Geographic Distribution of Scripps National Spelling Bee Spellers Resembles Geographic Distribution of Child Population in US States upon Implementation of the RSVBee “Wildcard” Program

Rosy E. Kannankeril¹ and Prince J. Kannankeril²
¹Franklin Classical School, Franklin, TN, ²Vanderbilt University, Nashville, TN

SUMMARY
The Scripps National Spelling Bee (SNSB) is an iconic academic competition for United States (US) schoolchildren, held annually since 1925. Historically, children qualified for the SNSB by winning a sponsored regional spelling bee. However, the sizes and geographic distributions of sponsored regions are uneven. One state may send more than twice as many spellers as another state, despite similar numbers in child population. In 2018, the SNSB introduced a wildcard program known as RSVBee, which allowed students to apply to compete as a national finalist, even if they did not win their regional spelling bee. This allowed more students to compete, but it was unknown how this would affect the geographic distribution of spellers. The purpose of our experiment was to test our hypothesis that the geographic distribution of SNSB national finalists more closely matched the child population of the US after RSVBee was implemented. We compared the number of sponsored (non-RSVBee) national finalists to the US child population from each of the 50 states, the District of Columbia, and Puerto Rico, for the years 2012–2019 and found a stable, strong correlation in each year, with correlation coefficients ranging from 0.741–0.774. With RSVBee in 2018 and 2019, we found the correlations increased significantly, with correlation coefficients ranging from 0.916–0.917 for all spellers (p-value = 0.032 by paired t-test). We conclude that the RSVBee program significantly improved the geographic distribution of the SNSB by matching the geographic distribution of SNSB finalists to the child population of the US.

INTRODUCTION
The Scripps National Spelling Bee (SNSB) is a high profile, annual academic competition primarily for United States (US) schoolchildren, run by the E.W. Scripps Company (1). To qualify for the SNSB, a speller must attend a school enrolled in the SNSB program and win their final regional spelling bee sponsored by a newspaper, a university, or another community organization. The sponsor typically pays for the speller and one chaperone to attend the week-long SNSB, which is held at the Gaylord National Resort and Convention Center in National Harbor, Maryland (2). There are, however, some geographic inequalities in the regional sponsor system, and Scripps has made recent attempts to address them. For instance, even though there are more than 200 sponsors every year, some regions of the US may not have any sponsors. Starting in 2015, spellers living in a region with no local sponsor could compete in the SNSB through the self-sponsorship program, in which the families of spellers who won their school spelling bee could pay their own way to attend the SNSB (3). Self-sponsorship did provide more spellers an opportunity to compete in the SNSB, however, from 2015—2017, only 30 spellers self-sponsored in all 3 years combined. The low number of self-sponsorships could possibly be attributed to the high cost of self-sponsorship, which, in 2016 cost $3,450 plus travel expenses. Furthermore, self-sponsoring was only available to those living in a region with no local sponsor. Another disparity with the regional sponsor system lies in the unequal size and geographic distribution of sponsored regions across the US. For instance, in 2018, the state of Massachusetts had 4 regional sponsors while the state of Maryland, which has a slightly smaller child population, had 11. Furthermore, the size of individual regions is highly variable—the median number of schools competing in a regional bee is 40, while the largest region has more than 1,000—leading to variability in the competitiveness of regional spelling bees (4). Thus, geographical factors can affect advancement to the SNSB as opposed to merit alone. In 2018, the SNSB introduced a wildcard program known as RSVBee with the intent of giving more spellers the experience of competing at the SNSB and providing more opportunities for students to improve their spelling and vocabulary. RSVBee allowed students who were either previous national finalists or winners of their school spelling bee to apply for a chance to compete as a national finalist, even if they did not win their regional spelling bee. In 2018, over 230 spellers competed through RSVBee, and the champion was an RSVBee speller (5). In 2019, the program continued with a slight increase in financial contribution from the spellers (6). Despite the increase, the number of RSVBee spellers increased to 285, once again giving more spellers the opportunity to compete at the SNSB. As illustrated, through the RSVBee program, the SNSB aimed to increase the number of spellers who could compete at the SNSB.

In this manuscript, we tested the hypothesis that the RSVBee program increased the strength of the correlation between the geographic distribution of national finalists and the child population of the United States. An increase in this correlation might suggest that RSVBee helped the 52 regions
become more appropriately represented at the SNSB. Overall, we were able to show that the RSVBee program improved the correlation between the geographic distribution of SNSB spellers and the child population of the US.

