Establishing the Denominator: The Challenges of Measuring Multiracial, Hispanic, and Native American Populations

Wendy D. Roth[*]
Associate Professor of Sociology
University of British Columbia

Abstract: The norms for how multiracial, Hispanic, and Native American populations self-identify are less well established than for other population groups in the United States. There is considerable variation and fluidity in how these groups identify their race and ethnicity, as well as how they are classified by others. This presents challenges for estimating and analyzing these populations based on their self-identification alone. I argue that self-identification should be treated as a numerator, but the challenge is for researchers to establish the denominator – the population that *could* identify as members of the group based on their ancestry. Considering how many people who could identify with these groups choose to do so sheds light on assimilation and emerging racial classification processes. Analyses of the larger potential populations further avoid bias stemming from non-random patterns of self-identification with the groups.

---

[*] Contact information: Wendy D. Roth, University of British Columbia, Sociology Department, 6303 N.W. Marine Drive, Vancouver, BC V6T 1Z1, Canada; wendy.roth@ubc.ca; 604-822-4845

Ethno-racial mixture is on the rise in the United States. This is occurring not only because of a growing population of children born from interracial unions (U.S. Census Bureau 2010), but also because of the growth of the Hispanic population (Krogstad and Lopez 2015), where racial mixture is so common over generations that relatively few consider themselves not to have any mixed race origins. Native Americans are another population where ethno-racial mixing is so common historically that large portions of the population consider themselves racially mixed (Snipp 1997). The Office of Management and Budget's 1997 decision to allow people to mark more than one race was an important recognition of the ways that population trends and ethno-racial classification norms are changing (U.S. Office of Management and Budget 1997; Wallman, Evinger, and Schechter 2000). These changes reflect an understanding that people with multiple racial ancestries often identify with more than one of them, but they also reveal that the norms for how these populations self-identify are less well established than for those who consider themselves solely Black or White. There is considerable variation in how these groups identify their race and ethnicity, as well as how they are classified by others (Alba, Insolera, and Lindeman 2016; Liebler et al. 2014; Porter, Liebler, and Noon 2016; Roth 2010, 2016). This presents challenges for estimating and analyzing these populations based on their ethno-racial self-identification alone.

Estimates of Multiracial, Hispanic, and Native American populations miss people who *could* identify in these ways

Current estimates of ethnic and racial groups from the U.S. Census are based on individuals' ethno-racial self-identification or classification by a member of their household who fills out the census form, presumably a relative or someone who knows them well. Yet relative to

White, Black, and Asian populations,[1] there is considerable fluidity in how multiracial, Hispanic, and Native American populations self-identify. In their comparison of individually-linked responses in the 2000 and 2010 censuses, Liebler and colleagues (2014) found that 97 percent of non-Hispanic Whites, 94 percent of non-Hispanic Blacks, and 91 percent of non-Hispanic Asians had the same self-identification in both censuses, but all other groups had much greater fluidity in responses. Only 53 percent of American Indians or Alaska Natives and 48 percent of Native Hawaiians or Pacific Islanders were consistent in both censuses. Similarly, based on calculations of their data, we see that only 23 percent of all non-Hispanic multiracials and 41 percent of all Hispanics listed the same race and Hispanic origin responses in 2000 and 2010.[2]

These findings are consistent with a number of other studies that show much greater fluidity in racial self-classification for multiracial, Hispanic, and Native American individuals, and greater stability in self-classification for non-Hispanic Whites, Blacks, and Asians (Doyle and Kao 2007; Liebler and Ortyl 2014; Saperstein and Penner 2012). Furthermore, this fluidity is seen not only over time, but also across contexts. Harris and Sim (2002) found that 12.4 percent of the non-Hispanic adolescents in the first wave of the National Longitudinal Survey of Adolescent Health (Add Health) survey gave different racial identifications in an in-school questionnaire and an in-home interview; most of these adolescents revealed multiracial origins in at least one of their responses. Although Harris and Sim's study focused only on non-Hispanic adolescents, when both Hispanic and non-Hispanic adolescents were included, about 15 percent answered the race question differently across these contexts (Brown, Hitlin, and Elder 2006).

