ANALYZING RACE AND ETHNICITY WITH THE GSS

INTRODUCTION: KEY CONCEPTS IN RACE AND ETHNICITY

One of the most important arguments in the sociology of race and ethnicity is that race is a social construct. While differences in skin color, hair texture, eye shape, and eye color certainly exist, society plays an important role in determining which differences we consider important (e.g., skin color or eye shape) and which we tend to overlook (e.g., shoe size or the presence or absence of freckles). The meanings that we attach (or fail to attach) to these differences are also deeply intertwined with our cultural beliefs and historical traditions as well as our religious, political, educational, and economic institutions. Indeed, many of the racial and ethnic differences that may appear natural, universal, or obvious are actually the result of these very institutions. Differences in phenotype may exist apart from society, but "racial groups," "racial identities," "racial ideologies," and "racial inequality" do not. When analyzing inequalities related to race and ethnicity, it is always important to remember that, rather than a biological and stable property of individuals, race and ethnicity are created and maintained through social processes.

There is an ongoing debate in the social sciences about how to best collect and analyze the data needed to document racial/ethnic inequalities without simultaneously essentializing racial difference.[1] One perspective argues that, by asking respondents about their racial identities, surveys reproduce and strengthen racial categories and ultimately perpetuate racial inequalities.[2] When an official survey asks people, "What race do you consider yourself to be?" it can make it seem as if racial groups are straightforward,

Learning objectives

By the end of this chapter, you should be able to:

1. Identify variables related to race, ethnicity, and citizenship.

2. Produce a bivariate table, also called a *cross-tab*.

3. Use control variables to examine how the relationship between two variables may be influenced by a third.

4. Interpret these analyses within a social justice framework.

self-evident, and natural. It can make race seem like an essential, natural, and obvious characteristic of individuals rather than a socially constructed idea. Another perspective, the perspective advanced here, insists that asking respondents about their racial/ethnic identities in the context of surveys is important because if we stop collecting data about race—if social surveys no longer ask respondents how they identify with respect to racial and ethnic categories—then we lose much of our ability to track racial/ethnic inequalities. Without information about respondents' racial/ethnic identities, we would risk losing our collective ability to document wage inequality, inequality in graduation rates and other educational outcomes, and inequalities in incarceration and crime victimization, just to name a few. As the American Sociological Association (ASA) stated in its 2003 "Statement on the Importance of Collecting Data and Doing Social Scientific Research on Race,"

> sociological scholarship on "race" provides scientific evidence in the current scientific and civic debate over the social consequences of the existing categorizations and perceptions of race; allows scholars to document how race shapes social ranking, access to resources, and life experiences; and advances understanding of this important dimension of social life, which in turn advances social justice. *Refusing to acknowledge the fact of racial classification, feelings, and actions, and refusing to measure their consequences will not eliminate racial inequalities. At best, it will preserve the status quo.*[3]

In short, collecting survey data about racial/ethnic identities, attitudes, and inequalities can provide valuable tools for fighting for social justice—if the surveys are designed well and if the results are interpreted appropriately.

Fortunately for us, the GSS includes a large number of carefully crafted questions concerning race and ethnicity, and many of these questions can and have been used to document and challenge racial and ethnic inequalities. The GSS includes multiple indicators of respondents' racial identity and ethnic identity as well as questions concerning their ancestry. A number of other questions focus on attitudes and beliefs about racial inequality. There are also questions concerning the racial composition of respondents' friendship networks, the racial/ethnic composition of respondents' neighborhoods, experiences of racial/ethnic discrimination, and citizenship status.

This chapter provides an introduction to analyzing race and ethnicity in the General Social Survey, with a focus on the dimensions of race and ethnicity described above.

IDENTIFYING VARIABLES RELATED TO RACE AND ETHNICITY

The GSS has included numerous questions related to race and ethnicity since its beginning in 1972. Just as our national conversations about race, ethnicity, and

citizenship have changed over the past four decades, so too have the questions included in the survey. Some questions concerning respondents' racial identity and ethnic background have been asked regularly since the survey's beginning, but new questions have been added regularly in response to political, social, and cultural events as well as the changing sociodemographic composition of US society.[4]

Chapter 2 presented an overview of how to search for variables using the SDA website. When using the GSS to analyze issues related to race and ethnicity, three key aspects of the survey design must be taken into consideration. First is that racial and ethnic identities are too complex to be assessed with a single survey question. As a result, the GSS has some questions that focus on race, others that focus on ethnicity, and a handful of questions that combine the two. The variable RACE, for example, indicates respondents' racial status but makes no distinction between white respondents who identify as ethnically Mexican and those who identify as Irish American.

The variable ETHNIC asks respondents, "From what countries or part of the world did your ancestors come?" and includes response categories like "Germany," "Africa," "French Canada," "Other Canada," "China," "Czechoslovakia," "England," and "Wales." While this variable includes more detailed response categories, it is nonetheless limited in some important ways. The variable ETHNIC tells us nothing about how central these ethnic identities are to the respondent's sense of self, nor does it give us a clear picture of when the respondent's family came to the US. A further limitation to the variable ETHNIC is that it lumps together people of "African" descent while providing a much greater level of nuance for people with European ancestry. In response to these limitations, researchers often use information from multiple variables to construct new variables with greater levels of nuance.

A second key point to remember is that the variables concerning racial and ethnic identities have changed significantly over time. As explained on the GSS homepage, "until 2000, the GSS measured race mostly by interviewer observation" using the variable RACE, which is coded into three broad categories: "White," "Black," and "Other." Interviewers who were "in doubt" about a respondent's race were instructed to ask the respondent, "What race do you consider yourself?"[5] This approach to asking about race has some very clear limitations. First, this design strategy rests on the assumption that interviewers can, in most cases, accurately determine someone's racial identity simply by looking at them. This approach seems to rest upon, and to perpetuate, the idea that race is an essential, stable, and obvious characteristic of individuals rather than a socially constructed system of inequality. A second limitation concerns the response categories of "White," "Black," and "Other." Obviously, the "Other" category includes a very wide range of individuals: people who describe themselves as Chinese American, Latinos and Latinas, Native Americans, and biracial and multiracial people, to name a few. Combining these diverse groups into a single category, "Other," makes it difficult to examine the differences among these groups. Moreover, the clear focus on Black and White racial

groups seems to marginalize non-Black minorities, which is particularly problematic, given the increasing percentage of the US population who identify as Latino/a, as Asian American, or with multiple racial/ethnic groups.

To address some of these limitations, beginning in 2002, the GSS included a new procedure for measuring race and ethnicity, following the procedures used in the US Census of 2000.[6] From this point forward, the survey asked all respondents, "What is your race? Indicate one or more races that you consider yourself to be." The GSS records up to three racial/ethnic groups mentioned by the respondent, and these groups are recorded in the variables RACECEN1, RACECEN2, and RACECEN3. For surveys from 2002 and later, the value of the variable RACE has been imputed based on how respondents answered RACCEN1 and other information.

A third key feature of the survey design, noted in Chapter 2, is that the GSS was administered to both English- and Spanish-speaking respondents beginning in 2006. This change is important because it significantly increased the percentage of respondents who identified as Hispanic and/or Latino/a. While between 2000 and 2004, approximately 9% of respondents identified as Hispanic, Latino, or Latina, this number increased to approximately 15% in 2006 and the following years. By allowing non-English-speaking Spanish speakers to participate in the GSS, the surveys from 2006 onward more fully represent the diverse experiences of Latinos and Latinas in the United States.

