

Claire Hubble

LIS 708 02

April 24, 2024

### Proposed Diversity Audit of the Physical Books in the Harvard Theatre Collection

Although there are many, unique library collections on the campus of Harvard University, none are as exciting as the Harvard Theatre Collection: composed of print books, photographs, designs, and digital objects documenting the history of the performing arts. Created in 1901, the collection covers a wide range of performing arts documentary material and is housed in the Houghton Library (*Harvard Theatre*). Houghton is one of over 70 libraries scattered across Cambridge and Boston that make up Harvard Library. The Theatre Collection, like all Harvard Library collections, is accessible by researchers and visitors of all kinds whether they are directly related to Harvard University or not.

Houghton Library emphasizes the accessibility of their collections through their stated value of “Welcome the World,” a phrase meant to encourage all people who are interested in Houghton materials to have easy access to those materials (*Houghton*). Due to the rare and unique nature of so many of the Theatre Collection items, users may view and handle these items inside the Houghton reading room only, limiting accessibility of non-digital items to the library’s hours of operation and its Cambridge location. Everyone with access to the Internet has access to HOLLIS and HOLLIS for Archival Discovery, the online public catalogs where the majority of Theatre Collection items have been described. Using simple search filters, a user could limit their findings to only print books within the Theatre Collection, for example.

The Theatre Collection is quite wide in scope, covering most areas of performing arts including theater, dance, opera, minstrelsy, magic, and puppetry. Within these categories, there

are a number of print books, manuscripts, audio/visual items, musical scores, ephemera, promptbooks, and digital objects. There are roughly 24,000 print books in the Theatre Collection, but the total number of items across all formats and content areas is in the millions.

Just as diverse as their Theatre Collection, the student body of Harvard is made up of almost 25,000 students. Of these students, the biggest demographic groups are students identifying as White (32.4%), as Asian (16.2%), and as Hispanic or Latino (9.1%), while roughly 27% of Harvard students are non-U.S. residents. The figures are not proportional to the faculty of Harvard, with an overwhelming majority of professors, administrators and support staff identifying as White (*Fact Book*).

### **Service Evaluation Inquiry**

The Houghton Library collections strategy emphasizes diversity as its second most important pillar after access, explicitly mentioning the Harvard Theatre Collection's aim to, “[focus] on developing collections relating to performances staged by or involving overlooked and marginalized individuals and communities” (*Houghton*). As an academic library that serves a diverse student population, it is very important that the collection demographically reflects its users. As the oldest and largest theater collection in the country that is open to the public, it is very important that the collection is demographically representative of the United States, including the marginalized communities the collection strategy mentions.

Although there is a push from Harvard Library for all of its collections to engage in diversifying, there have been no reports or evaluations created specifically about the Harvard Theatre Collection available to the public. This author proposes a diversity audit of the Theatre Collection in order to locate gaps in the representation of voices within the print books of the collection. The results of this project will be written into a report that will make selecting

materials to add to the collection easier. The evaluation will answer: are the Harvard Theatre Collection print books diverse in ways that accurately reflect the diversity of American theater and its audiences? Where are the gaps in representation within this collection?

### **Literature Review**

There have been a number of diversity audits that have taken place in public, academic, and special libraries all over the country within the past few years. For example, Jones, et al. (2023) describe strategies used by the University of Nevada, Las Vegas Special Collections and Archives (UNLV/SCA) to assess their collections, strategies and initiatives within the context of diversity, equity, and inclusion. For the collection portion of the evaluation, these researchers looked at digital, SCA accession records between 1967 and 2000, representing 1,627 manuscripts, photographs, and university archives. These records were then cross-referenced with corresponding archival collections to see exactly what subject headings were assigned within their finding aids. The content-specific audit is inspiring, however feels unmanageable for a collection the size of the Theatre Collection, almost 14 times the size of the Nevada Collection. With subject data compiled, the authors refined categories of diversity and identified special topics categories with significance to UNLV. Finally, the accession data was marked “diverse” or “non-diverse” based on the primary subject of each individual item (Jones, et al). The authors really emphasize the importance of quantitative data (made by the hard work of coding each item) within this study and their research goals of evaluating diversity were better served because of it.

While the Nevada librarians elected to focus on subject area, Emerson and Lehman (2022) focused their assessment on authors and voices within the collection, looking specifically at gender, sexuality, and race/ethnicity. To make the audit of the Thomas Tredway Library at

Augusta College manageable, Emerson and Lehman created a list of criteria each item had to follow in order to be included in the study such as formatting, author, and publisher specifications. Using autobiographical data and Internet resources, the evaluators gathered demographic data on all authors with books eligible for the study and stored this data in a central, collection spreadsheet. These methods reflect what might be more appropriate for the Theatre Collection, as this is the first diversity audit of the physical books. The results of this study were intended to assist librarians develop a better understanding of whose perspectives are represented in this collection and to “provide insight as to whether the library was meeting the goal of its Strategic Plan regarding building an inclusive collection” (Emerson). These are the results Harvard hopes to take away from this proposed evaluation and hope to use similar methods on a bigger scale.

