NAMES, Vol. 58 No. 1, March, 2010, 47–54

Racial and Gender Differences in Diversity of First Names

HERBERT BARRY III
University of Pittsburgh, USA

AYLENE S HARPER
Community College of Allegheny College, USA

First names and the number of individuals given each name were recorded on three populations, born in Pennsylvania in 1990, 1995, and 2000. Four categories of people were "black" females, "white" females, "black" males, and "white" males. Diversity of the name choices was measured by the 50 percent rank frequency, which divides the population equally between the individuals with more frequent and less frequent names. Diversity was highest for "black" females and lowest for "white" males. For all four categories of people, diversity increased progressively from 1990 to 1995 and from 1995 to 2000. Feminine phonetic attributes also were most frequent for "black" females and least frequent for "white" males. Racial divergence in choices of names was indicated by a smaller number of names in 2000 than in 1990 that were among the fifty most frequent for both "blacks" and "whites."

KEYWORDS First names, name diversity, name frequency, year born, racial differences, gender differences, Pennsylvania names

Many people choose a conventional or currently frequent first name for a baby. The advantage is that the named person is protected from ridicule or confusion concerning the name. The disadvantage is that the name fails to fulfill its basic purpose, to differentiate the named person from other people. Some people choose an unusual first name for a baby. The advantage is that the name is distinctive and memorable. The disadvantage is that it deviates from the cultural norm and therefore may be disliked by the named person and by other people.

In England and in English-speaking America, conventional first names have previously predominated. Prior to the twentieth century, the most frequent name continually was *Mary* for girls, *John* for boys. Most people were given these names or a small number of other frequent names.

 DOI 10.1179/175622710X12590782368107

Choices of first names have become more diverse. Tucker (2009) demonstrated progressively greater dispersion in choices of names in the United States during the latter part of the twentieth century and in the first few years of the twenty-first century. Diversity also is greater in choices of female than male names. Tucker emphasized that the proportion of the population given the most frequent name has decreased. Lieberson and Mikelson (1995) and Fryer and Levitt (2004) demonstrated that the proportion of the population given a unique name has increased, especially among "blacks" and females.

The most frequent name and unique names are limited to the opposite extreme diversities of name choices. A number of the most frequent names, such as 25 (Barry & Harper, 1995), 100 (Barry & Harper, 2003), or 500 (Barry & Harper, 2000), can measure diversity among frequent names but omit variations in diversity among the more numerous names that are seldom chosen. A measure described by Barry and Harper (2005; 2006) is the 50 percent rank diversity. It divides the population equally between the more frequent and less frequent names. It is therefore an equivalent measure whether a small or large number of the most frequent names are sufficient to include half of the population.

Names are given rank order frequencies, beginning with 1 for the most frequent. The cumulative number of individuals can be identified for any successive rank. The 50 percent rank diversity is the rank number at which the cumulative number of individuals is half the total number of individuals in the population. A higher 50 percent rank frequency number signifies greater diversity.

The present paper applies the 50 percent rank diversity to four categories of people: "black" females, "black" males, "white" females, and "white" males. Each category is divided into three populations, born in 1990, 1995, and 2000.

Barry and Harper (2005; 2006) reported greater diversity of names of females and males in the United States born from 1990 to 1999 than from 1900 to 1909. Diversity of names also is greater for "blacks" than "whites" (Eagleson & Clifford, 1946; Lieberson & Mickelson, 1995). Differentiation from names of "whites" has been more frequent for names given to "black" females than for names given to "black" males (Evans, 1989).

Many of the Africans who were transported to America as slaves were given names that were frequently given to "whites." In recent years, many "blacks" have emphasized their racial heritage by choosing names that are given frequently to "blacks" but rarely to "whites" (Lieberson & Mikelson, 1995; Fryer & Levitt, 2004). Frequent names for "blacks" therefore have become more often differentiated from frequent names for "whites." Lists of names given predominantly to "blacks" are included in an article by Black (1996) and in books by Lieberson (2000) and by Levitt and Dubner (2006).