In general, Hispanics reveal greater fluidity in their racial identification than in their identification as Hispanic. Although only 41 percent of Hispanics answered the race and Hispanic origin questions in the same way in 2000 and 2010, 87 percent of Hispanics consistently identified

as Hispanic (of any race) at both times (Liebler et al. 2014). However, there seems to be greater fluidity in Hispanic identification among younger people and those who are less immersed in a Hispanic community. Among adolescents who identified as Hispanic in the Add Health in-school survey, 80 percent identified as Hispanic in the in-home interview as well (Brown, Hitlin, and Elder 2006). In a study of Hispanic high school students from 1980-1982, 68 percent consistently identified as Hispanic at both times, and inconsistent identifications were more likely for those who were monolingual English speakers or attended schools with smaller percentages of Hispanic students (Eschbach and Gómez 1998).

There is also greater inconsistency between self-identity and observation by others for the multiracial, Hispanic and Native American populations (Roth 2016). Several studies in the health fields examined inconsistencies between individuals' self-reports and the observations of others in the form of medical records, interviewer classifications, or death certificates (Hahn, Truman, and Barker 1996; Kressin et al. 2003; West et al. 2005). These studies found the highest rates of consistency among self-reported Whites (91–98 percent) and Blacks (90–99 percent), and typically moderate-to-high rates among Asians (76–95 percent). Rates of consistent observations by others were lower for self-reported Hispanics (64–83 percent), and extremely low for Native Americans (0–23 percent). Studies of people who self-identify as multiracial also find relatively low rates of consistency in observers' classifications of their race. Observers tend to classify multiracial individuals as monoracial (Herman 2010). More observers classify those with mixed Black backgrounds as Black, while there is greater diversity in how they classify those with any other racial mixtures (Feliciano 2016).

What is particularly notable is that inconsistencies persist even when the observer knows the person they are classifying. In one study, people who previously self-identified their ancestry

but died before a follow-up study were identified by both a proxy – a next of kin or a nonrelative who knew them – and by funeral directors. Only 20 percent of those who had self-identified as Native American were classified as such by the proxies who knew them (and none of them were classified as Native American by the funeral directors) (Hahn, Truman, and Barker 1996). Porter, Liebler, and Noon (2016) also examined racial classifications by proxies in the individually-linked data from the 2000 and 2010 censuses. They examined cases that had a proxy report in one decennial census (usually from a neighbor) and a report from the individual or someone in his/her household in the other census. They found high consistency between the household reports and proxy reports for Whites (98 percent), Blacks (94 percent) and Asians (88 percent). They also found high consistency for a Hispanic designation, with 86 percent of household reports of Hispanic identity confirmed by proxies. However, the proxy reports matched household reports at lower rates for Native Americans (62 percent), Pacific Islanders (62 percent), and multiracial people (8–36 percent, depending on the mix of races reported). About half of Hispanics were classified as a different race by proxies than that reported by their households.

Although the proxies in the Porter et al. (2016) study typically had some knowledge of the individual, their classifications were guided by external cues, such as the racial and ethnic composition and class composition of the neighborhood and the person's age. Proxy classifications were more likely to match household reports if the person was living in an area where many others report the same race. This suggests that observers relied on the racial composition of the area to guide their assessments. Proxies more frequently reported children as multiple-race and adults as single-race, which may reflect a common perception that younger people are more likely to identify themselves as multiracial, or may stem from a greater tendency to see these young people with their parents and to draw upon this information in assigning their race.

The greater fluidity of multiracial, Hispanic, and Native American self-identification, and greater inconsistency between their self-identification and observation by others, indicates that norms of classification are less well formed for these groups than they are for Whites, Blacks, and Asians. These racial and ethnic designations are relatively flexible, and their salience is more dependent on the situational context. This means that estimates of multiracial, Hispanic, and Native American populations in the Census, American Community Survey (ACS), and many other surveys will miss people who *could* identify as members of these populations and may in some circumstances, but do not at the moment of data collection. Furthermore, when the individual in question is not the one filling out the household report for the census or ACS, there is likely to be greater discrepancy between what the household member reports and the individual's self-identification than for groups with more established norms of racial classification, such as Whites, Blacks, and Asians. Unfortunately, information about whether it is the individual himself or herself who is answering the questions or another member of the household is not available in ACS or Census data.

Analyzing everyone who *could* identify in these populations sheds light on assimilation and emerging racial classification processes

Noting the fluidity in multiracial adolescents' identifications, Harris and Sim argued that the 2000 census provided "a count of *a* multiracial population, not *the* multiracial population" (2002, 625; see also Morning and Saperstein, this volume). The same could be said for Hispanics and Native Americans, reflecting patterns of widespread interracial mixing across many generations. In this sense, determining a population based on ethno-racial self-identification needs to recognize that individuals self-select into the group. One of the most important issues for

scholars of race and immigration then becomes the ability to detect everyone who *could* identify in these populations, whether or not they do so. Analyzing these patterns can reveal how racial classification processes unfold and what identities people choose under what circumstances. In the case of Hispanics, this type of analysis can reveal which individuals have assimilated into the American mainstream by dropping a Hispanic identity.

