature, and 0.96 for history. In contrast, all topics in the DDC had, on average, at least double the number of descendants per western starting node than non-western starting node. Descendant bias was statistically significant for the topics language & literature ($p = 0.01$) and religion ($p < 0.001$) but not for history ($p = 0.62$). Although history had the largest absolute difference in means, its lack of significance can be attributed to the fact that it only has 18 identified starting nodes and one of them ("History of the North Americas") has a descendant count that is much larger than the other descendant counts.

The LCC has more starting nodes labelled as western. The DDC exhibits the opposite trend except in the case of history. The fact that there are more non-western starting nodes for language & literature and religion could account for the lower non-western descendant means in the DDC, however, the permutation test keeps the number of each type of starting node constant. Descendant bias is thus not an artifact of the number of starting nodes. The mean percentage of books per starting node is roughly equal between western and non-western ones. The biggest difference is found for history in the DDC, however the difference shows more books per non-western node than western (14.3% vs 9.1%). In the DDC, descendant bias is not explained by the idea that western starting nodes have more books than non-western starting nodes

Discussion

Category systems are subjective and can cause harm when they encode biases. Our goal was to illuminate these biases with new methods for quantifying structural biases in classification systems using western bias in library classifications as a case-study. Our results confirm previous findings that the DDC is biased in its categorization of non-western language and literature (Kua, 2008), and that both the DDC and LCC are biased in the space they afford to non-western religions (Zins & Santos, 2011; Westenberg, 2022). Beyond that, we found that the DDC is also biased in the number of categories it assigns to non-western history, and that the LCC is biased in the amount of space it affords to western history and in its placement of non-western languages and literature in the classification hierarchy. Our methods allow for the systematic comparison of bias in two systems, and we demonstrated that the DDC tends to show a higher degree of category bias than does the LCC. Specifically, there was evidence of strong node count and descendant bias in the DDC whereas there was no evidence of descendant bias in the LCC.

Our results highlight that library classification systems are not objective as the LCC and DDC are systems that reinforce a western perspective. We did not expect these systems to be unbiased as they were designed for a western population. However, the OCLC (2019) states that the DDC is used in libraries from at least 138 countries and has been translated into over 30 languages, and the Library of Congress[6] states that the LCC is one of the most widely used library classifications. Thus it is important to identify and remediate bias where possible, and our methods enable future work in this direction. For example, our methods identify that the classification of non-western language and literature encodes significant level bias. The LCC has no sub-category limits so a change such as moving the category for African languages and literatures so that it is at least at the same depth as the categories for English language and literature could help to remediate some level bias.

Our results have broader implications for natural category systems. Institutional categorizations such as library classifications often emerge from many contributors with diverse perspectives working to create a system that accurately characterizes the items it classifies. If we still find evidence of bias in these systems then we suspect that many natural category systems would show even more bias as they do not have the same goal of objectivity nor do they necessarily benefit from diverse perspectives. Future work could test this idea by applying our measures of category bias to hierarchical representations of natural category systems.

Finally, although our results indicate western bias in both classifications, our analysis is limited to a relatively coarse distinction between western and non-western because we are unable to determine if the actual books assigned to each category are written from a western or non-western perspective. It is likely that in a US-based library books classified as non-western still exhibit many kinds of western bias. Limited as they are, however, our metrics do succeed in capturing biases built into the hierarchical structure of the systems.

In summary, our methods allowed us to systematically quantify and compare the amount of western bias across different topics and between different systems. Although we focused on western bias in library classification systems, these methods are not restricted to these systems or biases. *View of the World from 9th Avenue* would not score well on our definitions of category bias, but quantifying its bias could serve as a first step towards creating a more balanced categorization.

[6] https://www.loc.gov/catdir/cpso/lcc.html

Acknowledgments

We would like to thank anonymous reviewers for constructive comments on an earlier version of the paper. KW is funded by the U of T–UoM IRTG program. This work was supported by NSERC Discovery Grant RGPIN-2018-05872 and ARC Future Fellowship FT19010020.

