Chapter 18: Categorical data

Labcoat Leni's Real Research

The impact of sexualized images on women's self-evaluations

Problem

Daniels, E., A. (2012). *Journal of Applied Developmental Psychology*, *33*, 79–90.

Women (and increasingly men) are constantly bombarded with images of 'idealized' women in the media and there is a growing concern about how these images affect our perceptions of ourselves. Daniels (2012) conducted an interesting study in which she showed young women images of successful female athletes (e.g., Anna Kournikova) that were either images of them playing sport (performance athlete images) or images of them posing in bathing suits (sexualized images). Participants completed a short writing exercise after viewing these types of images. Each participant saw only one type of image, but several examples. Daniels then coded these written exercises and identified themes, one of which was whether women commented on their own appearance or attractiveness. Daniels hypothesized that women who viewed the sexualized images (n = 140) would self-objectify (i.e., this theme would be present in what they wrote) more than those who viewed the performance athlete pictures (n = 117, despite what the participants section of the paper implies). These are the frequencies:

	Theme Present	Theme Absent	Total
Performance athletes	20	97	117
Sexualized athletes	56	84	140

Labcoat Leni wants you to enter these data in SPSS and test Daniel's hypothesis that there is an association between the type of image viewed, and whether or not the women commented on their own appearance/attractiveness in their writing exercise (**Daniels (2012).sav**).

Solution

Because the frequency data have been entered into SPSS, we must tell the computer that the variable **Self_Evaluation** represents the number of cases that fell into a particular combination of categories. To do this, access the *Weight Cases* dialog box by selecting Data Weight Cases... . Select Weight cases by and then drag the variable in which the number of cases is specified (in this case **Self_Evaluation**) to the box labelled *Frequency variable* (or click on). This process tells the computer that it should weight each category combination by the number in the column labelled **Self_Evaluation**. Your completed dialog box should look like Figure .

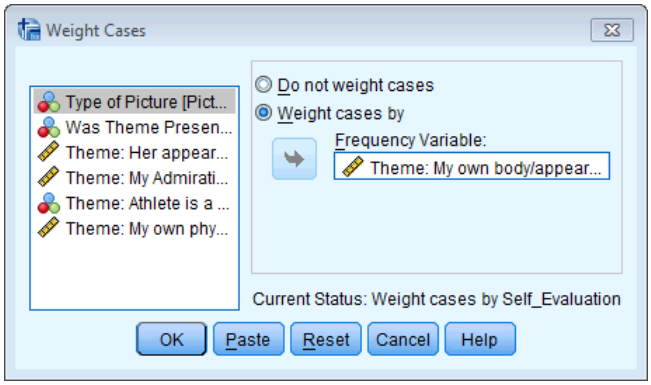

Figure 1

Next, select Analyze Descriptive Statistics Crosstabs... . First, drag one of the variables of interest from the variable list to the box labelled *Row(s)* (or select it and click on). For this example, I selected **Type of Picture** to be the rows of the table. Next, drag the other variable of interest (**Was Theme Present or Absent in what participant wrote**) to the box labelled *Column(s)* (or select it and click on). If you click on Statistics... a dialog box appears in which you can specify various statistical tests. Select the chi-square test, the contingency coefficient, phi and lambda and then click on Continue. If you click on Cells... a dialog box appears in which you can specify the type of data displayed in the crosstabulation table. It is important that you ask for expected counts because this is how we check the assumptions about the expected frequencies. It is also useful to have a look at the row, column and total percentages because these values are usually more easily interpreted than the actual frequencies and provide some idea of the origin of any significant effects. There are two other options that are useful for breaking down a significant effect (should we get one): (1) we can select a *z*-test to compare cell counts across columns of the contingency table (by checking Compare column proportions), and if we do we should use a Bonferroni correction (check Adjust p-values (Bonferroni method)); and (2) select standardized residuals. Once these options have been selected, click on Continue to return to the main dialog box. From here you can click on Exact... to compute Fisher's exact test if your sample is small or if your expected frequencies are too low. Select the *Exact* test option; we don't really need it for these data, but it will be a useful way to see how it is used. Click on Continue to return to the main dialog box and then click on OK to run the analysis (your completed dialog boxes should look like those in Figure).