In addition to the three key design issues noted above, several other features of the survey are worth noting here. First, it is important to remember that the GSS is a survey that includes both US citizens and non-US citizens living in the United States. The variable CITIZEN asks respondents about whether they are citizens of America, though this variable is asked only to a subsample of respondents and was included only in years 1996, 2004, and 2014.[7] Second, as discussed in Chapter 2, the survey design includes a deliberate oversample of African American respondents in years 1982 and 1987. For these years in particular, the percentage of survey respondents who identified as Black or African American was higher than the percentage of people who identified with these groups in the general population. And finally, as also discussed in Chapter 2, it is important to remember that the GSS is administered only to individuals who are not currently living in institutions. To the extent that racial/ethnic minority groups are overrepresented in institutions such as prisons, jails, and homeless shelters, the survey likely gives us a more positive representation of racial/ethnic minorities' experiences than exists in the general population.

Despite the limitations mentioned above, the GSS remains a valuable tool for analyzing inequalities related to race and ethnicity. Given the number of questions pertaining to issues of race and ethnicity and the additional questions being developed for each new survey, the opportunities for analyzing racial and ethnic inequalities with the GSS are virtually limitless. As is the case with all research projects, however, it is important to communicate the limitations of your data and analysis along with your results.

Searching for Variables Related to Race and Ethnicity

The first step in identifying variables related to race, ethnicity, and multiculturalism is to brainstorm possible key words. *Race* and *ethnicity* are good starting points, but consider some others: *white*, *Hispanic*, *Black*, *color*, *African American*, *Asian*, *racial*, *ethnic*, *citizen*, *immigrant*, *immigration*, *language*, *minorities*, *country*, *family origin*.

WHOSE RACE AND ETHNICITY?

As you search for variables using the key word *race* or *ethnicity*, you will come across variables that ask about the respondents' racial/ethnic identities. Make note, however, that there are also variables that ask about the respondents' spouses' racial/ethnic identity as well as the racial/ethnic identities of respondents' friends, neighbors, and coworkers. The variables INTRACE1 ("What is your race?

Indicate one or more races that you consider yourself to be") and INTETHN ("*race of the interviewer*") do not correspond to the race or ethnicity of the respondent! These variables pertain to the interviewer, not the interviewee. Be careful not to mix them up. And always "view" a variable before analyzing it to make sure that you know exactly what the variable represents.

Browsing for Variables Related to Race and Ethnicity

Another approach for identifying variables related to race and ethnicity is to "browse" for variables using the online codebook. There are several headings that deal specifically with race, ethnicity, and multiculturalism, but several of the other sections include relevant variables as well.

If you are interested in browsing for variables related to race and ethnicity, you may find the following subject headings and subheadings to be particularly useful:

- Respondent Background Variables
  - Age, Gender, Race, and Ethnicity

- Personal Concerns
  - Race Issues

- Controversial Social Issues
  - Race Part Two
  - Social Issues Scales

- Obligations and Responsibilities
  - Government Responsibility

In addition, a number of special modules have been included that focus on issues of race, ethnicity, and multiculturalism. While the variables topics below are included in

only a handful of surveys years, analyses of these variables can provide a particularly nuanced analysis of attitudes and beliefs about race, ethnicity, and multiculturalism.

- 1985: Social Networks

- 1987: Social and Political Participation

- 1990: Inter-group Relations

- 1994: Multiculturalism

- 2000: Multi-ethnic United States

- 2002: Prejudice

- 2004: Immigration

- 2004: National Identity

- 2004: Citizenship

- 2012: Jewish Identity

- 2012: Skin Tone

- 2014: National Identity III

- 2014: Citizenship

Remember that there are many variables related to race and ethnicity that occur outside the subject headings mentioned above. By browsing through the codebook and searching for terms related to race and ethnicity, you should be able to find hundreds of survey questions directly related to race and ethnicity.

COMBINING RACE AND HISPANIC ETHNICITY

The variable RACEHISP, which is included from the 2000 survey onward, combines information from the variables RACE and HISPANIC to create a variable about respondents' racial/ethnic identity with four response categories: "White," "Black," "Hispanic," and "Other." In this variable, the category "White" does not include those who said they were Hispanic, and the category "Hispanic" does not include those who said they were Black. When viewing the variable in SDA, the response categories appear to be "White," "Black," "Hispanic," and "Other." When analyzing this variable and writing up your results, be sure to specify that "White" refers to "non-Hispanic whites," "Hispanic" refers to "non-Black Hispanics," and "Black" refers to "Blacks and African Americans, both Hispanic and non-Hispanic."

Viewing a Variable

Once you have identified potential variables to analyze, the next step is to "view" the variable so that you can determine exactly what survey question the variable represents. Viewing a variable in SDA is easy and is discussed in detail in Chapter 2. From the GSS SDA homepage, simply type the variable name into the "Variable Selection" field (this is in the upper left corner) and then click "View." A new window will open, in which you will find (in almost all cases) the exact wording of the question as well as the frequency distribution of the responses. Remember that when you view a variable in this way, the resulting frequency distribution shows the combined responses from *all* survey years in which the question was asked.

For example, Figure 4.1 shows the frequency distribution for the variable that results from viewing the variable IMMIMP. By viewing this variable, we can see that it assesses respondents' beliefs about immigration. In particular, the survey asks respondents, "How much do you agree or disagree with each of the following statements? a. Immigrants improve American society by bringing in new ideas and cultures."

Figure 4.1

IMMIMP	IMMIGRANTS IMPROVE AMERICAN SOCIETY

Description of the Variable

1461. There are different opinions about immigrants from other countries living in America. (By "immigrants" we mean people who come to settle in America). How much do you agree or disagree with each of the following statements? a. Immigrants improve American society by bringing in new ideas and cultures

Percent	N	Value	Label
10.0	119	**1**	AGREE STRONGLY
47.2	563	**2**	AGREE
24.7	294	**3**	NEITHER AGREE NOR DISAGREE
14.8	176	**4**	DISAGREE
3.4	40	**5**	DISAGREE STRONGLY
	58,383	**0**	IAP
	23	**8**	CANT CHOOSE
	1	**9**	NA
100.0	**59,599**		**Total**

From this frequency distribution, we can see that 10% of respondents strongly agreed with the idea that immigrants improve American society, and another 47.2% agreed with this idea. Nearly a quarter of respondents neither agreed nor disagreed, and less than 20% disagreed. (For more information on how to interpret a frequency distribution, please review Chapter 2.)

An important limitation of viewing the variable in this way is that it does not provide information about *when* this variable was included in the GSS. The data could potentially have been collected anywhere between 1972 and 2014, or at multiple years within this time frame. To determine when the data were collected, it is necessary to construct a bivariate table. Bivariate tables, also called cross-tabulations or "cross-tabs" for short, will also allow us to assess the characteristics associated with beliefs about the effect of immigration on American society.

PRODUCING AND INTERPRETING A BIVARIATE TABLE OR "CROSS-TAB"

Background

Knowing how to produce and critically interpret a meaningful bivariate table is one of the most important tools of data analysis. Bivariate simple means "two variables," and bivariate tables are used to explore the relationship between two variables. If we were interested in knowing whether white, Black, and Hispanic adults report similar levels of home ownership, for example, we could make a cross-tab using the variables DWELOWN—"(Do you/Does your family) own your (home/apartment), pay rent, or what?"—and RACEHISP to explore this relationship. We could also examine the racial/ethnic differences in beliefs about racial inequality by analyzing the variables RACEHISP and WRKWAYUP. WRKWAYUP corresponds to the survey question "Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement? Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. [Blacks should do the same without special favors.]"[8]

Before creating a bivariate table, it is useful to consider whether you believe the relationship you are investigating is a causal relationship or if the relationship would be better understood as a correlation. (See Chapter 1 for a more in-depth discussion of causation and correlation.) In a causal relationship, the variable doing the causing is termed the independent variable and the variable that is affected is the dependent variable. In other words, the dependent variable is thought to depend on the independent variable.

In an analysis of racial/ethnic differences in home ownership, for example, respondents' racial/ethnic status would be the independent variable and home ownership would be the dependent variable. To the extent that a relationship between these two variables exists, it is racial/ethnic status that influences home ownership and not the other way around. Similarly, in an analysis of what whites, Blacks, Hispanics, and other racial/ethnic groups believe about the extent of racial

inequality in the US, racial/ethnic status would be the independent variable and beliefs about racial inequality would be the dependent variable.

