Chapter 18: Categorical data

Smart Alex's Solutions

Task 1

*Research suggests that people who can switch off from work (**Detachment**) during off-hours are more satisfied with life and have fewer symptoms of psychological strain (Sonnentag, 2012). Factors at work can affect your ability to detach when away from work. For example, a study looked at 1709 Swiss and German employees and measured job stress in terms of time pressure (**Time_Pressure**) at work (no time pressure, low, medium, high and very high time pressure). Data generated to approximate Figure 1 in Sonnentag (2012) are in the file **Sonnentag (2012).sav**. Carry out a chi-square test to see if time pressure is associated with the ability to detach from work.*

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 **Frequency**) 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 **Frequency** (see Figure 1).

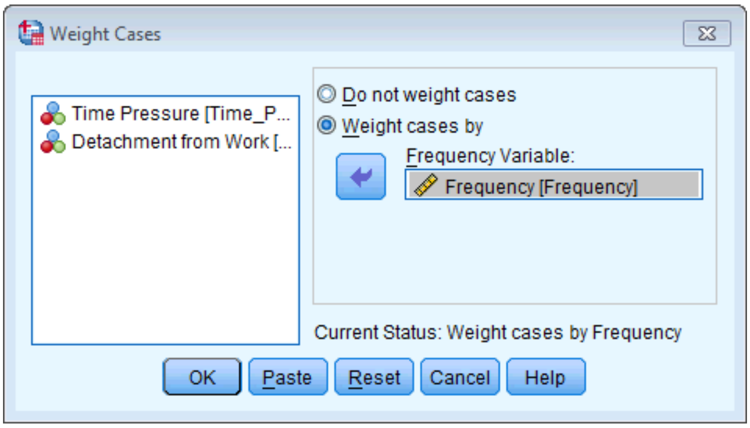

Figure 1

To conduct the chi-square test, use the *crosstabs* command by selecting Analyze Descriptive Statistics ▸ Crosstabs... . We have two variables in our crosstabulation table: **Detachment** and **Time pressure**. Select one of these variables and drag it into the box labelled *Row(s)* (or click on ➡). For this example, I selected **Time Pressure** to be the rows of the table. Next, select the other variable of interest (**Detachment**) and drag it to the box labelled *Column(s)* (or click on ➡). Use the book chapter to select other appropriate options.