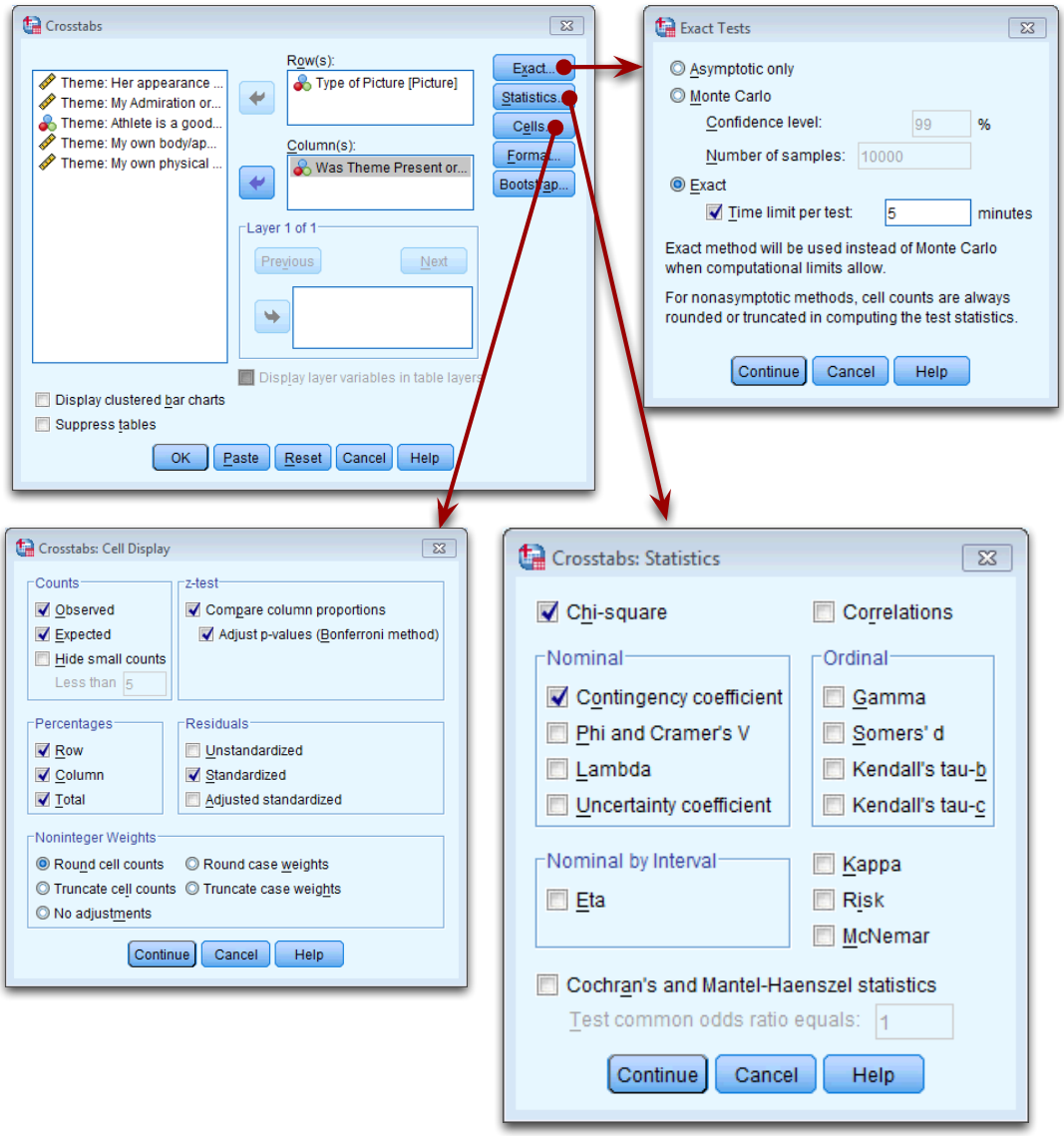
Crosstabs
Theme: Her appearance ...
Theme: My Admiration or...
Theme: Athlete is a good...
Theme: My own body/ap...
Theme: My own physical ...
Row(s):
Type of Picture [Picture]
Column(s):
Was Theme Present or...
Layer 1 of 1
Previous
Next
Display layer variables in table layers
Display clustered bar charts
Suppress tables
OK
Paste
Reset
Cancel
Help
Exact...
Statistics...
Cells...
Format...
Bootstrap...
Exact Tests
Asymptotic only
Monte Carlo
Confidence level:
99
%
Number of samples:
10000
Exact
Time limit per test:
5
minutes
Exact method will be used instead of Monte Carlo when computational limits allow.
For nonasymptotic methods, cell counts are always rounded or truncated in computing the test statistics.
Continue
Cancel
Help
Crosstabs: Cell Display
Counts
Observed
Expected
Hide small counts
Less than
5
z-test
Compare column proportions
Adjust p-values (Bonferroni method)
Percentages
Row
Column
Total
Residuals
Unstandardized
Standardized
Adjusted standardized
Noninteger Weights
Round cell counts
Round case weights
Truncate cell counts
Truncate case weights
No adjustments
Continue
Cancel
Help
Crosstabs: Statistics
Chi-square
Correlations
Nominal
Contingency coefficient
Phi and Cramer's V
Lambda
Uncertainty coefficient
Ordinal
Gamma
Somers' d
Kendall's tau-b
Kendall's tau-c
Nominal by Interval
Eta
Kappa
Risk
McNemar
Cochran's and Mantel-Haenszel statistics
Test common odds ratio equals:
1
Continue
Cancel
Help