RESULTS

We obtained child population data and the number of spellers at the SNSB for the years 2012 through 2019 for each of the 50 states, the District of Columbia (DC), and Puerto Rico (PR) from publicly available sources. For each year, we compared the percentage of sponsored spellers from a state/territory with the percentage of children residing in each state/territory as a ratio of sponsored spellers to child population, with 1 being the ideal ratio. For the years during which either self-sponsorship (2015–2017) or RSVBee (2018–2019) was available, we also calculated the ratio of the total spellers to child population. We found the Pearson correlation coefficient for each year between sponsored spellers and child population. For years during which self-sponsorship or RSVBee was available, we also found the correlation between the total spellers and the child population. We used the Fisher z-transformation and a paired sample t-test to assess the improvement in correlations attributable to self-sponsorship and RSVBee. We found that the RSVBee program improved the correlation between the geographic distribution of SNSB spellers and the child population of the US.

From 2012 through 2014, there was no form of self-sponsorship or RSVBee, so we only calculated a sponsored speller to child population ratio (Table 1). There was a strong positive correlation between the percentage of sponsored spellers and the percentage of child population by state in 2012 (Figure 1). If spellers were distributed proportionally to the population of children in the 52 regions, the ratios would be approximately 1, and the overall correlation would approach 1. We defined underrepresented states as those with a speller to population ratio less than 0.5 and overrepresented states as those with a speller to population ratio greater than 1.5. In 2012, there were 11 underrepresented regions, 16 overrepresented regions and the Pearson correlation coefficient between sponsored speller percentage and child population percentage was 0.772. In 2013, there were 11 underrepresented regions, 15 overrepresented regions and the Pearson correlation coefficient was 0.769. In 2014, there were 11 underrepresented regions, 15 overrepresented regions and the Pearson correlation coefficient was 0.763. Thus, the geographic representation in the SNSB was relatively stable across the 50 US states, DC, and PR from 2012–2014.

From 2015 through 2017, self-sponsorship was available only for spellers with no local sponsor. We calculated a sponsored speller to child population ratio and a total speller (sponsored plus self-sponsored) to child population ratio, with 1 being the ideal ratio. In 2015, when 6 spellers self-sponsored (3 from CT, 1 each from NJ, NY, and MO), 11 regions were underrepresented and 16 regions were overrepresented. The Pearson correlation coefficient was 0.744 for sponsored spellers to child population and increased minimally to 0.746 for total speller to child population. In 2016, when 8 spellers self-sponsored (2 each from MO and NY, 1 each from NJ, CA, OR, and TN), 10 regions were underrepresented and 14 regions were overrepresented. The correlation of sponsored speller to child population was 0.754 and increased slightly to 0.763 for total speller to child population. In 2017, when 16 spellers self-sponsored (7 from FL, 2 each from NJ and NY, 1 each from CA, OR, IL, IN, and MO), 11 regions were underrepresented and 15 regions were overrepresented. The correlation between sponsored speller and child population was 0.741 and increased to 0.766 for total speller and child population (Figure 2). We Fisher z-transformed the correlations and performed a paired t-test comparing the correlations before and after self-sponsorship in order to test for a statistically significant effect of self-sponsorship on the correlation. The sponsored to population and total to population correlations for 2015–2017 were not statistically significantly different (p-value = 0.219). Thus, the geographic representation across the 50 US states, DC, and PR was relatively stable from 2015–2017, with no statistically significant increase in the correlation between spellers and child population attributable to self-sponsorship.

For both 2018 and 2019, we calculated a sponsored speller to child population ratio (before RSVBee) and a total speller to child population ratio (after RSVBee), with 1 being the ideal ratio (Table 2). In 2018, there were 259 sponsored spellers and 242 RSVBee spellers. Before RSVBee, there