Similar to Emerson and Lehman in purpose, Stone’s evaluation (2020) of the University of California Irvine’s playscript collection is one of a kind. It is the only theater collection-specific diversity audit that is published as of today. Stone’s goals for this evaluation were to (1) determine the demographics of playwrights of plays in the UCI libraries, (2) compare how the collection has changed from FY2011 to FY2019, and (3) to reflect on how the playscripts collection is demographically representative of plays published in a given fiscal year. Using advanced search features in GOBI, Stone made a filtered list of plays in the UCI collection within the given years, removing newly published classics and digital copies specifically. Due to the vast amount of items in the Harvard Theatre Collection, a similar approach will be beneficial to all involved, however for the Harvard evaluation the filtering will focus on authors that appear multiple times within the collection. Just like Emerson and Lehman, Stone conducted demographic research on each individual playwright—focusing on gender, sexual orientation,

ethnicity and nationality—and gathered all the data into a Google spreadsheet. There is no mention of coding the items like the two previously mentioned evaluations, however Stone did limit his vocabulary when sorting as evidenced by the data chart titles and labels included in the study.

All three evaluations emphasize a need to closely follow definitions of demographic elements. Beyond defining diversity, all of these authors mention the complexities of putting humans into categories. This complexity can be made a bit simpler by having a good understanding of what constitutes specific categories, and being aware of when there is a need to break down general categories into smaller and more specific groups. For example, Stone only uses “Heterosexual” and “LGBTQ+” when evaluating sexuality of playwrights within his collection (Stone). Emerson and Lehman, on the other hand, identified authors who self-identified as being in opposite sex relationships, same sex relationships, or specifically identifying as being part of the LGBTQA+ community (Emerson). Although the methodology of these two studies was very similar, the data produced might be vastly different because of the categories used.

### **Evaluation Data and Methods**

The purpose of this evaluation of the Harvard Theatre Collection is to gain insights about what condition the collection is in and where there are gaps, in regards to diversity. Based on this goal and the information gleaned from the reviewed literature, the following research methods are recommended: pull a shelf list of physical books, establish data categories, assign codes, and perform comparative analysis of the data. Each of these steps will be explored in greater detail in the following paragraphs. Although these steps might not seem like hard tasks to complete, the size of the collection means this evaluation could take up to two years to complete, beginning in

Summer 2024 and completed by the end of Summer 2026. The evaluation needs to be a priority for the Library, not only to better assist current students and researchers in their learning, but to have a robust collection for future (and more diverse) generations.

Firstly, gather the appropriate, preliminary data needed for the study. The first group of data required relates to people. Demographic data, pulled from the U.S. Census and Harvard University, can be used to better understand the racial and ethnic makeup of the United States and the Harvard student body, respectively. The second group of data is regarding the Theatre Collection items in the form of a shelf list. This way, the print books—including some rare and unique items—can be assessed without handling every individual book and without having to go through items not included in this evaluation such as Theatre Collection videos and ephemera. All team members working on the project should have access to the shelf list at all times.

Team members must also have a good understanding of what they are assessing. Because theater is an embodied artform, the vast majority of playwrights create stories and publish plays with their lived experiences at the focus, featuring characters that act, talk, and look like them. It is very safe to assume when a white playwright finishes a script, the protagonist is white. This is not to say all playwright/play relationships are this way: *The Thanksgiving Play* (2015) by Indigenous playwright Larissa Fasthorse features four white characters only. However, inclusion of this title in the collection would provide an Indigenous perspective nonetheless, satirically subverting stereotypes (and theatrical history) with the choices she made while creating the story. With the addition of more playwrights of various races and ethnicities within the collection will come inclusion due to the nature of plays and playwriting. For these reasons, this evaluation will focus on playwright race and ethnicity when assessing diversity, instead of content of the plays themselves.

Keeping that framework in mind, categories of race and ethnicity should be created, along with codes for each categorical element. For the purposes of this evaluation, it is best to use eight of the nine races utilized by the U.S. Census Bureau. These categories are almost an exact match to the categories used by IPEDS to conduct demographic studies of the Harvard community. Data from this evaluation can then be directly compared to national and local demographics easily. The categories, along with 3-character codes for quick inputting, include:

- American Indian/Alaskan Native (INA)
- Asian (ASA)
- Black or African American (BLK)
- Hispanic or Latino (LAT)
- Native Hawaiian or Pacific Islander (HPI)
- White (WHT)
- Two or more races (XYZ)
- Non-U.S. Resident (N/R)
- Not specified (UNK)

Left off from the Census categories is “Non-Hispanic/Latino” because it will not yield data that is helpful for the purposes of the diversity audit. Codes were chosen so no two codes have the same starting letter in the hopes that keyboard shortcuts can be employed. Although there are many, more specific races and ethnicities in our world and on our shelves, the need to limit the amount of categories outweighs the need to be more distinct at this point in auditing. Hopefully, audits such as this continue in the future, when there is more opportunity to be more exact, and thus more representational, inclusive, and diverse.