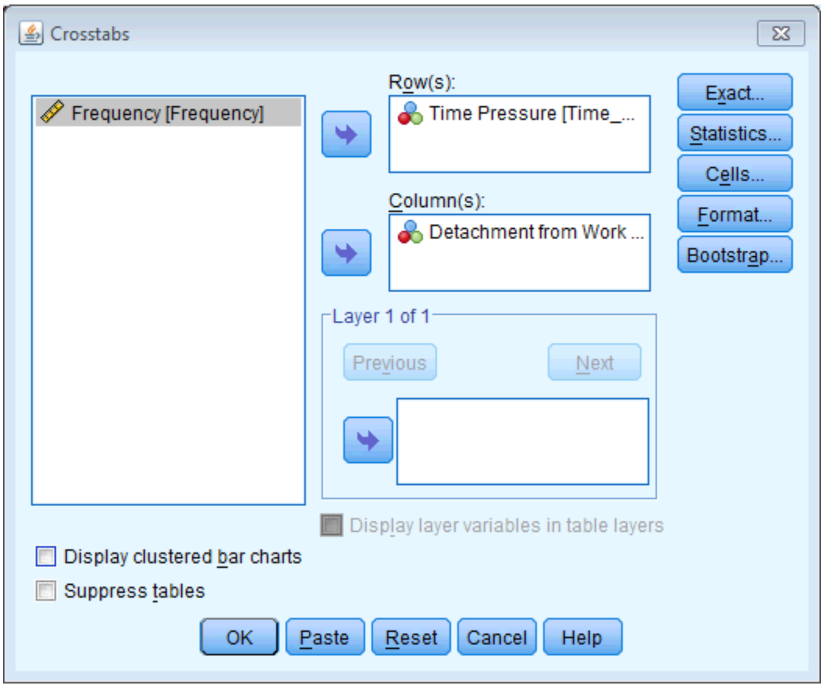

Figure 2

Time Pressure * Detachment from Work Crosstabulation

			Detachment from Work: Low Detachment	Detachment from Work: Very Low Detachment	Total
Time Pressure	No Time Pressure	Count	84a	94a	178
		Expected Count	75.5	102.5	178.0
		% within Time Pressure	47.2%	52.8%	100.0%
		% within Detachment from Work	11.6%	9.6%	10.4%
		% of Total	4.9%	5.5%	10.4%
		Std. Residual	1.0	-.8	
	Low Time Pressure	Count	89a	94a	183
		Expected Count	77.6	105.4	183.0
		% within Time Pressure	48.6%	51.4%	100.0%
		% within Detachment from Work	12.3%	9.6%	10.7%
		% of Total	5.2%	5.5%	10.7%
		Std. Residual	1.3	-1.1	
	Medium Time Pressure	Count	147a	175a	322
		Expected Count	136.6	185.4	322.0
		% within Time Pressure	45.7%	54.3%	100.0%
		% within Detachment from Work	20.3%	17.8%	18.8%
		% of Total	8.6%	10.2%	18.8%
		Std. Residual	.9	-.8	
	High Time Pressure	Count	206a	267a	473
		Expected Count	200.7	272.3	473.0
		% within Time Pressure	43.6%	56.4%	100.0%
		% within Detachment from Work	28.4%	27.1%	27.7%
		% of Total	12.1%	15.6%	27.7%
		Std. Residual	.4	-.3	
	Very High Time Pressure	Count	199a	354b	553
		Expected Count	234.6	318.4	553.0
		% within Time Pressure	36.0%	64.0%	100.0%
		% within Detachment from Work	27.4%	36.0%	32.4%
		% of Total	11.6%	20.7%	32.4%
		Std. Residual	-2.3	2.0	
Total		Count	725	984	1709
		Expected Count	725.0	984.0	1709.0
		% within Time Pressure	42.4%	57.6%	100.0%
		% within Detachment from Work	100.0%	100.0%	100.0%
		% of Total	42.4%	57.6%	100.0%

Each subscript letter denotes a subset of Detachment from Work categories whose column proportions do not differ significantly from each other at the .05 level.

Output 1

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	15.550[a]	4	.004	.004		
Likelihood Ratio	15.654	4	.004	.004		
Fisher's Exact Test	15.669			.003		
Linear-by-Linear Association	12.318[b]	1	.000	.000	.000	.000
N of Valid Cases	1709					

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 75.51.

b. The standardized statistic is 3.510.

Output 2

The chi-square test is highly significant, $\chi^2(4) = 15.55$, $p = .004$. This indicates that the profile of low-detachment and very low-detachment responses differed across different time pressures. Looking at the standardized residuals, the only time pressure for which these are significant is very high time pressure, which showed the greatest split of whether the employees experienced low detachment (36%) or very low detachment (64%). Within the other time pressure groups all of the standardized residuals are lower 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 all time pressure groups except very high time pressure, the residual for 'low detachment' was positive but for 'very low detachment' was negative; these are, therefore, people who responded more than we would expect that they experienced low detachment from work and less than expected that they experienced very low detachment from work. It was only under very high time pressure that opposite pattern occurred: the residual for 'low detachment' was negative but for 'very low detachment' was positive; these are, therefore, people who responded less than we would expect that they experienced low detachment from work and more than expected that they experienced very low detachment from work. In short, there are similar numbers of people who experience low detachment and very low detachment from work when there is no time pressure, low time pressure, medium time pressure and high time pressure. However, when time pressure was very high, significantly more people experienced very low detachment than low detachment.

Task 2

*Labcoat Leni describes a study (Daniels, 2012) that looked at the impact of sexualized images of atheletes compared to performance pictures on women's perceptions of the athletes and themselves. Women looked at different types of pictures (**Picture**) and then did a writing task. Daniels identified whether certain themes were present or absent in each written piece (**Theme_Present**). We have already looked at the self-evaluation theme, but Daniels also idetified others including: commenting on the athlete's body/appearance (**Athletes_Body**), indicating admiration or jealousy for the athlete (**Admiration**), indicating that the athlete was a role model or motivating (**Role_Model**), and their own physical activity (**Self_Physical_Activity**). The data are in the file **Daniels (2012).sav**. Carry out a chi-square test to see whether the type of picture viewed was associated with commenting on the athlete's body/appearance.*

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 **Her appearance and attractiveness**) 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 **Her appearance and attractiveness**.

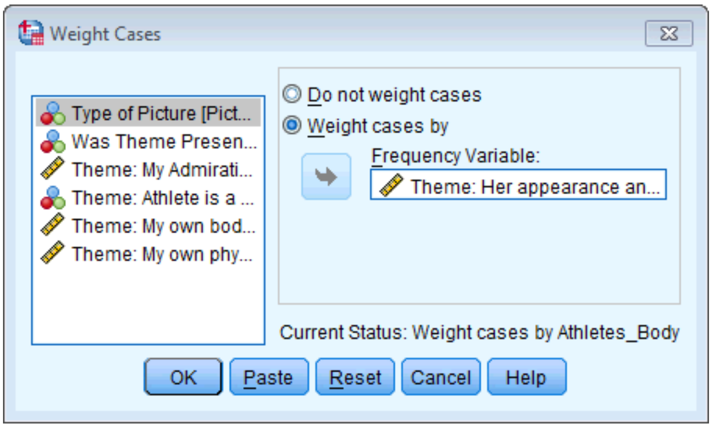

Figure 3

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 type of picture and was the theme present or absent in what the participant wrote. Select one of these variables and drag it into the box labelled *Row(s)* (or click on ➡). For this example, I selected **Picture** to be the rows of the table. Next, select the other variable of interest (**was the theme present or absent**) and drag it to the box labelled *Column(s)* (or click on ➡). Use the book chapter to select other appropriate options.