When testing for racial/ethnic differences in experiences, beliefs, attitudes, identities, and other social phenomena, racial/ethnic status will generally be the independent variable. The experiences, beliefs, attitudes, and identities that you believe will differ for people who hold different racial/ethnic identities will be dependent variables.
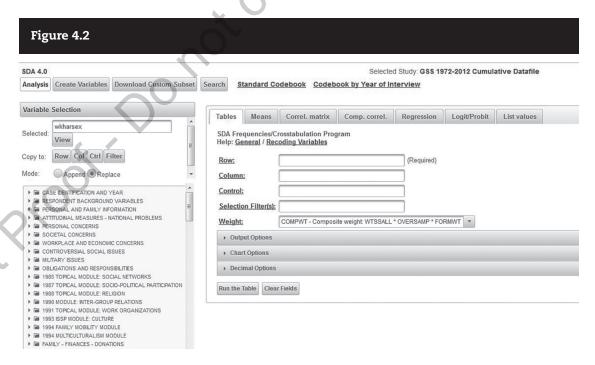
Creating a Cross-Tab in SDA

To produce a bivariate table in SDA, first make sure that you are on the "Analysis" section of the SDA website. In the upper left corner of the SDA page, click on the "Analysis" button.

The screen should look like Figure 4.2.

The left-hand side of the screen shows the codebook, and the right side is used to conduct the analysis. In cases where you believe there is a causal relationship, enter the name of your dependent variable in the "Row" field and the name of the independent variable in the "Column" field.

Let's use the example of racial/ethnic differences in home ownership as an example. Whether respondents own or rent their homes can be assessed with the variable DWELOWN: "(Do you/Does your family) own your (home/apartment), pay rent, or what?" This is a nominal-level variable and has been included regularly in the GSS,

Figure 4.2

beginning in 1985. If we are interested in knowing whether homeownership differs across racial/ethnic groups, DWELOWN is the dependent variable and RACEHISP is the independent variable.

In the "Weight" field, you will see that the default setting for weights is "COMPWT - Composite weight WTSSALL *Oversamp* FORMWT." To replicate the analyses as presented in this chapter, keep this selection as is.[9]

Enter the dependent variable, DWELOWN, in the "Row" field and your independent variable, RACEHISP, in the "Column" field, and then click the "Run the Table" button. A new window will appear that presents the cross-tab. Figure 4.3 shows the resulting cross-tab.

Figure 4.3

Variables					
Role	**Name**	**Label**	**Range**	**MD**	**Dataset**
Row	**DWELOWN**	DOES R OWN OR RENT HOME?	1-3	0,8,9	1
Column	**RACEHISP**	Race with Hispanic (2000 and later)	1-4	9	1
Weight	**COMPWT**	Composite weight = WTSSALL * OVERSAMP * FORMWT	.1913-11.1261		1

Frequency Distribution						
Cells contain: -**Column percent** -Weighted N		**RACEHISP**				
		1 White	2 Black	3 Hispanic	4 Other	*ROW TOTAL*
DWELOWN	1: OWN OR IS BUYING	**74.3** 5,847.2	**46.8** 730.4	**48.4** 657.8	**56.1** 297.1	*66.5* *7,532.5*
	2: PAYS RENT	**24.0** 1,886.9	**51.7** 806.8	**49.8** 677.0	**42.1** 222.9	*31.7* *3,593.6*
	3: OTHER	**1.8** 139.0	**1.4** 22.6	**1.8** 23.8	**1.8** 9.8	*1.7* *195.1*
	COL TOTAL	*100.0* *7,873.1*	*100.0* *1,559.7*	*100.0* *1,358.5*	*100.0* *529.8*	*100.0* *11,321.2*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

Interpreting a Cross-Tab

The window that appears contains three main parts. First is a description of the variables used in the analysis (labeled "Variables"), followed by the cross-tab (labeled "Frequency Distribution"). Below the cross-tab is a chart that gives a visual representation of the information presented in the cross-tab.

The "Variables" section, at the top, presents a summary of the analyses that follow, including variable names and short variable descriptions (variable labels). The "Variables" section also includes the range of valid response categories ("Range") and the response categories that correspond to nonvalid values ("MD" or "Missing Data").

The "Frequency Distribution" section shows your bivariate table. In the upper left corner of the frequency distribution chart, we see that the cells contain "column percents" and the number of cases (N) in each cell. Let's first look at those cells that are red or blue. The dark red cell in the upper left corner contains two numbers: 74.3 and 5,847.2. The N of 5,847.2 indicates that about 5,847 respondents in the GSS self-identified as non-Hispanic whites and also either owned their home outright or were in the process of paying off their home mortgage.

INTERPRETING THE N IN A CROSS-TAB THAT USES WEIGHTS

Recall that N stands for "number of respondents." While the number of respondents would typically be a whole number, the use of survey weights makes a slight adjustment to these numbers to make the results better reflect the US adult population. See Chapter 2 for a more detailed discussion of survey weights.

To the immediate right, we see a dark blue cell that contains the numbers 46.8 and 730.4. This tells us that about 730 respondents in the GSS identified as African American or Black and indicates that they too were in the process of paying off their mortgage or owned their home outright. Comparing the two Ns, we can see that far more white respondents owned their homes than did Black or African American respondents. These numbers can be misleading, however, because the number of non-Hispanic white people included in this analysis (7,873) was much greater than the number of Black or African American people included (1,560).

In addition to the number of cases in each cell (N), the cross-tab provides the column percents for each cell. Column percents are calculated by dividing the number of cases in a given cell by the total number of cases in each column and then multiplying the resulting proportion by 100. So, in the upper left corner, the column percent of 74.3 is calculated in the following way:

$$(5847.2 / 7873.1) * 100 = 74.3\%$$

The column percent tells us the percentage of individuals in each category of the column variable who fall into each category of the row variable. In this case, of the approximately 7,873 non-Hispanic white respondents, 74.3% reported that they were homeowners. The cell to the right indicates that about 46.8% of African American respondents indicated that they were homeowners. This number is calculated by dividing the total number of African Americans who answered the question by the number of African Americans who said they were homeowners:

$$(730.4 / 1559.7) * 100 = 46.8\%$$

The table also shows us that 48.4% of non-Black Hispanic respondents were in the process of paying off their mortgage or already owned their home outright. Of those respondents who identified with an "Other" racial/ethnic group, 56.1% owned their home or were in the process of paying off their mortgage.

The row totals at the far right-hand side of the table tell us the overall distribution of the row variable (DWELOWN). Taken on the whole, including people of all racial/ethnic groups, about 7,532 respondents indicated that they owned, or were in the process of buying, their homes, and this number corresponds to 66.5% of the total number of respondents:

$$(7532.5 / 11321.2) * 100 = 66.5\%$$

The column totals in the bottom row of the table tell us the overall distribution of the variable in the column (RACEHISP). Reading across the bottom row of the table, we see that approximately 7,873 respondents were non-Hispanic whites, about 1,560 were Black or African American, 1,358 were non-Black Hispanics, and about 530 were people who identified with other racial/ethnic groups.

The bottom right corner of the table shows a percentage of 100 and an N of approximately 11,321. This number corresponds to the total number of respondents included in the table. *Because we have not specified any specific time period for this analysis, the resulting bivariate table is drawing from all available data, anywhere from 1972 to 2014.* If we wanted to focus on a specific time period, we could use the "Filter" option in SDA, which is described below.

When we look at the percentages in the cross-tab, some big differences can be seen. For instance, 24% of whites indicate that they are currently renting their homes, but the percentages of Blacks and Hispanics who are renting are 51.7 and 49.8, respectively. In other words, compared to non-Hispanic whites, Blacks and Hispanics are more than twice as likely to be renters.