![Figure 1: Scatterplot of percentage of SNSB spellers by state vs percentage of child population by state in 2012 shows a strong positive correlation (Pearson r = 0.772).](image-url)
were 12 underrepresented regions; after RSVBee, 8 of those 12 had appropriate representation (speller to population ratio between 0.5 and 1.5). Before RSVBee, there were 16 overrepresented regions; after RSVBee, 8 of those 16 had appropriate representation. The correlation coefficient between sponsored spellers and child population was 0.747, and the correlation coefficient of total speller to child population was 0.916 (Figure 3). In 2019, there were 256 sponsored spellers and 285 RSVBee spellers from the 52 regions. Before RSVBee, there were 10 underrepresented regions; after RSVBee, 7 of those 10 were appropriately represented (1 became overrepresented). There were 14 overrepresented regions; after RSVBee, 7 of those 14 regions were appropriately represented. The correlation coefficient between sponsored spellers and child population was 0.774 and increased to 0.917 for total speller to child population. We compared the Fisher z-transformed correlations before and after RSVBee with a paired t-test and observed a statistically significant difference (p-value = 0.032) for 2018 and 2019. Thus, we observed a large and statistically significant increase in the correlation between the geographic distribution of national finalists and the child population of the United States attributable to the RSVBee program (Figure 4). With the RSVBee program, more individual states had appropriate representation based on the size of their child population (Figure 5). Overall, these results demonstrate that RSVBee improved the correlation between the geographic distribution of national finalists and the US child population.

**DISCUSSION**

The main findings of our analysis were that the geographic distribution of sponsored spellers from the 50 US states, DC, and PR at the SNSB was relatively stable from 2012–2019, that self-sponsorship in 2015–2017 had minimal impact on the geographic distribution of SNSB spellers, and that the RSVBee program (2018–2019) resulted in a statistically significant change in the geographic distribution of spellers that more closely resembled the geographic distribution of the US child population. The increased strength of correlation reflected that states and regions that were underrepresented by sponsors relative to their child population generally had increased speller to population ratios after counting RSVBee spellers. Similarly, states and regions that were relatively
overrepresented had more appropriate speller to population ratios after RSVBee. The wildcard program implemented in 2018 and 2019 appears to be a robust mechanism for SNSB national finalists to reflect the geographic distribution of school-aged children in the US more accurately.

Academic contests can take one of several approaches to determine qualifiers for a national competition. Some allow a single state-wide winner to compete at the national competition. The SNSB relies on regional sponsors to support qualification to the national finals. Even though the SNSB makes no explicit claim that national finalists should be proportional to child population, this regional system is already reasonably well-matched to the US child population, as evidenced by the strong positive correlations (0.741–0.774) we observed between the percentage of sponsored spellers and percentage of child population by state from 2012–2019. However, inequalities are still present. The two states with the largest percentage of child population are relatively underrepresented by sponsors at the SNSB (Figure 1). With the implementation of the RSVBee wildcard system, some of these inequities are mitigated as most of the under and overrepresented states had more appropriate representation at the SNSB.

Figure 5: Maps of the United States showing 2018 sponsored (panel A), 2018 total (panel B), 2019 sponsored (Panel C) and 2019 total (Panel D) speller/child population ratios for each state in the SNSB. Ratios < 0.5 (relative underrepresentation) are colored red, ratios between 0.5-1.5 (appropriate) are yellow, and ratios > 1.5 (relative overrepresentation) are colored green.

In conclusion, the RSVBee wildcard program positively impacted the SNSB by allowing more spellers to compete in the SNSB and significantly improving the correlation between geographic location of national finalists and child population of the United States. In 2018, 242 spellers came to the SNSB through RSVBee. In 2019, there was more awareness for RSVBee, and 285 spellers competed through RSVBee. In 2020, plans were underway to have a reduced number (about 140) of RSVBee spellers limited to 7th and 8th graders, with a $7.50 application fee and up to 18 financial aid packages (which would include the $1,500 participation fee, six-night hotel, and $1,000 for travel and meals) (8). The 2020 SNSB was canceled due to the COVID-19 pandemic (9).

