Connectivity vs. Canonicity: Data Science and Dance Studies in Historical Dialogue

Harmony Bench¹, Kate Elswit², and Antonio Jimenez-Mavillard²

¹The Ohio State University
²University of London, Royal Central School of Speech and Drama

1 Introduction

This paper is the joint effort of two dance scholars and a data scientist who work together on the AHRC-funded project Dunham’s Data: Katherine Dunham and Digital Methods for Dance Historical Inquiry (AHRC AH/R012989/1). Katherine Dunham was an African American dance artist, teacher, and activist whose career spanned eight decades and six continents. In the larger project, we turn to feminist and anti-racist digital methods following such scholars as D’Ignazio and Klein (2020) who encourage researchers to bring bodily materiality back into digital analyses, and Parham (2016) and Johnson (2018), who show how such approaches are particularly suited to engaging with African diasporic practices, both transnationally and transhistorically. We additionally build on data-informed research in theatre history, including projects that trace connections among performers, the roles they perform, and the circulation of theatrical repertoire (Balme, 2019; Bollen and Holledge, 2011; Caplan, 2016; Miller, 2020; Varela, 2021). The network-based approach to historical dance repertory in this project unsettles both a more familiar choreographic canonicity and the isolated “genius” of the modernist choreographer, emphasizing instead alternate systems of value that are foregrounded by attending to connectivity and the roles of performers in developing and maintaining repertory (Croft, 2015). Our research utilizes digital methods to shift disciplinary conversations in the performing arts about touring artists and their repertory, open new avenues of investigation, and illuminate aspects of specifically dance historical inquiry that would not otherwise be possible.

We focus on i) a dataset covering the repertory that Katherine Dunham’s dance company performed between 1947-60, which we manually curated primarily from undigitized archival programs; ii) the visualization of this data as an interactive network graph in order to represent both vertical (nested hierarchies of Dances-in-dances, Pieces, Containers, and Shows) and horizontal (components that are performed alongside each other, as well as remixed or swapped out) relationships; iii) network analyses such as centrality measures based on the connectivity of these entities; and in the longer paper, we address iv) the imperfect processes of inference and interpretation we draw from this messy data as we build toward historical argument (Rawson and Muñoz, 2019).
Dunham actively repurposed elements of her choreography throughout her career. In contrast to the seemingly stable identity of works that historical narrativization can erroneously suggest, a network approach to Dunham’s repertory better represents the fluidity and connection evident in her programs. Furthermore, our research reveals a tension between connectivity and canonicity. Whereas dance critics and scholars designate certain works as canonical on the basis of critical reception and perceived aesthetic or historical significance (Clark and Johnson, 2005; Franko, 2007), Dunham’s Data suggests there are other ways to evaluate the importance or value (Herrnstein Smith, 1983) of specific repertory within an artist’s oeuvre. Measures of connectedness within a network are one such possibility that performing arts scholars are beginning to explore as an alternative to the canon (Balme, 2019; Bollen, 2020). Such measures may or may not correlate to frequency of performance or the opinions of critics and historians, and divergences between the historical narrative and the results of data analysis are particularly intriguing.

2 Gathering Katherine Dunham’s Archival Data

Katherine Dunham was a prolific 20th century dance artist who choreographed and performed in operas, nightclubs, Broadway shows, Hollywood films, and modern concert dance. Over a career that spanned 80 years, Dunham worked across five continents in many contexts, from her early anthropological research in Haiti to her curatorial, pedagogical, activist, and administrative projects later in Dakar, New York, and East St. Louis. Dancers around the world today continue to cite her as an influence and inspiration. The mid-1930s to early 1960s represented her most active years as a choreographer heading the touring dance company Katherine Dunham and Her Dancers, Singers, and Musicians (Beckford, 1979; Clark and Johnson, 2005; Dee Das, 2017). As a mid-century celebrity in the field of dance, Dunham’s archives were robust enough to support a large-scale research effort to determine how digital methods might be employed toward the questions and problems that are meaningful to dance historians. In the case of Katherine Dunham, such questions include: What are the connections between where Dunham traveled, the cultural backgrounds of the dancers, singers, and musicians she employed, and how she represented real and imagined places onstage? Who performed together, and what are the implications for how embodied knowledge was shared within the company and then beyond as performers left to pursue other opportunities? Given that Dunham had a decades-long career during which she structured her shows as revues, what did she maintain in her active repertory and how did her repertory change over time? How can we better understand her repertory through the interconnectedness in how she represented choreographic entities over time?