Figure 2

Output

Type of Picture * Was Theme Present or Absent in what participant wrote? Crosstabulation[1]

			Was Theme Present or Absent in what participant wrote?		
			Absent	Present	Total
Type of Picture	Performance Athletes	Count	97a	20b	117
		Expected Count	82.4	34.6	117.0
		% within Type of Picture	82.9%	17.1%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	53.6%	26.3%	45.5%
		Std. Residual	1.6	-2.5	
	Sexualized Athletes	Count	84a	56b	140
		Expected Count	98.6	41.4	140.0
		% within Type of Picture	60.0%	40.0%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	46.4%	73.7%	54.5%
		Std. Residual	-1.5	2.3	
Total		Count	181	76	257
		Expected Count	181.0	76.0	257.0
		% within Type of Picture	70.4%	29.6%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	100.0%	100.0%	100.0%

Each subscript letter denotes a subset of Was Theme Present or Absent in what participant wrote? categories whose column proportions do not differ significantly from each other at the .05 level.

1. Type of Theme = Self-evaluation

Output 1

First, let's check that the expected frequencies assumption has been met. We have a 2×2 table, so all expected frequencies need to be greater than 5. If you look at the expected counts in the contingency table, we see that the smallest expected count is 34.6 (for women who saw pictures of performance athletes and did self-evaluate). This value exceeds 5 and so the assumption has been met.

The other thing to note about this table is that because we selected *Compare column proportions* our counts have subscript letters. For example, in the row labelled *Performance Athletes* the count of 97 has a subscript letter *a* and the count of 20 has a subscript letter *b*. These subscripts tell us the results of the *z*-test that we asked for: columns with different subscripts have significantly different column proportions. We need to look within rows of the table. So, for *Performance Athletes* the columns have different subscripts as I just explained, which means that proportions within the column variable (i.e., *Was the theme present or absent in what they wrote?*) are significantly different. The *z*-test compares the *proportion* of the total frequency of the first column that falls into the first row against the *proportion* of the total frequency of the second column that falls into the first row. So, of all the women who did self-evaluate (theme present), 26.3% saw pictures of performance athletes, and of all the women who didn't self-evaluate (theme absent), 53.6% saw pictures of performance athletes. The different subscripts tell us that these proportions are significantly different. Put another way, the proportion of women who self-evaluated after seeing pictures of performance athletes was significantly less than the proportion who didn't self-evaluate after seeing pictures of performance athletes.

If we move on to the row labelled *Sexualized Athletes*, the count of 84 has a subscript letter *a* and the count of 56 has a subscript letter *b*; as before, the fact they have different letters tells us that the column proportions are significantly different. The proportion of

women who self-evaluated after seeing sexualized pictures of female athletes (73.7%) was significantly greater than the proportion who didn't self-evaluate after seeing sexualized pictures of female athletes (46.4%).

As we saw earlier, Pearson's chi-square test examines whether there is an association between two categorical variables (in this case the type of picture and whether the women self-evaluated or not). The value of the chi-square statistic is 16.057. This value is highly significant ($p < .001$), indicating that the type of picture used had a significant effect on whether women self-evaluated.

Underneath the chi-square table there are several footnotes relating to the assumption that expected counts should be greater than 5. If you forgot to check this assumption yourself, SPSS kindly gives a summary of the number of expected counts below 5. In this case, there were no expected frequencies less than 5, so we know that the chi-square statistic should be accurate.