Scholars have attempted various strategies to consider everyone who could potentially identify in these populations by garnering information on the ethno-racial identity or origins of their parents and grandparents. One approach is to focus on children from interracial marriages by aggregating census data to the household level and compiling information about the race and ethnicity of the parents and the children. Several studies have used this approach, focusing on how these children who could potentially be labelled as multiracial are in fact racially classified on the census (Brunsma 2005; Kana'iaupuni and Liebler 2005; Liebler 2004, 2010; Qian 2004; Roth 2005; Xie and Goyette 1997).[3] In general, this research shows that many of these children are not classified as part of the two or more races population, despite having parents of two or more races. In 2000, only 43 percent of children with parents of different single races were identified as multiracial, while the remainder were identified as the race of one of their parents (Roth 2005).[4]

There are limitations to this approach due to the survey structure and available information in the census, however. It can only focus on interracial couples where one member is 'Person 1', and it focuses on children (rather than the adult multiracial population) because the only way to get information about the race of individuals and their parents is for them to be living in the same household. This approach also has to rely on assumptions about the biological relatedness of the child and parents. The distinction made beginning in the 2000 census between 'natural-born son/daughter' and 'adopted son/daughter' in the question on each person's relationship to Person

1 is extremely valuable to this type of analysis. But researchers still have to assume that the child is also the natural-born child of the spouse or partner of Person 1. These studies typically focus only on married households for this reason, which limits our knowledge of the potential multiracial population further and yet still draws on assumptions about relatedness rather than data confirming it.

These studies are also limited by the lack of information on the person filling out the form, as well as valuable information about the context of survey completion. We cannot tell who in the household is completing the form, although this is likely to be an important aspect of how the child is classified. The parent completing the form may be more likely to emphasize his or her own racial background in the child's designation (Roth 2005). Analysis of Current Population Survey (CPS) data on Mexican Americans showed that when children's Mexican ancestry derives from the mother's side of the family, the children are more likely to be identified as Mexican if their mother completes the survey form (Duncan and Trejo 2011).

Another aspect of the context which we are currently unable to evaluate is whether the survey is self-completed and mailed in or completed by an enumerator or with an enumerator present. Since some adolescents will identify their race differently when completing the questions on their own or in an in-home interview (Harris and Sim 2002), we might expect this context to influence some household members' classification of their multiracial children as well. And since the race of the interviewer or observer can influence how they perceive the respondent (Feliciano 2016; Hill 2002), it is possible that there could be race-of-enumerator effects in how multiracial children are classified in cases where an enumerator is present or helping to complete the form. These types of contextual factors are likely to matter much more in the racial classification of ethno-racially mixed populations than for Whites, Blacks, or Asians.

Another approach to analyzing all those who potentially *could* identify as multiracial, Hispanic, or Native American focuses on ancestry data. By identifying people who list ancestries indicating multiple races, or Hispanic or Native American origins, scholars can analyze which individuals have chosen to identify with these groups or to simplify their ethno-racial identities in ways that shed light on broader processes of assimilation and racial boundary formation. Gullickson and Morning (2011), for example, used the census ancestry question to identify adults who report having ancestry in more than one racial group. They found that individuals with part-Asian-ancestry were the most likely to identify themselves as multiracial in 2000, and for those with no Asian ancestry, part-Black ancestry groups were more likely to identify themselves as Black rather than multiracial. Those with White and Native American ancestry were much more likely to identify their race as White than as any other option. Similarly, Emeka and Vallejo (2011) used the ancestry question on the ACS to identify adults with Latin American ancestry, and found that 6 percent of them did not identify as Hispanic. A non-Hispanic identification was most common among those who listed Latin American and non-Latin American ancestries, suggesting intermarriage in earlier generations, as well as those who speak only English. These studies attempt to portray the individuals' self-identification but are again limited by the lack of information over who is completing the survey. Gullickson and Morning (2011) restricted their analysis to the household head on the assumption that this is most likely the person completing the form, but it is not clear that the household head is always the person answering the questions.