The project team compiled three primary datasets: one documenting Dunham’s daily travels from 1947-60, one documenting the nearly 200 performers who joined and left her company during that time, and one documenting the nearly 200 repertory numbers they performed (Bench and Elswit, 2020). These datasets were manually curated from archival sources held in the Dunham collections at Southern Illinois University at Carbondale, Missouri Historical Society, Library of Congress, New York Public Library, and Harry Ransom Center. The data representing Dunham’s repertory came from the way she framed her work for mid-century audiences, as archived through undigitized concert programs, and supplemented by correspondence, lighting plots for theatrical presentation, scripts for introducing each dance number from the stage, and other materials from public archives.
This is an annotated excerpt of a 1955 Bal Negre program using colored dots to indicate the level each repertory item occupies in our nested hierarchy. At the highest level are Shows (not pictured, but which are indicated with orange), followed by Containers (red), Pieces (blue), Mixed works, which sometimes appear in programs as Pieces and sometimes as Containers (purple), and at the smallest level are Dances-in-dances (aqua), which are generally named social dances. This program contains the Container "Americana," the Pieces "Field Hands," "Spirituals," "Barrelhouse," "Flaming Youth," "Cakewalk," and the Mixed work Plantation Dances, which is an excerpt from "Southland" presented under the auspices of "Americana." Image courtesy Special Collections Research Center, Morris Library, Southern Illinois University Carbondale.

Although dance historians commonly represent choreographers’ bodies of work as a simple list of works in repertory and their premiere dates, this approach was insufficient to capture the shifting relationships among these entities over time as elements were recycled and repurposed for new contexts. While dance historians largely remain suspicious of data-driven research, this conventional practice of representing repertory as a list of works is its own form of “cleaning” that obscures the complexity of that historical data. By contrast, our Repertory Dataset retains complexity by foregrounding a nested hierarchy of performance elements, as well as including the various titles and descriptions by which each choreography might be known; the years in which it was performed; whether it appeared in concert halls, nightclubs, or both; all documented durations; and all of the singers, dancers, and drummers who are ever listed as having performed in it. The dataset also captures both Dunham’s particular language in describing (times and) places of her choreographic work, and the real and imagined geographies that they index.

We identified a nested hierarchy of these performance elements in Dunham’s programs, which became a key element in the structure for the Repertory Dataset, the focus of our network visualizations (see Figure 1). Following disciplinary conventions, we adhered to the evening-length “Show” and the substantive but comparatively short “Piece,” but we added the category of “Container,” as an act-length entity into which pieces are sometimes gathered. “Dances-in-dances,” which are generally social dance forms that are identified in the programs, and “Mixed-works,” which describes entities that sometimes appear in archival documents as pieces, and sometimes as containers. Analyzing repertory as a network allows us to see the interconnections among these
various entities, and challenges the stability of the choreographic work as such (Pakes, 2020; Pouillaude, 2017).