Chi-Square Tests[a]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	16.057[b]	1	.000		
Continuity Correction[c]	14.976	1	.000		
Likelihood Ratio	16.629	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	15.994	1	.000		
N of Valid Cases	257				

a. Type of Theme = Self-evaluation

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 34.60.

c. Computed only for a 2x2 table

Output 2

The highly significant result indicates that there is an association between the type of picture and whether women self-evaluated or not. In other words, the pattern of responses (i.e., the proportion of women who self-evaluated to the proportion who did not) in the two picture conditions is significantly different.

Below is an excerpt from Daniels's (2012) conclusions:

> *My body/looks.* Hypothesis 3 predicted that participants in the sexualized athlete and sexualized model conditions would self-objectify more than participants in the performance athlete condition. As expected, participants who saw the sexualized athletes (40.0%) made more *self-evaluation* statements than did participants who saw the performance athletes (17.1%), χ^2 (1, n = 257) = 16.06, $p < .001$. Cramer's V = .25. The planned chi-square analysis to inves-

Is the black American happy?

Problem

Beckham, A. S. (1929). *Journal of Abnormal and Social Psychology*, *24*, 186–190.

When I was doing my psychology degree I spent a lot of time reading about the civil rights movement in the USA. Although I was supposed to be reading psychology, I became more interested in Malcolm X and Martin Luther King Jr. This is why I find Beckham's 1929 study of black Americans such an interesting piece of research. Beckham was a black American academic who founded the psychology laboratory at Howard University, Washington, DC. His wife Ruth was the first black woman ever to be awarded a Ph.D. (also in psychology) at the University of Minnesota. To put some context on Beckham's study, it was published 36 years before the Jim Crow laws were finally overthrown by the Civil Rights Act of 1964, and at a time when black Americans were segregated, openly discriminated against and were victims of the most abominable violations of civil liberties and human rights. For a richer context I suggest reading James Baldwin's superb novel *The fire next time*. Even the language of the study and the data from it are an uncomfortable reminder of the era in which it was conducted.

Beckham sought to measure the psychological state of black Americans with three questions put to 3443 black Americans from different walks of life. He asked them whether they thought black Americans were happy, whether they personally were happy as a black American, and whether black Americans *should* be happy. They could answer only *yes* or *no* to each question. Beckham did no formal statistical analysis of his data (Fisher's article containing the popularized version of the chi-square test was published only 7 years earlier in a statistics journal that would not have been read by psychologists). I love this study, though, because it demonstrates that you do not need elaborate methods to answer important and far-reaching questions; with just three questions, Beckham told the world an enormous amount about very real and important psychological and sociological phenomena.

The frequency data (number of yes and no responses within each employment category) from this study are in the file **Beckham(1929).sav**. Labcoat Leni wants you to carry out three chi-square tests (one for each question that was asked). What conclusions can you draw?

Are black Americans happy?

Let's run the analysis on the first question. First we must remember to tell SPSS which variable contains the frequencies by using the *weight cases* command. Select Data Weight Cases..., then in the resulting dialog box select Weight cases by and then select the variable in which the number of cases is specified (in this case **Happy**) and drag it to the box labelled *Frequency variable* (or click on). This process tells the computer that it should weight each category combination by the number in the column labelled **happy**.

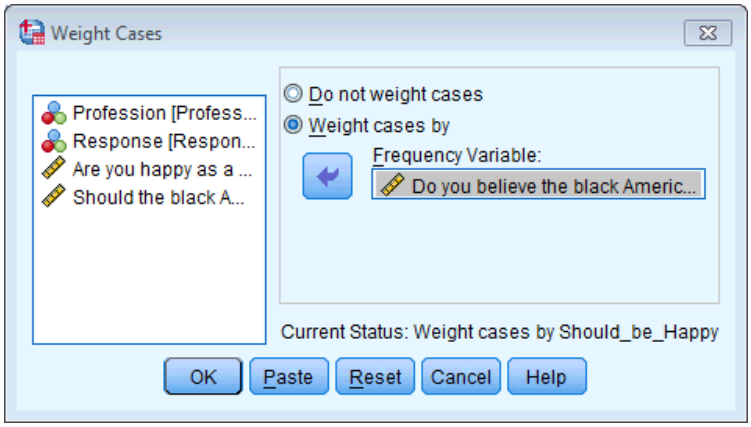

Figure 3

To conduct the chi-square test, use the *crosstabs* command by selecting Analyze Descriptive Statistics Crosstabs... . We have two variables in our crosstabulation table: the occupation of the participant (**Profession**) and whether they responded yes or no to the question (**Response**). Select one of these variables and drag it into the box labelled *Row(s)* (or click on). For this example, I selected **Profession** to be the rows of the table. Next, select the other variable of interest (**Response**) and drag it to the box labelled *Column(s)* (or click on). Use the book chapter to select other appropriate options (we do not need to use the exact test used in the chapter because our sample size is very large; however, you could choose a Monte Carlo test of significance if you like).