Sample and methods

Birth certificates in Pennsylvania were used because of the availability of first-name frequencies. The year 1990 is the first that distinguished between "black" and "white" mothers. Changes in choices of names were determined by recording births also in two subsequent years, 1995 and 2000. Children born in the three years constituted separate populations. The rank order frequencies were determined separately for four

categories of people: "black" females, "white" females, "black" males, and "white" males.

Frequencies of names were obtained in the form of magnetic records that contain the information compiled by the Pennsylvania State Health Data Center. The original information consists of the first name recorded on the birth certificate for each male and female born alive in Pennsylvania in 1990, 1995, and 2000. The mothers were recorded as "black" or "white." The word "baby" is excluded from the numbers and data analyses because it signifies that no name was given.

Pennsylvania has nationally representative attributes. The state contains one of the nation's most populous cities, Philadelphia, and many smaller cities, municipalities, and rural communities. The percentage of "blacks" in Pennsylvania is similar to the percentage in the United States. Among all the births in Pennsylvania in 1992, "blacks" constituted 15 percent and other "non-whites" 2 percent (Pennsylvania, 1992). "Whites" include Hispanics because there was no separate category for that ethnic group.

The 50 percent rank diversity (Barry & Harper, 2005; 2006) is the rank order number for frequency of names that divides the population equally between the number of individuals with more frequent names and less frequent names. A higher number signifies that more names are included among 50 percent of the individuals who have the most frequent names. If the most frequent name is given to more than half of the individuals, the 50 percent diversity is 1. If each individual is given a unique name, all the names share the same rank frequency and the 50 percent rank diversity number is half of the total number of individuals.

The 50 percent rank diversity can be applied to other rank frequencies of choices in a population of individuals. In addition to surnames, the population can be divided into equal numbers with higher and lower rank frequencies for other variables, such as days of vacation or travel per year, annual income, and value of real estate. A different technique for dividing a population into two equal halves is the pharmacological measure of LD_{50} or ED_{50}. It is the minimum lethal dose (LD) or minimum therapeutically effective dose (ED) of a drug. A small number of higher and lower doses in a small number of individuals can determine the LD_{50} or ED_{50}.

The information on the names was recorded in an electronic file (Barry, 1995). The SPSS statistical package (SPSS, 1986) was used for summaries and statistical analyses of the information.

Fifty percent rank diversity

The 50 percent rank diversity is the rank frequency number that divides all the individuals as equally as possible between higher and lower rank frequencies. Table 1 shows the 50 percent rank diversity for three populations, consisting of birth in 1990, 1995, or 2000, separately for "black" females, "white" females, "black" males, and "white" males. Table 1 also shows the number of named individuals in each of the twelve populations.

The highest 50 percent rank diversity was 804 for "black" females in 2000. Rank frequency 804 contained 521 names, ranks 586–1106, each given to two individuals, and therefore given to 1042 individuals. There were 4477 individuals whose names were shared with two or more other "black" females, ranks 1–585, and 4313 indi-

TABLE 1
DIVERSITIES OF NAMES AND NUMBERS OF INDIVIDUALS

Year	Female Names		Male Names	
	"Black"	"White"	"Black"	"White"
	50% rank diversity			
1990	392	48	117	29
1995	601	73	218	35
2000	804	95	308	48
	Total number of individuals			
1990	12,592	71,223	12,728	75,539
1995	10,449	62,629	10,869	66,132
2000	9832	58,437	10,245	60,958

viduals, beginning with rank 1107, whose names were given to no other "black" female in Pennsylvania in 2000. The lowest 50 percent rank diversity was 29 for "white" males born in 1990. The name with rank 29 was Jacob, given to 740 individuals. There were 37,069 individuals with more frequent names, ranks 1–28, and 37,730 individuals with less frequent names, beginning with rank 30.