WHAT DO THE COLORS IN A BIVARIATE TABLE MEAN?

When you produce a cross-tab, you will notice that the cells are different shades of blue and red. These colors are meant to help analysts to identify patterns in the data and provide an indication of whether the relationship between two variables is statistically significant. The color of each cell reflects the Z-statistic, which shows whether the frequencies in a cell are greater or fewer than we would expect if there were no relationship between the variables in the general population.[10]

Again, because we have not yet limited the analysis to a particular time period, the analysis of racial/ethnic differences in homeownership (shown in Figure 4.3) combines information from every year in which data for the variables RACEHISP and DWELOWN are available. But let's imagine that we are interested in what's been happening more recently.

Applying a Filter

To limit the analysis so that it includes data from only recent years, add a **filter** by typing the following into the "Selection Filter" field:

YEAR (2010-2014)

To make a cross-tab of the variables DWELOWN and RACEHISP that focuses on data from 2010 to 2014, now type the variable DWELOWN into the row field and RACEHISP into the column field. Remember to keep the default of COMPWT in the "Weight" field. The resulting cross-tab is shown below in Figure 4.4. Notice that the table now includes a row for "Filter," which reminds us that the analysis includes only data from the 2010 to 2014 surveys. In this particular case, the results from our original analysis (Figure 4.3) are very similar to our results that use the filter. This is not always the case!

Filters can be used to limit the analysis to data from a particular time period (as shown above), but they can also be used to limit the analysis to particular social groups, to people within a specific age range, to those who identify with particular political parties, or to people who reside in particular types of communities.

When presented with the above analysis, a critical researcher might say something like, "The discrepancy between homeownership isn't so much about racial/ethnic

Figure 4.4

	Variables					
Role	**Name**	**Label**	**Range**	**MD**	**Dataset**	
Row	**DWELOWN**	DOES R OWN OR RENT HOME?	1-3	0,8,9	1	
Column	**RACEHISP**	Race with Hispanic (2000 and later)	1-4	9	1	
Weight	**COMPWT**	Composite weight = WTSSALL * OVERSAMP * FORMWT	.1913-11.1261		1	
Filter	**YEAR(2010-2014)**	GSS YEAR FOR THIS RESPONDENT	1972-2014		1	

Frequency Distribution						
Cells contain: -**Column percent** -Weighted N		**RACEHISP**				
		1 White	2 Black	3 Hispanic	4 Other	**ROW TOTAL**
DWELOWN	1: OWN OR IS BUYING	**73.4** 2,122.9	**45.3** 295.7	**46.3** 295.8	**56.0** 120.8	**64.4** 2,835.3
	2: PAYS RENT	**25.2** 728.0	**53.8** 350.7	**51.4** 328.3	**41.8** 90.3	**34.0** 1,497.3
	3: OTHER	**1.5** 41.9	**.9** 6.0	**2.2** 14.3	**2.2** 4.7	**1.5** 66.9
	COL TOTAL	**100.0** 2,892.9	**100.0** 652.4	**100.0** 638.3	**100.0** 215.9	**100.0** 4,399.5

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

inequality but rather a reflection of the fact that racial/ethnic minorities are more likely than whites to reside in large cities, and in urban areas, people tend to rent more frequently." This critical researcher makes a reasonable point. Would the inequality we see in homeownership in Figure 4.4 remain if we were to examine only those living in large cities or only those living in small towns? Using a second filter can help us to answer this question.

By viewing the variable XNORCSIZ, we can see that this variable corresponds to the size of the town or city in which the respondent lives and that the value of "1" for this variable represents a city with a population of greater than 250,000 people. To examine whether racial/ethnic differences in homeownership persist when only those respondents who live in big cities are included, begin by entering DWELOWN in the "Row" field and the variable RACEHISP in the "Column" field. In the "Selection Filter" field, type:

YEAR (2010-2014) XNORCSIZ (1)

This filter will restrict the analysis to data collected between 2010 and 2014 and will now include only those respondents who lived in cities with populations greater than 250,000 people.

Figure 4.5 shows the resulting analysis. Notice that there are now two rows for filters and that the total number of cases in the table has been significantly reduced. In Figure 4.4, approximately 4,400 people were included in the analysis. Because we include only people living in large cities in Figure 4.5, the analysis is now restricted to approximately 790 people.

Figure 4.5 shows that, even when the analysis is focused only on those living in cities of more than 250,000 people, racial/ethnic differences in homeownership persist. Among non-Hispanic white respondents, the majority (61.4%) either own their home outright or are in the process of paying off their mortgage. Among Black and Hispanic respondents, the rate of homeownership is much lower (36.8% and 33.4%, respectively). For those respondents who live in big cities and who identify with other racial/ethnic groups, the rate of home ownership (61.1%) is similar to that of non-Hispanic whites.

For many people, the cross-tab above raises more questions than it answers. For example:

1. What factors in addition to respondents' racial/ethnic status might be influencing homeownership?

2. How might rates of homeownership differ for men and women of diverse racial/ethnic groups?

3. Does the unequal access to homeownership that we see here remain when we control for differences in education?

The first of these questions can be answered by constructing additional cross-tabs using DWELOWN as the dependent variable and including independent variables like DEGREE (highest degree earned), MARITAL (respondent's marital status), and SEX (respondent's gender). The second and third questions above can be assessed with control variables, which are discussed below.

Figure 4.5

		Variables			
Role	**Name**	**Label**	**Range**	**MD**	**Dataset**
Row	**DWELOWN**	DOES R OWN OR RENT HOME?	1-3	0,8,9	1
Column	**RACEHISP**	Race with Hispanic (2000 and later)	1-4	9	1
Weight	**COMPWT**	Composite weight = WTSSALL * OVERSAMP * FORMWT	.1913-11.1261		1
Filter	**YEAR(2010-2014)**	GSS YEAR FOR THIS RESPONDENT	1972-2014		1
Filter	**XNORCSIZ(1)**	EXPANDED N.O.R.C. SIZE CODE(=CITY GT 250000)	1-10	0	1

Frequency Distribution						
Cells contain: -**Column percent** -Weighted N		**RACEHISP**				
		1 White	2 Black	3 Hispanic	4 Other	*ROW TOTAL*
DWELOWN	1: OWN OR IS BUYING	61.4 201.3	36.8 71.0	33.4 69.7	61.1 36.8	*48.0* *378.7*
	2: PAYS RENT	37.5 122.8	62.1 119.9	64.4 134.4	38.1 23.0	*50.7* *400.1*
	3: OTHER	1.1 3.6	1.1 2.1	2.3 4.7	.8 .5	*1.4* *10.9*
	COL TOTAL	*100.0* *327.7*	*100.0* *192.9*	*100.0* *208.9*	*100.0* *60.2*	*100.0* *789.7*

Applying Control Variables

A **control variable** is an independent variable that is included in the analysis in order to determine whether a relationship between two variable holds true when variation in a third variable (that is, the control variable) is held constant.

In the above example, we first examined the relationship between respondents' racial/ethnic group and their status as homeowners (Figure 4.4). We saw vast differences in homeownership across racial/ethnic groups and concluded that non-Hispanic whites were significantly more likely than African Americans, Hispanics, or other

racial/ethnic minorities to own homes. In Figure 4.5, we then examined whether this relationship held true when only those respondents who were living in large cities were included in the analysis. But what about people living in medium-sized cities and small towns? We could rerun the analyses of RACEHISP and DWELOWN several times, each time using a different value for the filter value. We could start with the filter "XNORCSIZ (1)," and then replace that filter with "XNORCSIZ (2)," and then replace that filter with "XNORCSIZ (3)," and so forth.