References

Angell, K., & Price, C. (2012). Fat bodies in thin books: Information bias and body image in academic libraries. *Fat Studies*, *1*(2), 153-165. doi: 10.1080/21604851.2012.641895

Bowker, G. C., & Star, S. L. (2000). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

de Espinosa, E. L. (2017). *Is Latin America part of the West?* Elcano Royal Institute. Retrieved from https://www.realinstitutoelcano.org/en/work-document/is-latin-america-part-of-the-west/

Glushko, R. J., Maglio, P. P., Matlock, T., & Barsalou, L. W. (2008). Categorization in the wild. *Trends in Cognitive Sciences*, *12*(4), 129-135. doi: https://doi.org/10.1016/j.tics.2008.01.007

Gould, S. J. (1990). Taxonomy as politics: The harm of false classification. *Dissent*, 73–78.

Hall, S. (2018). The West and the Rest: Discourse and power [1992]. In *Essential essays, volume 2* (pp. 141–184). Duke University Press.

Higgins, M. (2016). Totally Invisible: Asian American Representation in the Dewey Decimal Classification, 1876-1996. *Knowledge Organization*, *43*(8), 609-621. doi: https://doi.org/10.5771/0943-7444-2016-8-609

Howard, S. A., & Knowlton, S. A. (2018). Browsing through bias: The library of congress classification and subject headings for African American studies and LGBTQIA studies. *Library Trends*, *67*(1), 74–88. doi: doi:10.1353/lib.2018.0026.

Kam, D. V. (2007). Subject headings for Aboriginals: The power of naming. *Art Documentation: Journal of the Art Libraries Society of North America*, *26*(2), 18-22.

Knowlton, S. A. (2005). Three Decades Since Prejudices and Antipathies: A Study of Changes in the Library of Congress Subject Headings. *Cataloging & Classification Quarterly*, *40*(2), 123–145. doi: https://doi.org/10.1300/J104v40n02_08

Kua, E. (2008). Non-Western Languages and Literatures in the Dewey Decimal Classification Scheme. *Libri*, *54*, 256–265. doi: https://doi.org/10.1515/LIBR.2004.256

Loehrlein, A. J. (2012). *Priming effects associated with the hierarchical levels of classification*. Doctoral dissertation, School of Library and Information Science and the Department of Cognitive Science, Indiana University.

Mai, J.-E. (2010). Classification in a social world: bias and trust. *Journal of Documentation*, *66*(5), 627–642.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology*, *32*(1), 89–115.

OCLC. (2019). *Introduction to the Dewey Decimal Classification.* OCLC, Inc. Retrieved from https://www.oclc.org/content/dam/oclc/dewey/versions/print/intro.pdf

OhioLINK Collection Building Task Force, Gammon, J., & O'Neill, E. T. (2011). *OhioLINK–OCLC collection and circulation analysis project 2011.* OCLC Research. doi: https://doi.org/10.25333/C30341

Park, B., & Rothbart, M. (1982). Perception of out-group homogeneity and levels of social categorization: Memory for the subordinate attributes of in-group and out-group members. *Journal of Personality and Social Psychology*, *42*, 1051-1068.

Svenonius, E. (2000). Subject Languages: Introduction, Vocabulary Selection, and Classification. In *The Intellectual Foundation of Information Organization.* Cambridge, MA: MIT Press.

Trubetskoy, S. (2017). *List of western countries.* Retrieved from https://sashamaps.net/docs/maps/list-of-western-countries/

United Nations, DESA, Population Division. (2022). World population prospects 2022: Summary of results. *UN DESA/POP/2022/TR/NO. 3*.

Westenberg, J. (2022). Where 'Dewey' go from here? perpetuation of difference in public library knowledge organization systems. *Aletheia: The Arts and Science Academic Journal*, *2*(2). doi: https://doi.org/10.15173/a.v2i2.3002

Zins, C., & Santos, P. L. (2011). Mapping the knowledge covered by library classification systems. *Journal of the American Society for Information and Technology*, *62*(5), 877—901. doi: https://doi.org/10.1002/asi.21481