After categories and codes are established, the assessing can begin. Using the pulled shelf list, race/ethnicity codes could be given to each author. Researching this information will be very hard for authors who died before author websites and social media bios were popular, so reliance on primary and secondary resources is necessary. For contemporary and alive playwrights, demographic information can also be found on publisher sites and within editorial profiles. This

information would need to be stored on a central Excel sheet, accessible by the many staff members working to complete this project. (Ideally, this spreadsheet is presented to library catalogers at the end of the project so this information can be stored in each book's metadata within the catalog, eliminating the need to go through and manually assign codes to authors for future audits.) Because the focus is authors, the number of items needed to be researched will be smaller than the amount of physical books in the collection (24,000) because there are many authors with multiple titles and editions present. However, the number of titles per author will need to be noted in order to create accurate percentages in the next section of the project. It is in the Theatre Collection's best interest to take the time needed to conduct this research.

After the labor-intensive section is concluded, charts and tables can be created in order to better visually understand the state of the collection in regards to diversity. Since we are working on Excel, pivot tables, pie charts, and bar graphs can be created quickly and exported into any third-party softwares. Comparative analysis can be conducted with the demographic data and the newly-complete list of author races within the collection. Although the hope is never to have too big of gaps anywhere in the collection, in terms of diversity, we are expecting to have much room to grow. In terms of data, this will mean looking at the same category of race across demographic and collection data and identifying where ratios and percentages are off.

Concluding this project will give the Harvard Theatre Collection a much better understanding of who is represented on our shelves, and where there needs to be significant work done in order to accurately reflect American theater history and its audiences.

### **Final Reflections**

An evaluation of this magnitude and for this purpose will always be worth the time, energy, and money required from the library and its workers. Until there is a standardization in

this field of noting diversity elements in books as they are cataloged, diversity audits are bound to labor-heavy research methods. (But even this, demanding authors put their identities on display, is problematic.) This proposal has allowed me to fully understand the importance of nuanced research methods when approaching any and all evaluations—of collections, of services, spaces, or programs. There were also some areas of this project where I conceded my true wants to serve the workers and end goal of the project, such as not having more specific categories of race when assigning codes. Finally, it would be best for everyone if this proposal was more influenced by librarians of color. As a white woman, I might be unaware of the biases some of the above choices reflect, so input from people whose lived experiences are a central theme of this evaluation is absolutely crucial.

## Bibliography

Emerson, María Evelia, and Lauryn Grace Lehman. "Who Are We Missing? Conducting a Diversity Audit in a Liberal Arts College Library." *The Journal of Academic Librarianship*, vol. 48, no. 3, May 2022, p. 102517. <https://doi.org/10.1016/j.acalib.2022.102517>.

*Harvard Theatre Collection* | *Harvard Library*,

[library.harvard.edu/collections/harvard-theatre-collection](https://library.harvard.edu/collections/harvard-theatre-collection).

*Houghton Library Collections Strategy* | *Harvard Library*, 5 Oct. 2020,

[library.harvard.edu/houghton-library-collections-strategy](https://library.harvard.edu/houghton-library-collections-strategy).

*Fact Book: Enrollment* | *Office of Institutional Research & Analytics*,

<https://oira.harvard.edu/factbook/fact-book-enrollment/>. Accessed 25 Mar. 2024.

Jones, Sarah, et al. "Assessing Diversity in Special Collections and Archives." *College &*

*Research Libraries*, vol. 84, no. 3, 2023. <https://doi.org/10.5860/crl.84.3.335>.

Powe, Kymberlee. "From Diversity to Inclusion: How to audit your collection and why."

Connecticut State Library. January 4, 2022. (1 hr 14 mins).

[https://www.youtube.com/watch?v=vzK0dL\\_QSC0&t=1s](https://www.youtube.com/watch?v=vzK0dL_QSC0&t=1s)

*Shelflisting Introduction* | *Yale University Library*.

<https://web.library.yale.edu/cataloging/shelflisting-intro#>. Accessed 24 Apr. 2024.

Stone, Scott M. "Whose Play Scripts Are Being Published? A Diversity Audit of One Library's

Collection in Conversation with the Broader Play Publishing World." *Collection*

*Management*, vol. 45, no. 4, Oct. 2020, pp. 304–20.

<https://doi.org/10.1080/01462679.2020.1715314>.

*United States - Census Bureau Profile.*

[https://data.census.gov/profile/United\\_States?g=010XX00US#race-and-ethnicity](https://data.census.gov/profile/United_States?g=010XX00US#race-and-ethnicity).

Accessed 16 Apr. 2024.