A more efficient strategy is to use XNORCSIZ as a control variable rather than as a filter. To do this, simply enter DWELOWN in the "Row" field and RACEHISP in the "Column" field. Enter the variable XNORCSIZ in the "Control" field. In the "Selection Filter" field, keep the filter for year (2010–2014). Run the table, and in the window that opens, you will see 11 bivariate tables, each followed by a bar chart. Each table examines the relationship between the main independent variable (RACEHISP) and the dependent variable (DWELOWN), but these tables show the relationship between these two variables, *holding community size (XNORCSIZ) constant.* The first table you see should be identical to that shown in Figure 4.5. Below that, the second table shows the racial/ethnic differences in homeownership in medium-sized cities (with populations of 50,000 to 250,000), as shown in Figure 4.6. The third shows this relationship for people living in the suburbs of large cities (Figure 4.7), and so forth. Scrolling all the way down, the last table will show a summary table with a cross-tab identical to that shown in Figure 4.4.

As you make your way through these tables, you may notice that the number of respondents in each table varies. This is because more people live in big cities and in

Figure 4.6

Statistics for XNORCSIZ = 2(CITY,50-250000)						
Cells contain: -**Column percent** -Weighted N		**RACEHISP**				
		1 White	2 Black	3 Hispanic	4 Other	*ROW TOTAL*
DWELOWN	1: OWN OR IS BUYING	**62.8** 252.9	**38.6** 48.3	**58.5** 68.4	**42.0** 13.4	*56.6* *383.0*
	2: PAYS RENT	**35.2** 141.7	**60.8** 76.1	**40.0** 46.8	**58.0** 18.5	*41.8* *283.0*
	3: OTHER	**2.0** 8.0	**.7** .8	**1.5** 1.8	**.0** .0	*1.6* *10.6*
	COL TOTAL	*100.0* *402.5*	*100.0* *125.2*	*100.0* *117.0*	*100.0* *31.9*	*100.0* *676.6*

Figure 4.7

Statistics for XNORCSIZ = 3(SUBURB, LRG CITY)						
Cells contain: -**Column percent** -Weighted N		**RACEHISP**				
		1 White	2 Black	3 Hispanic	4 Other	*ROW TOTAL*
DWELOWN	1: OWN OR IS BUYING	**72.8** 428.4	**47.1** 60.8	**49.9** 85.6	**62.5** 35.5	*64.5* *610.4*
	2: PAYS RENT	**26.5** 156.1	**51.8** 66.9	**48.5** 83.3	**37.5** 21.3	*34.6* *327.6*
	3: OTHER	**.6** 3.6	**1.1** 1.4	**1.6** 2.8	**.0** .0	*.8* *7.8*
	COL TOTAL	*100.0* *588.2*	*100.0* *129.1*	*100.0* *171.7*	*100.0* *56.8*	*100.0* *945.8*

the suburbs of big cities than in small towns. Notice also that the pattern of racial/ethnic inequality in homeownership remains similar throughout each of these tables. The percentages of whites, African Americans, Hispanics, and other minorities who own their homes vary a little depending on the size of town or city considered. In almost all cases, however, non-Hispanic whites show significantly higher rates of homeownership than Hispanics, Blacks, and other racial/ethnic minorities.

APPLICATION: IS EDUCATION ASSOCIATED WITH BELIEFS ABOUT RACIAL INEQUALITY?

This application uses a series of cross-tabs to examine the relationship between respondents' educational attainment and their beliefs about racial inequality. While respondents' educational attainment and beliefs about racial/ethnic inequality can both be assessed with multiple variables, here we focus on the variables DEGREE and WRKWAYUP.

Step 1. Restate the research question and identify the independent and dependent variables.

The research question is "How and to what extent is educational attainment related to beliefs about racial inequality? The dependent variable, beliefs about racial inequality, is measured with the variable WRKWAYUP. The independent variable is DEGREE—respondents' highest educational degree.

Step 2. View each variable to make sure it means what you think it means. By viewing the variable, you can see the precise wording of the survey question that corresponds to the variable.

By viewing the variable WRKWAYUP, we can see that this is an ordinal-level variable corresponding to the following survey question:

> Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement? Irish, Italian, Jewish and many other minorities overcame prejudice and worked their way up. [Blacks should do the same without special favors.]

The variable DEGREE is a special case where viewing the variable does not reveal the exact survey question because the variable values were based on respondents' answers to multiple other questions.[11] DEGREE is an ordinal-level variable with five categories ranging from 0, which represents less than a high school degree, to 4, which represents a graduate degree.

Step 3. Determine which time period the data is from and choose a selection filter based on the time frame you wish to analyze. Remember that when you view a variable, you are viewing the combined data across all survey years for which there are data. If the question was asked regularly since 1972, this is four decades of data. If the question was asked only once—say, in 1985—then we could still analyze the data, but we would want to specify that the data are more than 30 years old. It is thus always crucial to determine the survey years in which the variable you are analyzing was included.

The easiest way to do this is by creating a quick cross-tab of the variables in your analysis by the variable YEAR (the variable that corresponds to the survey year). Producing a cross-tab with WRKWAYUP in the "Row" field and YEAR in the "Column" field (with COMPWT selected in the "Weight" field) will result in Figure 4.8.

As shown in Figure 4.8, the variable WRKWAYUP has been included in the GSS regularly since 1994. Unless we limit our analysis to recent data using the "Filter" command, the resulting cross-tabs will be drawing from all of this data.

DEGREE is also a core variable in the GSS and is included with every survey year. To double-check, simply produce a cross-tab of the variable DEGREE by the variable YEAR.

For this exercise, let's limit the analysis to very recent years: 2012 and 2014.

Step 4. Conduct the relevant analysis. Create a cross-tab of the variable WRKWAYUP by the variable DEGREE with a selection filter for YEAR 2012-2014. When examining how respondents' beliefs about racial inequality are shaped by respondents' educational attainment, we are assuming that beliefs about racial

Figure 4.8

Frequency Distribution (WRKWAYUP)

Cells contain: -Column percent -Weighted N	YEAR											ROW TOTAL
	1994	1996	1998	2000	2002	2004	2006	2008	2010	2012	2014	
1: AGREE STRONGLY	**44.7** 638.1	**44.7** 866.4	**42.4** 777.7	**44.3** 808.4	**47.7** 425.0	**41.0** 365.3	**44.6** 873.7	**44.4** 586.3	**41.1** 582.9	**41.7** 535.1	**39.4** 651.0	**43.3** 7,110.1
2: AGREE SOMEWHAT	**30.1** 429.2	**28.9** 560.6	**30.7** 563.8	**28.5** 519.6	**26.5** 235.9	**27.7** 246.7	**29.0** 568.9	**30.2** 399.1	**30.7** 436.1	**29.1** 373.4	**29.0** 478.5	**29.3** 4,811.8
3: NEITHER AGREE NOR DISAGREE	**10.1** 144.5	**11.7** 227.1	**12.8** 234.9	**11.0** 200.5	**12.7** 113.2	**15.5** 138.2	**13.0** 253.7	**12.6** 166.5	**13.1** 186.4	**14.5** 185.9	**14.9** 246.8	**12.8** 2,097.7
4: DISAGREE SOMEWHAT	**9.7** 138.0	**9.2** 178.7	**8.4** 153.5	**9.4** 170.7	**7.5** 66.4	**10.6** 94.7	**8.4** 164.4	**7.9** 104.7	**8.6** 121.3	**9.1** 117.0	**10.0** 165.7	**9.0** 1,475.2
5: DISAGREE STRONGLY	**5.4** 76.3	**5.4** 104.3	**5.8** 106.2	**6.9** 126.3	**5.6** 49.6	**5.1** 45.7	**5.0** 98.6	**4.8** 63.5	**6.5** 92.4	**5.5** 70.6	**6.7** 110.7	**5.7** 944.2
COL TOTAL	*100.0* 1,426.2	*100.0* 1,937.0	*100.0* 1,836.0	*100.0* 1,825.6	*100.0* 890.2	*100.0* 890.7	*100.0* 1,959.3	*100.0* 1,320.2	*100.0* 1,419.2	*100.0* 1,282	*100.0* 1,652.7	*100.0* 16,439