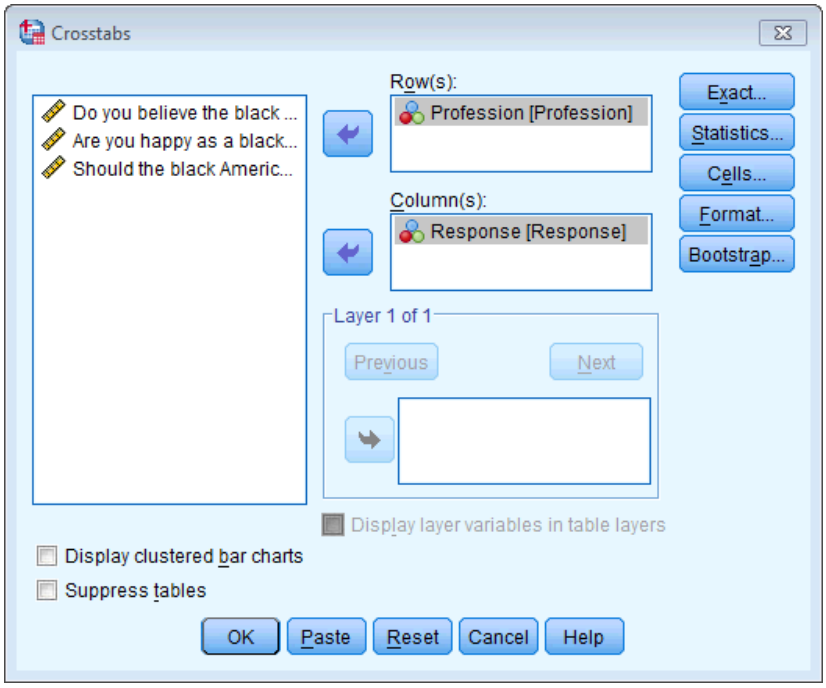

Figure 4

Profession * Response Crosstabulation

			Response		
			No	Yes	Total
Profession	College Students	Count	1610	390	2000
		Expected Count	1316.3	683.7	2000.0
		% within Profession	80.5%	19.5%	100.0%
		% within Response	70.3%	32.8%	57.5%
		Std. Residual	8.1	-11.2	
	Unskilled Laborers	Count	122	378	500
		Expected Count	329.1	170.9	500.0
		% within Profession	24.4%	75.6%	100.0%
		% within Response	5.3%	31.8%	14.4%
		Std. Residual	-11.4	15.8	
	Preachers	Count	265	35	300
		Expected Count	197.4	102.6	300.0
		% within Profession	88.3%	11.7%	100.0%
		% within Response	11.6%	2.9%	8.6%
		Std. Residual	4.8	-6.7	
	Physicians	Count	51	159	210
		Expected Count	138.2	71.8	210.0
		% within Profession	24.3%	75.7%	100.0%
		% within Response	2.2%	13.4%	6.0%
		Std. Residual	-7.4	10.3	
	Housewives	Count	122	78	200
		Expected Count	131.6	68.4	200.0
		% within Profession	61.0%	39.0%	100.0%
		% within Response	5.3%	6.6%	5.7%
		Std. Residual	-.8	1.2	
	School Teachers	Count	38	108	146
		Expected Count	96.1	49.9	146.0
		% within Profession	26.0%	74.0%	100.0%
		% within Response	1.7%	9.1%	4.2%
		Std. Residual	-5.9	8.2	
	Lawyers	Count	64	11	75
		Expected Count	49.4	25.6	75.0
		% within Profession	85.3%	14.7%	100.0%
		% within Response	2.8%	.9%	2.2%
		Std. Residual	2.1	-2.9	
	Musician	Count	19	31	50
		Expected Count	32.9	17.1	50.0
		% within Profession	38.0%	62.0%	100.0%
		% within Response	.8%	2.6%	1.4%
		Std. Residual	-2.4	3.4	
	Total	Count	2291	1190	3481
		Expected Count	2291.0	1190.0	3481.0
		% within Profession	65.8%	34.2%	100.0%
		% within Response	100.0%	100.0%	100.0%