The 50 percent rank diversity was greater for "black" males than for "white" females. Differences in 50 percent rank diversity were greater between the racial categories than between the genders. For all four categories identified in Table 1, increases in 50 percent rank diversity were consistent from 1990 to 1995 and from 1995 to 2000. From 1990 to 2000, the 50 percent rank diversity almost tripled for "black" males, more than doubled for "black" females, almost doubled for "white" females, and less than doubled for "white" males.

Names with rank frequencies 1 to 10

Table 2 lists the names with the 10 highest rank frequencies given to "black" females, "white" females, "black" males, and "white" males, born in 1990 and in 2000. The comparison between the two years omits the intermediate year, 1995. The first name in each column had rank frequency 1. Two names shared rank frequency 10 for "black" males in 1990 and for "white" females in 2000, resulting in 11 instead of 10 names.

A comparison between 1990 and 2000 shows more continuation of rank frequencies between 1 and 10 for male than female names. There were 6 names of "black" males and 6 names of "white" males in 1990 that continued to be listed in 2000. There were 2 names of "black" females and 3 names of "white" females that continued to be listed in 2000.

Several names had rank frequencies between 1 and 10 for both "black" and "white" females and for both "black" and "white" males in the same year. For female names, they were *Ashley*, *Brittany*, and *Jessica* in 1990, *Taylor* and *Alexis* in 2000. For male names, they were *Michael*, *Christopher*, *David*, and *Joshua* in 1990, *Joshua* and *Michael* in 2000.

TABLE 2
THE TEN MOST FREQUENT NAMES

Female names		Male names	
"Blacks"	"Whites"	"Blacks"	"Whites"
	Births in 1990		
Jasmine	Amanda	Michael	Michael
Ashley	Jessica	Anthony	Matthew
Brittany	Ashley	Christopher	Christopher
Tiffany	Brittany	Brandon	Joshua
Amber	Samantha	James	Andrew
Jessica	Sarah	Robert	Ryan
Courtney	Lauren	William	Joseph
Danielle	Megan	Kevin	John
Dominique	Nicole	Eric	Daniel
Nicole	Emily	David Joshua	David
	Births in 2000		
Kayla	Emily	Isaiah	Jacob
Jada	Madison	Joshua	Michael
Destiny	Hannah	Nasir	Matthew
Brianna	Sarah	Michael	Nicholas
Taylor	Alexis	Anthony	Ryan
Jasmine	Samantha	Jordan	Tyler
Diamond	Olivia	Christopher	Joseph
Alexis	Lauren	James	Zachary
Imani	Alyssa	Elijah	John
Ashley	Abigail Taylor	David	Joshua

Names with rank frequencies 1 to 50

Rank frequencies 1 to 50 were calculated for each of the four categories of people and for all three years of birth. The large number of names enables more precise comparisons among the categories of people and between births in 1990 and births in 2000. The rank frequencies 1–50 equate the relative frequencies of the smaller numbers of names chosen for "blacks" than for "whites."

The percentage of names with rank frequencies 1 to 50 in only one of the two years of birth, 1990 or 2000, was 85 percent for "black" females, 56 percent for "white" females, 54 percent for "black" males, and 45 percent for "white" males. The contrast between 85 percent for "black" females and 46 percent for "white" males resembles the contrast in 50 percent rank diversity, shown in Table 1.

Most of the names with rank frequencies of 1 to 50 were chosen exclusively for females or males. The only exceptions were *Dominique* in 1990 and *Jordan* in 2000, for "black" females and "black" males.

Table 3 shows other differences among the names with rank frequencies 1 to 50 given to "black" females, "white" females, "black" males, and "white" males in any of the three years of birth, 1990, 1995, and 2000. The total number of names is increased when names with rank frequencies 1 to 50 in 1990 are replaced by new names in 1995 and in 2000. In accordance with the 50 percent rank diversity in Table 1, the most names were given to "black" females and the fewest names to "white" males.

Barry and Harper (1995) reported a quantitative phonetic femininity index, combining measures of the final letter and structure of the name. The scores can range from −4.00 to +4.00. Tables 3 shows that the average score was higher for "black" than "white" females and also higher for "black" than "white" males. One of the components of the femininity index is whether or not the spoken stress is after the first syllable. Table 3 shows that the stress was after the first syllable most often for names given to "black" females and least often for names given to "white" males.