To study ethnic attrition among Mexican-Americans, Duncan and Trejo (2011) took advantage of the question on parents' place of birth in the CPS. They examined whether the children of Mexican-born parents identified themselves as Mexican and found that 90 percent of U.S.-born individuals aged 25-59 with a Mexican-born parent did identify as Mexican on the

Hispanic origin question. However, those who did not self-identify as Mexican had more education than those who did, showing that estimates of socioeconomic status that rely on self-identification will underestimate the achievements of people of Mexican descent. Furthermore, by using a household aggregation approach described above and focusing on children in their parents' households, the authors were able to determine the places of birth of the children's grandparents. When three or four of their grandparents were born in Mexico, these children were identified as Mexican nearly all the time (95 percent and 96 percent, respectively). But only 79 percent of those with two grandparents born in Mexico and 58 percent of those with one grandparent born in Mexico were identified as Mexican on the Hispanic origin question. And children who were not identified as Mexican had parents with much higher levels of educational attainment than those who retained a Mexican identification. The question on parents' place of birth in the CPS offers one of the best ways to study those who could identify as Hispanic, but the relatively small sample size of the CPS, even when pooled over many years, limits the use of these data to study other Hispanic subgroups.

Further information about the ethno-racial identities and places of birth of individuals' parents and grandparents are central to expanding research on the identity choices of adults who could consider themselves part of the multiracial, Hispanic and Native American populations. The challenge is to not define the populations by these individuals' ethno-racial identification (or ethno-racial classification by someone in the household) but to recognize that this identification is a dependent variable to a much greater extent than with other population groups. Accurately estimating and analyzing these groups therefore requires new strategies for determining the denominator – all the people who could identify in these populations – and treating self-identification as a numerator. With less-established classification norms for multiracial, Hispanic,

and Native American populations, current measures of racial and ethnic identification, the numerator, will increasingly fail to serve as stable measures that reliably delineate these populations.

[1] White and Black self-identifications are regularly found to be the most consistent over time and context; Asian self-identifications are also quite stable, generally showing rates of consistency just slightly below those of Whites and Blacks (Liebler et al. 2014). The rates for inconsistency between Asian self-identification and classification by others as Asian is similar, with Asians typically falling slightly below Blacks and Whites but higher than Hispanics or Native Americans in terms of consistency (Hahn, Truman, and Barker 1996; Kressin et al. 2003; West et al. 2005).

[2] These percentages are based on my calculations of the figures in Liebler et al. (2014)'s Table 2.

[3] Studies have also used this analytical approach with other data sets, including Add Health (Bratter and Heard 2009), Brazilian national household surveys (Schwartzman 2007), and the Chilean census (Valenzuela and Unzueta 2015).

[4] This percentage is my calculation based on data in Roth (2005), Tables 4 and 5.

References

Alba, Richard D., Noura E. Insolera, and Scarlett Lindeman. 2016. Is race really so fluid? Revisiting Saperstein and Penner's empirical claims. *American Journal of Sociology* 122 (1): 247–62.

Bratter, Jenifer, and Holly E. Heard. 2009. Mother's, father's, or both? Parental gender and parent-child interactions in the racial classification of adolescents. *Sociological Forum* 24 (3): 658–88.

Brown, J. Scott, Steven Hitlin, and Glen H. Elder. 2006. The greater complexity of lived race: An extension of Harris and Sim. *Social Science Quarterly* 87: 411–31.

Brunsma, David L. 2005. Interracial families and the racial identification of mixed-race children: Evidence from the early childhood longitudinal study. *Social Forces* 84 (2): 1131–57.

Doyle, Jamie Mihoko, and Grace Kao. 2007. Are racial identities of multiracials stable? Changing self-identification among single and multiple race individuals. *Social Psychology Quarterly* 70 (4): 405–23.

Duncan, Brian, and Stephen J. Trejo. 2011. Intermarriage and the international transmission of ethnic identity and human capital for Mexican Americans. *Journal of Labor Economics* 29 (2): 195–227.

Emeka, Amon, and Jody Agius Vallejo. 2011. Non-Hispanics with Latin American ancestry: Assimilation, race, and identity among Latin American descendants in the US. *Social Science Research* 40 (6): 1547–63.

Eschbach, Karl, and Christina Gómez. 1998. Choosing Hispanic identity: Ethnic identity switching among respondents to High School and Beyond. *Social Science Quarterly* 79 (1): 74–90.

Feliciano, Cynthia. 2016. Shades of race: How phenotype and observer characteristics shape racial classification. *American Behavioral Scientist* 60 (4): 390–419.

Gullickson, Aaron, and Ann Morning. 2011. Choosing race: Multiracial ancestry and identification. *Social Science Research* 40 (2): 498–512.

Hahn, Robert A., Benedict I. Truman, and Nancy D. Barker. 1996. Identifying ancestry: The reliability of ancestral identification in the United States by self, proxy, interviewer, and funeral director. *Epidemiology* 7 (1): 75–80.