Output 3

Chi-Square Tests

				Monte Carlo Sig. (2-sided)			Monte Carlo Sig. (1-sided)		
					99% Confidence Interval			99% Confidence Interval	
	Value	df	Asymp. Sig. (2-sided)	Sig.	Lower Bound	Upper Bound	Sig.	Lower Bound	Upper Bound
Pearson Chi-Square	936.139[a]	7	.000	.000[b]	.000	.000			
Likelihood Ratio	929.369	7	.000	.000[b]	.000	.000			
Fisher's Exact Test	927.231			.000[b]	.000	.000			
Linear-by-Linear Association	184.891[c]	1	.000	.000[b]	.000	.000	.000[b]	.000	.000
N of Valid Cases	3481								

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.09.

b. Based on 10000 sampled tables with starting seed 2000000.

c. The standardized statistic is 13.597.

Output 4

The chi-square test is highly significant, $\chi^2(7) = 936.14$, $p < .001$. This indicates that the profile of yes and no responses differed across the professions. Looking at the standardized residuals, the only profession for which these are non-significant are housewives who showed a fairly even split of whether they thought black Americans were happy (40%) or not (60%). Within the other professions all of the standardized residuals are much higher than 1.96, so how can we make sense of the data? What's interesting is to look at the direction of these residuals (i.e., whether they are positive or negative). For the following professions the residual for 'no' was positive but for 'yes' was negative; these are therefore people who responded more than we would expect that black Americans were not happy and less than expected that black Americans were happy: college students, preachers and lawyers. The remaining professions (labourers, physicians, school teachers and musicians) show the opposite pattern: the residual for 'no' was negative but for 'yes' was positive; these are, therefore, people who responded less than we would expect that black Americans were not happy and more than expected that black Americans were happy.

Are they Happy as black Americans?

We run this analysis in exactly the same way except that we now have to weight the cases by the variable **You_Happy**. Select Data Weight Cases...; then in the resulting dialog box Weight cases by should already be selected from the previous analysis. Select the variable in the box labelled *Frequency Variable* and click on to move it back to the variable list and clear the box. Then, we need to select the variable in which the number of cases is specified (in this case **You_Happy**) and drag it to the box labelled *Frequency Variable* (or click on). This process tells the computer that it should weight each category combination by the number in the column labelled **You_Happy**. Then carry out the analysis through *crosstabs* exactly as before.