Table 3 includes two linguistic origins of the names, obtained from several sources. Most of them are the same as in a book by Evans (2006). Hebrew was most often the origin of names given to "white" males. Greek was most often the origin of names given to "white" females.

Frequent names for both "blacks" and "whites"

The names with rank frequency between 1 and 50 for female or male "blacks" or "whites" included 30 female names and 37 male names for both racial categories in either 1990 or 2000. Among the 30 female names, 57 percent were included for both racial categories in 1990, 30 percent in 2000, and 13 percent in both years. Among the 37 male names, 27 percent were included for both racial categories in 1990, 16 percent in 2000, and 57 percent in both years. The principal difference between the genders was a much higher percentage of male than female names included in both years. For both genders, more names were included for both racial categories in 1990 than in 2000.

The change from 1990 to 2000 in proportion of names that were frequently chosen for both "blacks" and "whites" is a more precise comparison between the two years. Among the female names, 19 were more frequent and 9 less frequent for both "blacks" and "whites" in 1990 than in 2000. Among the male names, 22 were more frequent and 5 were less frequent for both "blacks" and "whites" in 1990 than in 2000. Racial divergence of the frequent female and male names therefore increased from 1990 to 2000.

TABLE 3
FIFTY MOST FREQUENT NAMES FOR "BLACKS" AND "WHITES"

	Female Names		Male Names	
	"Black"	"White"	"Black"	"White"
Total names	99	77	74	65
Average femininity index	+2.34	+1.48	+0.04	−0.48
% stress after first syllable	57	34	38	17
Linguistic origin				
% Hebrew	7	14	22	32
% Greek	13	26	12	18

The minority of names were more frequent for both racial categories in 2000 than in 1990. They were 9 female names, *Alexis*, *Alyssa*, *Brianna*, *Destiny*, *Gabrielle*, *Morgan*, *Taylor*, *Sydney*, and *Victoria*. They were 5 male names, *Alexander*, *Cameron*, *Christian*, *Jordan*, and *Tyler*.

Discussion

The 50 percent rank diversity is a single number that divides the population into two equal halves. This measure therefore is applicable whether the name choices are highly concentrated on a small number of the most frequent names or are distributed among many names. Table 1 shows large differences in 50 percent rank diversity among the four categories of people and among the populations born in 1990, 1995, and 2000.

Names given to African Americans have two major sources, frequent names of "white" European descendants and distinctively "black" names. The two sources contribute to the greater diversity of names of "blacks" than "whites." A more effective source of greater diversity is a tremendous variety of names that are chosen only for "black" females. The 50 percent rank diversity of 805 for "black" females in 2000 contains names that were given to only two individuals. The population thereby was divided approximately equally between names given to three or more individuals and names given to a single individual.

The most novel finding is the consistent increase in diversity of first names from births in 1990 to 1995 and from births in 1995 to 2000, for all four categories of people. The rapidly increasing diversity of first names is part of a cultural pattern of more diverse choices for many other objects and activities. Advances in technological capabilities have enabled people to satisfy individual preferences. Large grocery stores contain many brands of the same foods. Television sets offer choices among dozens of channels with different programs. Many movie theatres contain multiple screens showing different films concurrently. The Internet offers a vast variety of discussion groups and other opportunities for information and interaction. Cell phones enable conversations among people at any time, in any location.

Divergence between names given to "blacks" and "whites" is indicated by fewer female names that were predominantly given to one racial category in 2000 than in 1990. Contrary to this trend, several names were chosen more often in 2000 than in 1990 for both "blacks" and "whites." Further increase in frequency for both "blacks" and "whites" after 2000 would contribute to racial convergence instead of divergence of names.

Acknowledgments

The data on frequencies of first names were supplied by the State Health Data Center, Pennsylvania Department of Health, Harrisburg, Pennsylvania. The Department specifically disclaims responsibility for any analyses, interpretations, or conclusions.