Harris, David R., and Jeremiah Joseph Sim. 2002. Who is multiracial? Assessing the complexity of lived race. *American Sociological Review* 67 (4): 614–27.

Herman, Melissa R. 2010. Do you see what I am? How observers' backgrounds affect their perceptions of multiracial faces. *Social Psychology Quarterly* 73 (1): 58–78.

Hill, Mark E. 2002. Race of the interviewer and perception of skin color: Evidence from the Multi-City Study of Urban Inequality. *American Sociological Review* 67 (1): 99–108.

Kana'iaupuni, Shawn Malia, and Carolyn A. Liebler. 2005. Pondering poi dog: Place and racial identification of multiracial Native Hawaiians. *Ethnic and Racial Studies* 28 (July): 687–721.

Kressin, Nancy R., Bei-Hung Chang, Ann Hendricks, and Lewis E. Kazis. 2003. Agreement between administrative data and patients' self-reports of race/ethnicity. *American Journal of Public Health* 93 (10): 1734–39.

Krogstad, Jens Manuel, and Mark Hugo Lopez. 2015. Hispanic population reaches record 55 million, but growth has cooled. *Pew Research Center*. June 25. Available from www.pewresearch.org.

Liebler, Carolyn A. 2004. Ties on the fringes of identity. *Social Science Research* 33 (4): 702–23.

Liebler, Carolyn A. 2010. Homelands and indigenous identities in a multiracial era. *Social Science Research* 39 (4): 596–609.

Liebler, Carolyn A., and Timothy Ortyl. 2014. More than one million new American Indians in 2000: Who are they? *Demography* 51 (3): 1101–30.

Liebler, Carolyn A., Sonya Rastogi, Leticia E. Fernandez, James M. Noon, and Sharon R. Ennis. 2014. *America's churning races: Race and ethnic response changes between Census 2000 and the 2010 Census.* Revision of Center for Administrative Records Research and Applications (Census Bureau) Working Paper 2014-09.

Porter, Sonya R., Carolyn A. Liebler, and James M. Noon. 2016. An outside view: What observers say about others' races and Hispanic origins. *American Behavioral Scientist* 60 (4): 465–97.

Qian, Zhenchao. 2004. Options: Racial/ethnic identification of children of intermarried couples. *Social Science Quarterly* 85: 746–66.

Roth, Wendy D. 2005. The end of the one-drop rule? Labeling of multiracial children in black intermarriages. *Sociological Forum* 20 (1): 35–67.

Roth, Wendy D. 2010. Racial mismatch: The divergence between form and function in data for monitoring racial discrimination of Hispanics. *Social Science Quarterly* 91 (5): 1288–1311.

Roth, Wendy D. 2016. The multiple dimensions of race. *Ethnic and Racial Studies* 39 (8): 1310–38.

Saperstein, Aliya, and Andrew M. Penner. 2012. Racial fluidity and inequality in the United States. *American Journal of Sociology* 118 (3): 676–727.

Schwartzman, Luisa Farah. 2007. Does money whiten? Intergenerational changes in racial classification in Brazil. *American Sociological Review* 72 (6): 940–63.

Snipp, C. Matthew. 1997. Some observations about racial boundaries and the experiences of American Indians. *Ethnic and Racial Studies* 20 (4): 667–89.

U.S. Census Bureau. 2010. *Households and families: 2010.* 2010 Census Brief. Available from www.census.gov.

U.S. Office of Management and Budget. 1997. Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Register Notice,* October 30. Available from http://www.whitehouse.gov/omb.

Valenzuela, Eduardo, and M. Belén Unzueta. 2015. Parental transmission of ethnic identification in mixed couples in Latin America: The Mapuche case. *Ethnic and Racial Studies* 38 (12): 2090–2107.

Wallman, Katherine K., Suzann Evinger, and Susan Schechter. 2000. Measuring our nation's diversity: Developing a common language for data on race/ethnicity. *American Journal of Public Health* 90: 1704–8.

West, Carmen N., Ann M. Geiger, Sarah M. Greene, Emily L. Harris, In-Lu A. Liu, Mary B. Barton, Joann G. Elmore, et al. 2005. Race and ethnicity: Comparing medical records to self-reports. *Journal of the National Cancer Institute Monographs* 35: 72–74.

Xie, Yu, and Kimberly Goyette. 1997. The racial identification of biracial children with one Asian parent: Evidence from the 1990 Census. *Social Forces* 76 (2): 547–70.