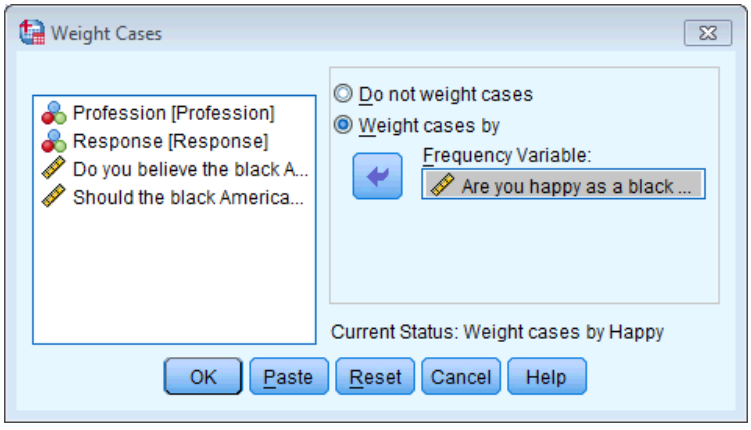

Figure 5

Profession * Response Crosstabulation

			Response		
			No	Yes	Total
Profession	College Students	Count	48	1822	1870
		Expected Count	271.0	1599.0	1870.0
		% within Profession	2.6%	97.4%	100.0%
		% within Response	10.5%	67.4%	59.2%
		Std. Residual	-13.5	5.6	
	Unskilled Laborers	Count	195	305	500
		Expected Count	72.5	427.5	500.0
		% within Profession	39.0%	61.0%	100.0%
		% within Response	42.6%	11.3%	15.8%
		Std. Residual	14.4	-5.9	
	Preachers	Count	0	230	230
		Expected Count	33.3	196.7	230.0
		% within Profession	.0%	100.0%	100.0%
		% within Response	.0%	8.5%	7.3%
		Std. Residual	-5.8	2.4	
	Physicians	Count	7	203	210
		Expected Count	30.4	179.6	210.0
		% within Profession	3.3%	96.7%	100.0%
		% within Response	1.5%	7.5%	6.6%
		Std. Residual	-4.2	1.7	
	Housewives	Count	146	17	163
		Expected Count	23.6	139.4	163.0
		% within Profession	89.6%	10.4%	100.0%
		% within Response	31.9%	.6%	5.2%
		Std. Residual	25.2	-10.4	
	School Teachers	Count	28	79	107
		Expected Count	15.5	91.5	107.0
		% within Profession	26.2%	73.8%	100.0%
		% within Response	6.1%	2.9%	3.4%
		Std. Residual	3.2	-1.3	
	Lawyers	Count	0	30	30
		Expected Count	4.3	25.7	30.0
		% within Profession	.0%	100.0%	100.0%
		% within Response	.0%	1.1%	.9%
		Std. Residual	-2.1	.9	
	Musician	Count	34	16	50
		Expected Count	7.2	42.8	50.0
		% within Profession	68.0%	32.0%	100.0%
		% within Response	7.4%	.6%	1.6%
		Std. Residual	9.9	-4.1	
	Total	Count	458	2702	3160
		Expected Count	458.0	2702.0	3160.0
		% within Profession	14.5%	85.5%	100.0%
		% within Response	100.0%	100.0%	100.0%

Output 5

Chi-Square Tests

				Monte Carlo Sig. (2-sided)			Monte Carlo Sig. (1-sided)		
					99% Confidence Interval		4	99% Confidence Interval	
	Value	df	Asymp. Sig. (2-sided)	Sig.	Lower Bound	Upper Bound	Sig.	Lower Bound	Upper Bound
Pearson Chi-Square	1390.740[a]	7	.000	.000[b]	.000	.000			
Likelihood Ratio	1144.171	7	.000	.000[b]	.000	.000			
Fisher's Exact Test	1134.318			.000[b]	.000	.000			
Linear-by-Linear Association	454.605[c]	1	.000	.000[b]	.000	.000	.000[b]	.000	.000
N of Valid Cases	3160								

a. 1 cells (6.3%) have expected count less than 5. The minimum expected count is 4.35.

b. Based on 10000 sampled tables with starting seed 957002199.

c. The standardized statistic is -21.321.

Output 6

The chi-square test is highly significant, $\chi^2(7) = 1390.74$, $p < .001$. This indicates that the profile of yes and no responses differed across the professions. Looking at the standardized residuals, these are significant in most cells with a few exceptions: physicians, lawyers and school teachers saying 'yes'. Within the other cells all of the standardized residuals are much higher than 1.96. Again, we can look at the direction of these residuals (i.e., whether they are positive or negative). For labourers, housewives, school teachers and musicians the residual for 'no' was positive but for 'yes' was negative; these are, therefore, people who responded more than we would expect that they were not happy as black Americans and less than expected that they were happy as black Americans. The remaining professions (college students, physicians, preachers and lawyers) show the opposite pattern: the residual for 'no' was negative but for 'yes' was positive; these are, therefore, people who responded less than we would expect that they were not happy as black Americans and more than expected that they were happy as black Americans. Essentially, the former group are in low-paid jobs in which conditions would have been very hard (especially in the social context of the time). The latter group are in much more respected (and probably better-paid) professions. Therefore, the responses to this question could say more about the professions of the people asked than their views of being black Americans.

Should black Americans be happy?

We run this analysis in exactly the same way except that we now have to weight the cases by the variable **Should_Be_Happy**. Select Data Weight Cases...; then in the resulting dialog box Weight cases by should already be selected from the previous analysis. Select the variable in the box labelled *Frequency Variable* and click on  to move it back to the variable list and clear the box. Then, we need to select the variable in which the number of cases is specified (in this case **Should_Be_Happy**) and drag it to the box labelled *Frequency Variable* (or click on ). This process tells the computer that it should weight each category combination by the number in the column labelled **Should_Be_Happy**. Then carry out the analysis through *crosstabs* exactly as before.