Figure 4.9

Frequency Distribution							
		DEGREE					
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW* *TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**40.8** 153.9	**46.2** 689.7	**49.2** 113.4	**28.7** 155.1	**25.3** 74.1	*40.4* *1,186.2*
	2: AGREE SOMEWHAT	**33.6** 126.8	**29.1** 435.4	**25.7** 59.3	**29.8** 160.8	**23.7** 69.6	*29.0* *851.9*
	3: NEITHER AGREE NOR DISAGREE	**14.0** 52.9	**13.1** 195.0	**15.1** 34.8	**19.1** 103.1	**16.0** 46.9	*14.7* *432.6*
	4: DISAGREE SOMEWHAT	**7.3** 27.6	**7.2** 108.2	**4.9** 11.3	**13.4** 72.6	**21.5** 63.1	*9.6* *282.7*
	5: DISAGREE STRONGLY	**4.2** 15.8	**4.4** 65.7	**5.1** 11.8	**9.0** 48.4	**13.5** 39.6	*6.2* *181.3*
	COL TOTAL	*100.0* *377.1*	*100.0* *1,493.9*	*100.0* *230.6*	*100.0* *539.9*	*100.0* *293.2*	*100.0* *2,934.7*

inequality are the dependent variable and education level is the independent variable.

Begin constructing your cross-tabs by entering WRKWAYUP in the "Row" field (since it is the dependent variable) and DEGREE in the "Column" field (since it is the independent variable).

Use the Selection Filter field to restrict the analysis to very recent data by typing:

YEAR (2012-2014)

Figure 4.9 shows the resulting cross-tab.

Step 5. Interpret your results. There are five basic step to interpreting a cross-tab.

1. Remind your audience of the basics.

When presenting your analyses to an audience, it is important to:

a. *Restate your research question.* In this case, the research question is "Among adults in the contemporary US, is there a relationship between educational attainment and beliefs about racial inequality?"

b. *Remind your audience of the data source and the specific variables that you used to answer this question.* In this case, the data source is the 2012–2014 General Social Surveys. Respondents were asked, "Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement? Irish, Italian, Jewish and many other minorities overcame prejudice and worked their way up. [Blacks should do the same without special favors.]" Respondents were also asked about the highest degree they earned.

c. *Identify and describe the dependent and independent variables, clearly stating the level of measurement for each variable and how each variable was coded.* In this case, beliefs about racial inequality are the dependent variable and are assessed with a five-category ordinal-level variable, where higher values indicate stronger disagreement. DEGREE is the independent variable, which is also an ordinal-level variable with five categories, but here values range from 0 to 4.

d. *Specify the number of cases included in the analysis.* The overall number of valid cases (N) included in the cross-tab is presented in the bottom right corner of each cross-tab. In Figure 4.9, the weighted number of cases in the analysis consists of about 2,935 respondents.

2. Focus on specifics.

The first step in interpreting a cross-tab is to look carefully at the specific numbers in the tables and interpret them as specifically as you can. Focusing on the cell frequencies (the bottom number in each cell of the table) can be useful, but in most cases it is more helpful to examine the column percents—the top number in each cell.

For example, the numeric cell in the bottom left corner of Figure 4.9 contains the number 377.1, and this number is the column frequency. This number tells us that there were approximately 377 individuals with less than a high school education who provided information about their highest degree and their beliefs about racial ideology. The top row of Figure 4.9 corresponds with a response of "strongly agree." In the upper left corner of Figure 4.9, we see that approximately 154 people with less than a high school education responded that they "strongly agreed" with this statement. The 40.8 in this cell is the column percent and shows us the percentage of people with less than a high school degree who strongly agreed with this statement.

$$(153.9 / 377.1) * 100 = 40.8\%$$

In the top right cell, we see that 25.3% of respondents who had a graduate degree strongly agreed with the idea that Blacks should overcome prejudice and work their way up without special favors.

Looking at the row that corresponds with "strongly disagree," we can see that 4.2% of respondents with less than a high school degree strongly disagreed with this idea. This increased to 13.5% for respondents with a graduate degree.

The modal category (the category containing the highest number of cases) is "strongly agree" or "agree" for all levels of education.

3. Consider the big picture.

After examining individual percentages within the table, it is important to step back and take a larger view of the overall relationship presented in the tables. When examining each table individually, can you see any patterns or trends in the column percentages, or do the percentages seem to go up and down at random? If you see clusters of dark blue or dark red cells in the table, then there is probably an identifiable pattern. If there are very few darkly colored cells or if they seem randomly scattered across the table, then there may not be an easily identifiable pattern. It's also important to think about how the individual analyses work together. In situations where you are analyzing multiple bivariate tables, do they all point to the same overall conclusion, or are the results more mixed? If the individual analyses seem to point to different conclusions, how do you make sense of these results?

Look again at Figure 4.9 and focus on the row representing individuals who strongly agree that Blacks should "work their way up without special favors." Moving from the least educated to the most educated groups (that is, left to right across the table), the percentages tend to get smaller. The percentage of respondents who strongly agree decreases as education increases. Looking at the percentages of people who somewhat disagree or strongly disagree, we can see that those with a college or graduate degree are more likely to disagree than are those with lower levels of education.

4. Consider limitations.

An important part of all scientific research is to be clear about the limitations of the research. The analyses described above should not be understood as the unquestioned "final word" on the relationship between education and racial ideology. Every research project has limitations, some more than others, and it is important to make these clear when interpreting the results.

When considering the limitations of any survey research project, it is important to consider issues of survey design, possible sample biases, and generalizability. For example, the results here apply only to adults in the US who are aged 18 and older. Since the analyses are restricted to the years 2012 through 2014, respondents include

English- and Spanish-speaking adults, but adults in the US who are not able to speak either language are not included in the analysis. Nor are those living in institutions such as prisons or mental institutions included in this analysis.

It is also important to consider potential **confounding variables**, variables that are not included in the analyses but might be affecting the relationship between the independent and dependent variables. For example, how might differences in respondents' racial/ethnic statuses shape the relationship between education and racial ideology? This question is addressed further in the Exercises, which "control" for respondents' racial/ethnic status when examining the relationship between education and racial ideology.

Finally, while it is crucial to explain what the analyses reveal about the larger research question ("Is education associated with beliefs about racial inequality?"), it is often useful to clarify what the analyses do *not* tell us about the research question. Anticipate possible misinterpretations. Clarify what the findings suggest and what they don't suggest. For example, while respondents with greater levels of education are less likely to believe that Blacks should "work their way up without special favors," almost half of respondents with a graduate degree either agree or strongly agree with this idea (25.3% + 23.7% = 49%). While there does appear to be a relationship between education and racial ideology as it is measured here, it is also important to note that individuals' beliefs about racial issues are complicated and sometimes contradictory. A single indicator of racial ideology likely does not capture the full complexity of how people think about racial issues.

5. Summarize your conclusions.

Interpreting these results within a social justice framework requires thinking through issues of power and inequality, socially constructed differences, links between the micro and macro levels of society, and the importance of intersecting inequalities. We might first ask, why would individuals with lower levels of education be more likely to believe that Blacks should "work their way up without special favors" compared to individuals with higher levels of education? The analyses presented in Figures 4.8 and 4.9 do not take racial/ethnic differences into account. How would the patterns we see in these beliefs change if we were to examine the intersections of race and class? Among every level of education considered, the modal category for WRKWAYUP was either agree strongly or agree somewhat, likely speaking to a broader ideology of the American Dream— the idea that through hard work, persistence, and creative problem-solving, individuals can work their way up in society. Analyzing these results within a social justice framework, we might ask, to what extent are ideals about the American Dream and respondents' reluctance to extend "special favors" to disadvantaged social groups racialized? These questions can be answered by combining further analyses of GSS data with research that is qualitative, historical, comparative, and theoretical.