Helpful services were received from the Pennsylvania State Health Data Center, in particular by Jerry Orris, and from the University of Pittsburgh Computer Center, in particular data transfer and programming by Eileen S. Kopchik. Helene Borke, PhD, contributed helpful criticisms of earlier versions of this article.

Partial versions of this paper were reported at annual meetings of the American Name Society, in New York City, on December 30 1992, and in Chicago, Illinois, on January 3 2008.

Bibliography

Barry, H. III. 1995. Computers and Research on Personal Names. *Names*, 43: 315–24.

Barry, H. III, & Harper, A.S. 1995. Increased Choice of Female Phonetic Attributes in First Names. *Sex Roles*, 32: 809–19.

Barry, H. III, & Harper, A.S. 2000. Three Last Letters Identify Most Female Names. *Psychological Reports*, 87: 48–54.

Barry, H. III, & Harper, A.S. 2003. Final Letter Compared with Final Phoneme in Male and Female Names. *Names*, 51:13–33.

Barry, H. III, & Harper, A.S. 2005. The Majority of Female First Names Ended in *A* or *E* Throughout the Twentieth Century. *Gender Roles*. Ed. by J. W. Lee. Hauppauge, NY: Nova Science, pp. 117–43.

Barry, H. III, & Harper, A.S. 2006. The Majority of Female First Names Ended in *A* or *E* Throughout the Twentieth Century. *Gender Identity, Psychology and Life Style*. Ed. by A.J. Lauber. Hauppauge, NY: Nova Science, pp. 91–116.

Black, K. 1996. Afro-American Personal Naming Traditions. *Names*, 44: 105–25.

Eagleson, O.W. & Clifford, A.D. 1945. A Comparative Study of the Names of "White" and Negro Women College Students. *Journal of Social Psychology*, 21: 57–64.

Evans, C.K. 1989. Adam and Andre, Lindsay and Lakeisha: Racial Differences in First Names. 1987. *Bulletin of the North Central Name Society*, Spring: 43–63.

Evans, C.K. 2006. *The Great Big Book of Baby Names: A Complete Guide from A to Z*. Lincolnwood, IL: Publications International Ltd.

Fryer, R.G. Jr, & Levitt, S.D. 2004. The Causes and Consequences of Distinctively Black Names. *Quarterly Journal of Economics*, 119: 767–805.

Levitt, S.D. & Dubner, S. J. 2006. *Freakonomics* (revised and expanded). New York: HarperCollins.

Lieberson, S. 2000. *A Matter of Taste; How Names, Fashions, and Culture Change*. New Haven, CT: Yale University Press.

Lieberson, S. & Mikelson, K.S. 1995. Distinctive African American Names: an Experimental, Historical, and Linguistic Analysis of Innovation. *American Sociological Review*, 60: 928–46.

Pennsylvania 1992. *Pennsylvania Vital Statistics Annual Report*, 1990. Harrisburg: Pennsylvania State Health Data Center.

SPSS Inc. 1986. *SPSSX User's Guide*. 2nd edn. Chicago: SPSS.

Tucker, D.K. 2009. Increased Competition and Reduced Popularity: US Given Name Trends of the Twentieth and Early Twenty-First Centuries. *Names*, 57: 52–62.

Notes on contributors

Herbert Barry III, PhD, Professor Emeritus at the University of Pittsburgh since 2001, was a member of the American Name Society Executive Council 2000–2003. In addition to his research with Aylene S. Harper on unisex and frequent first names, he has published articles on names of fictional characters in novels.

Aylene S. Harper, PhD, is Professor of Psychology at Community College of Allegheny County, South Campus, West Mifflin, Pennsylvania, 15122. Topics of her articles with Herbert Barry III, beginning in 1982, include unisex first names and differences between frequent first names of males and females.

Correspondence to: Herbert Barry III, 552 North Neville St, Apt 83, Pittsburgh, PA 15213-2830, USA. Email: barryh@pitt.edu