Profession * Response Crosstabulation

			Response		
			No	Yes	Total
Profession	College Students	Count	1810	141	1951
		Expected Count	1270.8	680.2	1951.0
		% within Profession	92.8%	7.2%	100.0%
		% within Response	81.9%	11.9%	57.5%
		Std. Residual	15.1	-20.7	
	Unskilled Laborers	Count	104	396	500
		Expected Count	325.7	174.3	500.0
		% within Profession	20.8%	79.2%	100.0%
		% within Response	4.7%	33.5%	14.7%
		Std. Residual	-12.3	16.8	
	Preachers	Count	36	264	300
		Expected Count	195.4	104.6	300.0
		% within Profession	12.0%	88.0%	100.0%
		% within Response	1.6%	22.3%	8.8%
		Std. Residual	-11.4	15.6	
	Physicians	Count	36	174	210
		Expected Count	136.8	73.2	210.0
		% within Profession	17.1%	82.9%	100.0%
		% within Response	1.6%	14.7%	6.2%
		Std. Residual	-8.6	11.8	
	Housewives	Count	120	90	210
		Expected Count	136.8	73.2	210.0
		% within Profession	57.1%	42.9%	100.0%
		% within Response	5.4%	7.6%	6.2%
		Std. Residual	-1.4	2.0	
	School Teachers	Count	33	75	108
		Expected Count	70.3	37.7	108.0
		% within Profession	30.6%	69.4%	100.0%
		% within Response	1.5%	6.3%	3.2%
		Std. Residual	-4.5	6.1	
	Lawyers	Count	57	7	64
		Expected Count	41.7	22.3	64.0
		% within Profession	89.1%	10.9%	100.0%
		% within Response	2.6%	.6%	1.9%
		Std. Residual	2.4	-3.2	
	Musician	Count	14	36	50
		Expected Count	32.6	17.4	50.0
		% within Profession	28.0%	72.0%	100.0%
		% within Response	.6%	3.0%	1.5%
		Std. Residual	-3.3	4.4	
	Total	Count	2210	1183	3393
		Expected Count	2210.0	1183.0	3393.0
		% within Profession	65.1%	34.9%	100.0%
		% within Response	100.0%	100.0%	100.0%

Output 7

Chi-Square Tests

				Monte Carlo Sig. (2-sided)			Monte Carlo Sig. (1-sided)		
					99% Confidence Interval		4	99% Confidence Interval	
	Value	df	Asymp. Sig. (2-sided)	Sig.	Lower Bound	Upper Bound	Sig.	Lower Bound	Upper Bound
Pearson Chi-Square	1784.226[a]	7	.000	.000[b]	.000	.000			
Likelihood Ratio	1928.327	7	.000	.000[b]	.000	.000			
Fisher's Exact Test	1924.651			.000[b]	.000	.000			
Linear-by-Linear Association	564.081[c]	1	.000	.000[b]	.000	.000	.000[b]	.000	.000
N of Valid Cases	3393								

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 17.43.

b. Based on 10000 sampled tables with starting seed 1993510611.

c. The standardized statistic is 23.750.

Output 8

The chi-square test is highly significant, $\chi^2(7) = 1784.23$, $p < .001$. This indicates that the profile of yes and no responses differed across the professions. Looking at the standardized residuals, these are nearly all significant. Again, we can look at the direction of these residuals (i.e., whether they are positive or negative). For college students and lawyers the residual for 'no' was positive but for 'yes' was negative; these are, therefore, people who responded more than we would expect that they thought that black Americans should not be happy and less than expected that they thought black Americans should be happy. The remaining professions show the opposite pattern: the residual for 'no' was negative but for 'yes' was positive; these are, therefore, people who responded less than we would expect that they did not think that black Americans should be happy and more than expected that they thought that black Americans should be happy.

What is interesting here and in the first question is that college students and lawyers are in vocations in which they are expected to be critical about the world. Lawyers may well have defended black Americans who had been the subject of injustice and discrimination or racial abuse, and college students would likely be applying their critically trained minds to the immense social injustice that prevailed at the time. Therefore, these groups can see that their racial group should not be happy and should strive for the equitable and just society to which they are entitled. People in the other professions perhaps adopt a different social comparison.

It's also possible for this final question that the groups interpreted the question differently: perhaps the lawyers and students interpreted the question as 'should they be happy given the political and social conditions of the time?', while the others interpreted the question as 'do they *deserve* happiness?'

It might seem strange to have picked a piece of research from so long ago to illustrate the chi-square test, but what I wanted to demonstrate is that simple research can sometimes be incredibly illuminating. This study asked three simple questions, yet the data are utterly fascinating. It raises further hypotheses that could be tested, it unearths very different views in different professions, and it illuminates a very important social and psychological issue. There are others studies that sometimes use the most elegant paradigms and the highly complex methodologies, but the questions they address are utterly meaningless for the real world. They miss the big picture. Albert Beckham was a remarkable man, trying to understand important and big real-world issues that mattered to hundreds of thousands of people.