EXERCISES

To what extent does the relationship between education and racial ideology differ for respondents of different racial/ethnic groups? Figure 4.10 examines the relationship between educational attainment and racial ideology, controlling for respondents' racial/ethnic group. As discussed above, RACEHISP is a nominal-level variable that indicates respondents' self-identification as non-Hispanic white, Black or African American, non-Black Hispanic, or another racial/ethnic minority group.

The top portion of the figure shows the variables used to construct the analysis and indicates that RACEHISP was used as a control variable and YEAR was used as a filter.

1. What is the dependent variable in this analysis?

    a. YEAR

    b. RACEHISP

    c. WRKWAYUP

    d. DEGREE

2. Of the racial/ethnic groups analyzed in this figure, which group has the highest number of respondents?

    a. non-Hispanic whites

    b. Blacks

    c. non-Black Hispanics

    d. other racial/ethnic groups

3. The analysis presented in Figure 4.10 uses a filter that:

    a. restricts the analysis to survey years 2012 through 2014.

    b. restricts the analysis to survey years 2000 through 2014.

    c. restricts the analysis to people who describe themselves as white.

    d. restricts the analysis to people who describe themselves as non-white.

4. Of the racial/ethnic groups analyzed in this figure, which group has the smallest number of respondents?

    a. non-Hispanic whites

    b. Blacks

    c. non-Black Hispanics

    d. other racial/ethnic groups

5. Looking at the first cross-tab in Figure 4.10, the 53.5 in the upper left corner tells us that:

    a. 53.5% of non-Hispanic white respondents strongly agreed that "Blacks should work their way up without special favors."

    b. 53.5% of non-Hispanic whites responded that they had not earned a high school diploma.

    c. Of the people who strongly agreed that "Blacks should work their way up without special favors," 53.5% had less than a high school diploma.

    d. Among non-Hispanic white respondents, 53.5% of those with less than a high school diploma strongly agreed that "Blacks should work their way up without special favors."

6. Approximately how many non-Black Hispanic respondents reported that they had a bachelor's degree but not a graduate degree?

    a. 409

    b. 46

    c. 36

    d. 170

7. Looking at the first cross-tab in Figure 4.10, the 44.2 in the upper right corner tells us that:

    a. 44.2% of all respondents strongly agreed that "Blacks should work their way up without special favors."

    b. 44.2% of non-Hispanic white respondents strongly agreed that "Blacks should work their way up without special favors."

    c. 44.2% of non-Hispanic whites responded that they had earned a graduate degree.

    d. 44.2% of respondents surveyed identified as non-Hispanic whites.

8. Comparing the first and second cross-tabs in Figure 4.10, which of the following statements is correct?

    a. Across all education levels, the percentage of non-Hispanic white respondents who strongly agreed that "Blacks should work their way up without special favors" is higher than the percentage of Black respondents who strongly agreed with this statement.

b. Across all education levels, the percentage of non-Hispanic white respondents who strongly agreed that "Blacks should work their way up without special favors" is lower than the percentage of Black respondents who strongly agreed with this statement.

c. In terms of who is represented in the table, the number of Black respondents with less than a high school degree is larger than the number of white respondents with less than a high school degree.

d. In terms of who is represented in the table, the number of Black respondents with a graduate degree is larger than the number of white respondents with less than a high school degree.

9. Which of the following statements best describes the relationship between educational attainment and racial ideology among non-Black Hispanic respondents?

a. Non-Black Hispanic respondents who have less than a high school education are more likely to "strongly agree" than are non-Black Hispanic respondents with a higher level of education.

b. For all education levels, the modal category for the dependent variable WRKWAYUP is "agree strongly."

c. Non-Black Hispanic respondents with a bachelor's degree or higher are significantly more likely to disagree (strongly or somewhat) with the idea that "Blacks should work their way up without special favors" than are non-Black Hispanic respondents with lower levels of education.

d. The percentage of non-Black Hispanic respondents who "strongly disagree" with the idea that "Blacks should work their way up without special favors" is higher than the percentage of Black respondents who "strongly disagree" with this idea.

10. What limitations are worth noting when interpreting this table?

a. The analysis does not represent individuals who are currently residing in institutions.

b. Individuals who are unable to speak either English or Spanish are not represented in the GSS.

c. Beliefs about whether Blacks should be able to "work their way up without special favors" is only one aspect of racial ideology and likely does not capture the full complexity of how people think about racial issues.

d. All of the above are important limitations for interpreting this table.

Figure 4.10

Variables					
Role	**Name**	**Label**	**Range**	**MD**	**Dataset**
Row	**WRKWAYUP**	BLACKS OVERCOME PREJUDICE WITHOUT FAVORS	1-5	0,8,9	1
Column	**DEGREE**	RS HIGHEST DEGREE	0-4	7,8,9	1
Control	**RACEHISP**	Race with Hispanic (2000 and later)	1-4	9	1
Weight	**COMPWT**	Composite weight = WTSSALL * OVERSAMP * FORMWT	.1913-11.1261		1
Filter	**YEAR(2012-2014)**	GSS YEAR FOR THIS RESPONDENT	1972-2014		1

Statistics for RACEHISP = 1(White)						
		DEGREE				
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW* *TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**53.5** 81.5	**51.2** 506.0	**54.3** 82.7	**28.9** 118.1	**27.2** 57.5	*44.2* *845.8*
	2: AGREE SOMEWHAT	**30.7** 46.8	**28.7** 283.8	**26.4** 40.1	**30.8** 126.0	**22.4** 47.3	*28.4* *544.0*
	3: NEITHER AGREE NOR DISAGREE	**8.9** 13.5	**11.7** 115.8	**12.8** 19.5	**20.8** 85.1	**18.0** 38.0	*14.2* *271.9*
	4: DISAGREE SOMEWHAT	**3.9** 6.0	**5.7** 56.8	**5.4** 8.1	**12.7** 51.9	**21.2** 44.6	*8.8* *167.5*
	5: DISAGREE STRONGLY	**3.0** 4.6	**2.7** 26.8	**1.1** 1.6	**6.9** 28.1	**11.2** 23.6	*4.4* *84.8*
	COL TOTAL	*100.0* *152.5*	*100.0* *989.3*	*100.0* *152.1*	*100.0* *409.2*	*100.0* *211.0*	*100.0* *1,914.0*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

CHAPTER FOUR • ANALYZING RACE AND ETHNICITY WITH THE GSS

Statistics for RACEHISP = 2(Black)							
		DEGREE					
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**35.6** 27.5	**29.9** 69.2	**39.6** 14.9	**22.6** 11.0	**5.4** 1.8	*29.0* *124.3*
	2: AGREE SOMEWHAT	**16.5** 12.7	**24.7** 57.2	**13.2** 5.0	**22.0** 10.7	**26.5** 8.9	*22.0* *94.5*
	3: NEITHER AGREE NOR DISAGREE	**18.6** 14.4	**17.3** 40.1	**21.7** 8.2	**7.9** 3.9	**7.9** 2.6	*16.1* *69.2*
	4: DISAGREE SOMEWHAT	**17.8** 13.7	**14.2** 32.9	**3.6** 1.3	**24.8** 12.1	**27.5** 9.2	*16.1* *69.2*
	5: DISAGREE STRONGLY	**11.6** 9.0	**14.0** 32.3	**21.9** 8.3	**22.6** 11.0	**32.8** 11.0	*16.7* *71.5*
	COL TOTAL	*100.0* *77.3*	*100.0* *231.6*	*100.0* *37.7*	*100.0* *48.6*	*100.0* *33.5*	*100.0* *428.7*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