**MATERIALS AND METHODS**

**Child Population Data**

We collected child population data from the 50 US states, the District of Columbia, and Puerto Rico from the United States Census Bureau, Population Division (datacenter.kidscount.org) (10). We added the data for children ages 5–11 and 12–14 to determine the number of children aged 5–14 years in each state/territory. We used the population data from the previous year for each corresponding SNSB (for example, 2011 population data for the 2012 SNSB) to estimate the number of children ages 6–15 years during the SNSB. This spans the ages of the youngest (6 years) and oldest (15 years) spellers at the SNSB to date. For example, for New Jersey (NJ) from the 2011 population data, the number of children from the age 5–11 cohort was 792,356 and the number of children from the age 12–14 cohort was 350,977. We added these numbers to get 1,143,333 children ages 5–14. We added the number of children ages 5–14 from each of the 52 regions to find the total number of children ages 5–14 in the 52 regions combined. When using the 2011 population data, this number was 41,534,282. We then calculated a percentage for each region by dividing the number of children from each region ages 5–14 by the total number of children ages 5–14. The population percentage from the 2011 population data (used for the 2012 SNSB) for NJ was 1,143,333/41,534,282 = 2.75%.
SNSB Data

We used the SNSB round results from each year to find the number of spellers from each state at the SNSB. The results from 2019 are available online (11). Results from 2012 through 2018 are available using the internet archive: Wayback Machine (12–18). For each year, we counted the number of sponsored spellers and, if present, self-sponsored spellers or RSVBee spellers from each region. The sponsored spellers could be distinguished from self-sponsored and RSVBee spellers because the name of the sponsor was listed for each speller—self-sponsored and RSVBee spellers have their school listed as their sponsor. In cases where it was not clear whether a speller was sponsored, we used open source intelligence to collect information regarding regional spelling bee winners to identify sponsored spellers. In cases where a sponsor covered more than one state or sponsored spellers from a state other than where the sponsor was based, the speller was counted as being from the state in which they attend school.

Analysis

For each year, we calculated the percentage of sponsored spellers from each region by dividing the number of sponsored spellers from each region by the total number of sponsored spellers. For example, in 2012, there were 7 sponsored spellers from NJ and 264 sponsored spellers from the 52 regions. The sponsored speller percentage for NJ was 7/264 = 2.65%. In NJ for the 2012 SNSB, the sponsored speller to population ratio was 2.65%/2.75% = 0.96. For the years 2015–2017, in addition to finding the sponsored speller to population ratio, we also added the number of self-sponsored spellers to the number of sponsored spellers to find the new total number of spellers. From there, we calculated a new percentage by dividing the total number of spellers (with any self-sponsored spellers added) by the new total. We then found the total speller to population ratio. For the years 2018 and 2019, in addition to finding the sponsored speller to population ratio, we added the number of RSVBee spellers from each region to the number of sponsored spellers from each region. We calculated a new percentage of spellers from each region by dividing the new number of spellers from each region by the new total number of spellers. We then found the total speller to population ratio. For 2012–2014, we found the correlation between sponsored spellers and population. For 2015–2017, we found the correlation between sponsored spellers and population and the correlation between the total spellers (sponsored plus self-sponsored) and population. For 2018 and 2019, we found the correlation between sponsored spellers and population and the correlation between the total spellers (sponsored plus RSVBee) and population. We used a Fisher z-transformation on the correlation coefficients and a paired sample t-test to compare sponsored spellers/population to total spellers/population (2015–2017) to assess the effect of self-sponsorship and sponsored spellers/population to total spellers/population (2018–2019) to assess the effect of RSVBee. A two-tailed p-value of < 0.05 was considered statistically significant. Calculations and correlations were performed using Microsoft Excel 2019. Fisher z-transformation and paired t-test were performed using SPSS for Windows version 26.

Table 1: Child population aged 5-14 years (2011), 2012 spellers, and speller to child population ratios for 2012–2014 at the Scripps National Spelling Bee (SNSB) by US region. Column 2 shows the percentage of US children aged 5–14 years living in a specific state or territory in 2011. Column 3 shows the percentage of spellers from each state or territory at the 2012 SNSB. Column 4 shows the ratio of 2012 speller percentage to the 2011 child population percentage (column 3 divided by column 2) for each state or territory for 2012. Columns 5 and 6 show the speller to population ratios for 2013 and 2014 respectively.
REFERENCES


4. Communication with the E.W. Scripps Company.


ACKNOWLEDGEMENTS
“We would like to thank Ann Vessey Kannankeril and Sarah Kannankeril for their continual support. Above all, we would like to thank our Lord and Savior Jesus Christ.”

Copyright: © 2020 Kannankeril and Kannankeril. All JEI articles are distributed under the attribution non-commercial, no derivative license (http://creativecommons.org/licenses/by-nc-nd/3.0/). This means that anyone is free to share, copy and distribute an unaltered article for non-commercial purposes provided the original author and source is credited.