3 From Canonicity to Connectivity in Katherine Dunham’s Repertory

Focusing on repertory in this way complicates Dunham’s canon. Proportional representation in dance scholarship suggests that Dunham’s works of greatest historical importance include “L’Ag’Ya,” a love story centered around a Martiniquan fighting dance, and “Barrelhouse,” which portrays dancers in an American juke joint. Both of these were in active repertory for decades. The imagined rituals of personal and communal transformation depicted in “Rites de Passage,” which grew from two pieces to four that were almost never performed all together, receive a substantial proportion of scholarly attention. Finally, “Southland,” Dunham’s most politically pointed work addressing racist violence, has been considered extensively by scholars even though it was performed only twice in Santiago, Chile in 1950, and included in a short run in Paris in 1953 (Hardin, 2016; Hill, 1994). Dunham’s canonical choreography as represented in scholarship does not reliably correlate to data such as the number of years certain repertory was performed, indicating popularity with audiences. If the construction of canons is always determined by a mixture of affective and arbitrary logics, other possibilities emerge to determine the relative importance of repertory, in particular in relation to connectivity. For example: Which dance numbers were performed most frequently? Which stayed in active repertory the longest? Which had the most performers pass through? Secondary questions then follow: Which clusters of repertory were performed together most often? Which changed the most and least over time? Which form crucial linkages across the repertory as a whole or across generations of performers?

Every choreographic entity listed in archival documents and represented in the Repertory Dataset became a node in the Repertory Network (see Figure 2). Each was assigned one of the five choreographic types based on the previously identified nested hierarchy (Shows, Containers, Mixed works, Pieces, Dances-in-dances). The diameter of each node in the visualization positively correlates to its frequency of occurrence in the referenced archival documents. The Repertory Network preserves complexity of the data by maintaining the connections among all elements of which each show, container, or piece was composed without promoting a single, definitive version. All alternate names for each entity are available on right-click.

Node connectivity defines which nodes are positioned in the center of a network and which ones stay in the periphery, and their patterns can be captured by network statistics. More central nodes tend to connect otherwise disconnected areas of the network, and therefore centrality can assist in determining how “important” different nodes are in the network. This “importance” can be quantified through centrality measures. We calculated four of the most typical centrality metrics for network analysis: degree, eigenvector, closeness, and betweenness centralities. Our starting hypothesis was that the most frequently occurring Pieces and Containers would also be the most canonical as recognized by dance historians, as well as the most connected within the network. In the analysis of social networks, centrality can indicate influence, which might seem to parallel the historical and cultural influence of canonicity. However, we discovered that in the case of Dunham’s repertory, centrality is not a reliable predictor of canonicity and vice versa.
For Containers, the frequency of occurrence correlated with centrality metrics as expected, in other words, the most represented Containers are the most central in the network (although there are remarkable outliers). For example, “Brazilian Suite,” “Motivos,” and “Americana” are the three most frequent Containers in programs and also have the best scores for degree, eigenvector centrality, and closeness centrality. However, Pieces did not behave as expected. For example, we found that “Barrelhouse” and “Bahiana (Brazil) / Batucada” are the two most frequently performed Pieces listed in Dunham’s programs, but they have only 3 and 4 connections to Containers, respectively. Another example is “Para que tu veas,” with 49 occurrences in programs and an eigenvector value of 16 (almost the highest of the network), which contrasts to “Cakewalk,” which has almost double the number of occurrences in programs and an eigenvector value of only 3.

From the network visualization, we can see that the largest nodes in the network are not necessarily the most connected (highest degree) or the best positioned (most central position). This means that when considering Dunham’s pieces, the repertory that historians have identified as the most significant are not necessarily the ones performed most frequently, nor what network analysis reveals to be the most connected. Similarly, the pieces that emerge as the most central from a network analysis perspective are not necessarily the ones best represented in dance scholarship.

4 Conclusion

Analyzing how deeply choreographed works are connected using centrality metrics (degree, eigenvector, closeness, and betweenness centralities) offers a different view of the importance and impact of specific repertory than proposed by traditional humanities
methods. We consider how emphasizing connectivity over canonicity, supported by approaches to network analysis in the digital humanities (Painter et al., 2019), can radically shift dance historians’ evaluations of repertory, and how the network graph itself makes the interconnections across repertory visible.

References