Statistics for RACEHISP = 3(Hispanic)						
		DEGREE				
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**28.2** 38.0	**44.9** 98.5	**48.0** 10.3	**30.8** 14.2	**36.0** 8.7	*38.1* *169.7*
	2: AGREE SOMEWHAT	**48.2** 64.8	**33.4** 73.2	**12.0** 2.6	**28.7** 13.2	**14.3** 3.5	*35.3* *157.3*
	3: NEITHER AGREE NOR DISAGREE	**17.6** 23.6	**14.1** 31.0	**28.7** 6.2	**13.0** 6.0	**8.1** 2.0	*15.4* *68.8*
	4: DISAGREE SOMEWHAT	**4.7** 6.3	**5.8** 12.7	**4.2** .9	**15.9** 7.3	**21.3** 5.2	*7.3* *32.4*
	5: DISAGREE STRONGLY	**1.3** 1.8	**1.8** 3.9	**7.1** 1.5	**11.6** 5.3	**20.4** 5.0	*3.9* *17.5*
	COL TOTAL	*100.0* *134.6*	*100.0* *219.3*	*100.0* *21.5*	*100.0* *46.0*	*100.0* *24.3*	*100.0* *445.7*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

Statistics for RACEHISP = 4(Other)						
		DEGREE				
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**54.3** 7.0	**29.8** 16.0	**28.5** 5.5	**32.8** 11.8	**24.8** 6.1	*31.7* *46.3*
	2: AGREE SOMEWHAT	**19.4** 2.5	**39.5** 21.2	**60.1** 11.6	**30.1** 10.9	**40.9** 10.0	*38.4* *56.2*
	3: NEITHER AGREE NOR DISAGREE	**10.5** 1.3	**15.1** 8.1	**4.6** .9	**22.3** 8.1	**17.8** 4.3	*15.5* *22.7*
	4: DISAGREE SOMEWHAT	**12.4** 1.6	**10.7** 5.7	**4.6** .9	**3.7** 1.3	**16.6** 4.0	*9.3* *13.6*
	5: DISAGREE STRONGLY	**3.5** .4	**4.9** 2.6	**2.1** .4	**11.0** 4.0	**.0** .0	*5.1* *7.5*
	COL TOTAL	*100.0* *12.8*	*100.0* *53.6*	*100.0* *19.4*	*100.0* *36.1*	*100.0* *24.4*	*100.0* *146.3*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

Statistics for all valid cases						
		DEGREE				
Cells contain: -**Column percent** -Weighted N		0 LT HIGH SCHOOL	1 HIGH SCHOOL	2 JUNIOR COLLEGE	3 BACHELOR	4 GRADUATE	*ROW* *TOTAL*
WRKWAYUP	1: AGREE STRONGLY	**40.8** 153.9	**46.2** 689.7	**49.2** 113.4	**28.7** 155.1	**25.3** 74.1	*40.4* *1,186.2*
	2: AGREE SOMEWHAT	**33.6** 126.8	**29.1** 435.4	**25.7** 59.3	**29.8** 160.8	**23.7** 69.6	*29.0* *851.9*
	3: NEITHER AGREE NOR DISAGREE	**14.0** 52.9	**13.1** 195.0	**15.1** 34.8	**19.1** 103.1	**16.0** 46.9	*14.7* *432.6*
	4: DISAGREE SOMEWHAT	**7.3** 27.6	**7.2** 108.2	**4.9** 11.3	**13.4** 72.6	**21.5** 63.1	*9.6* *282.7*
	5: DISAGREE STRONGLY	**4.2** 15.8	**4.4** 65.7	**5.1** 11.8	**9.0** 48.4	**13.5** 39.6	*6.2* *181.3*
	COL *TOTAL*	*100.0* *377.1*	*100.0* *1,493.9*	*100.0* *230.6*	*100.0* *539.9*	*100.0* *293.2*	*100.0* *2,934.7*

Color coding:	<-2.0	<-1.0	<0.0	>0.0	>1.0	>2.0	**Z**
N in each cell:	Smaller than expected			Larger than expected			

ANALYSES & ESSAYS

1. *What sociodemographic characteristics are related to respondents' beliefs about racial or ethnic inequality?* Identify one variable that you think plays a role in determining beliefs about racial inequality. Identify three questions that assess respondents' beliefs about racial/ethnic inequality. Construct three separate cross-tabs that examine how the sociodemographic characteristic you have chosen relates to these beliefs about racial/ethnic inequality. Interpret your results.

2. *To what extent do people of different racial/ethnic groups have different ideas about citizenship and patriotism?* Create three different cross-tab analyses in which RACEHISP is the independent variable in each one and NTCITVTE, CRIMLOSE, SHORTCOM are the different dependent variables. Interpret your results.

3. *To what extent do people of different racial/ethnic groups differ in their confidence in social institutions?* Create three different cross-tab analyses in which RACEHISP is the independent variable in each one and CONEDUC, CONPRESS, and CONFINAN are the different dependent variables. Use a filter to restrict your analysis to data from the years 2010 through 2014. Interpret your results, giving particular attention to how your findings highlight the connections between the micro and macro levels of society.

4. *To what extent are contemporary marriage and family arrangements structured by race, ethnicity, and gender?* Create a cross-tab where MARITAL is the dependent variable, SEX is the independent variable, and RACEHISP is the control variable. Use a filter to restrict your analysis to data from the years 2010 through 2014. Interpret your results, giving particular attention to the intersections of gender and racial/ethnic inequality.

5. *To what extent do people of different racial/ethnic groups have different experiences at work?* Create three different cross-tab analyses in which RACEHISP is the independent variable in each one and WKRACISM, RESPECT, and SATJOB1 are the different dependent variables. Use a filter to restrict your analysis to data from the years 2010 through 2014. Interpret your results, giving particular attention to how your findings highlight the connections between the micro and macro levels of society.

NOTES

1. See, for example, Zuberi, Tukufu and Eduardo Bonilla-Silva, eds. 2008. *White Logic, White Methods: Racism and Methodology*. Lanham, MD: Rowman & Littlefield.

2. This perspective is dominant in contemporary France, for example, and as a result there are relatively little governmental data available for documenting racial inequality.

3. American Sociological Association. 2003. *The Importance of Collecting Data and Doing Social Scientific Research on Race*. Washington, DC: American Sociological Association. (Emphasis added.)

4. Describing the value of the GSS for analyzing racial/ethnic relations in US society, Tom W. Smith (2002:7) writes that the GSS includes "359 different items a total of 836 times, from 1972 to 2000 . . ." Smith reports that there are at least 69 questions related to intergroup relations that have been asked regularly over time, and there are a number of special modules focusing on intergroup relations. Smith, Tom W. 2002. "GSS Methodological Report No. 96 Measuring Racial and Ethnic Discrimination." National Opinion Research Center University of Chicago. Retrieved September 18, 2016 (http://gss.norc.org/Documents/reports/methodological-reports/MR096.pdf).

5. See "FAQ #12: What Happened to Information on Race After 2000?" (http://www3.norc.org/GSS+Website/FAQs/).

6. Ibid.

7. The variable USCITZN also asks about citizenship but uses slightly different wording and was asked to an even smaller subset of cases in these years.

8. There are a handful of instances where viewing the variable in SDA does not produce the precise wording of the survey question. The variable WRKWAYUP is one such example. As shown on the GSS Data Explorer, the full question reads, "Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement? (HAND CARD TO RESPONDENT) Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without special favors" (https://gssdataexplorer.norc.org/variables/424/vshow).

9. For more information on weighting options, see Chapter 2.

10. For more information, see http://www.icpsr.umich.edu/icpsrweb/content/SAMHDA/help/helpan.htm#Tcolor.

11. Appendix D of the GSS Cumulative Codebook (2015) provides the precise questions that respondents were asked. To generate the value of degree, for example, respondents were asked, "What is the highest grade in elementary school or high school that you finished and got credit for? Did you ever get a high school diploma or a GED certificate? Did you ever complete one or more years of college for credit—not including schooling such as business college, technical or vocational school? If yes, how many years did you complete? Do you have any college degrees? If yes, what degree or degrees?"