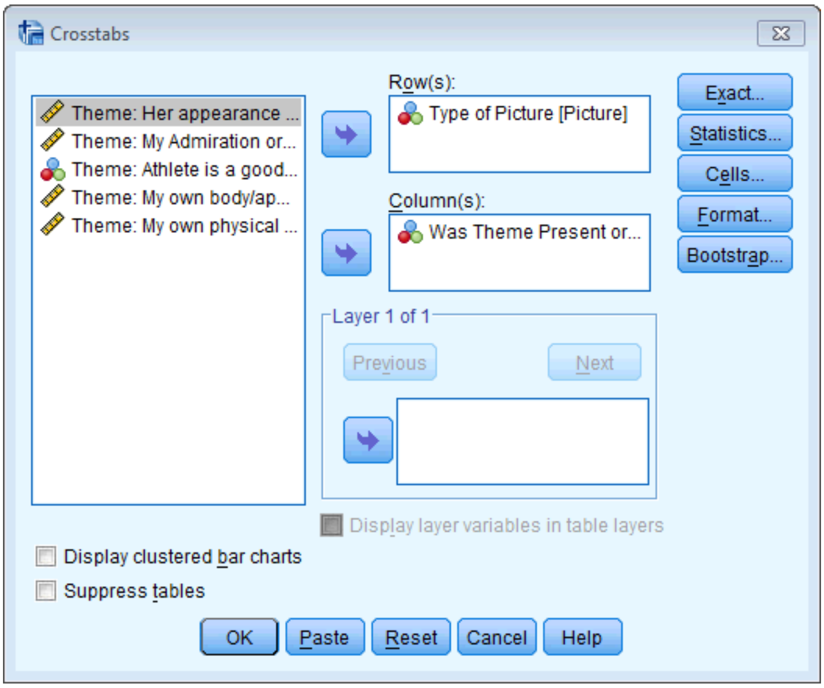

Figure 4

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	88a	29b	117
		Expected Count	47.8	69.2	117.0
		% within Type of Picture	75.2%	24.8%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	83.8%	19.1%	45.5%
		Std. Residual	5.8	-4.8	
	Sexualized Athletes	Count	17a	123b	140
		Expected Count	57.2	82.8	140.0
		% within Type of Picture	12.1%	87.9%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	16.2%	80.9%	54.5%
		Std. Residual	-5.3	4.4	
Total		Count	105	152	257
		Expected Count	105.0	152.0	257.0
		% within Type of Picture	40.9%	59.1%	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 = Her Body

Output 3

Chi-Square Tests[a]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	104.923[b]	1	.000		
Continuity Correction[c]	102.329	1	.000		
Likelihood Ratio	113.066	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	104.515	1	.000		
N of Valid Cases	257				

a. Type of Theme = Her Body

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 47.80.

c. Computed only for a 2x2 table

Output 4

The chi-square test is highly significant, $\chi^2(1) = 104.92$, $p < .001$. This indicates that the profile of theme present vs. theme absent differed across different pictures. Looking at the standardized residuals, they are significant for both pictures of performance athletes and sexualized pictures of athletes. If we look at the direction of these residuals (i.e., whether they are positive or negative), we can see that for pictures of performance athletes, the residual for 'theme absent' was positive but for 'theme present' was negative; this indicates that in this condition, more people than we would expect did not include the theme *her appearance and attractiveness* and fewer people than we would expect did include this theme in what they wrote. In the sexualized picture conditio,n on the other hand, the opposite was true: the residual for 'theme absent' was negative and for 'theme present' was positive. This indicates that in the sexualized picture condition, more people than we would expect included the theme *her appearance and attractiveness* in what they wrote and fewer people than we would expect did not include this theme in what they wrote.

Daniels says:

Her appearance and attractiveness. Hypothesis 1 predicted that participants in the sexualized athlete and sexualized model conditions would objectify the women in the photographs more than participants in the performance athlete condition. As expected, participants who saw the sexualized athletes made more *her appearance and attractiveness* statements than did participants who saw the performance athletes. Specifically, participants who saw the sexualized athletes (87.9%) made more statements about *her body* than did participants who saw the performance athletes (24.8%), χ^2 (1, n = 257) = 104.92, *p* < .001, Cramer's V = .64. The planned

Task 3

Using the same data, carry out a chi-square test to see whether the type of picture viewed was associated with indicating admiration or jealousy for the athlete.

We run this analysis in exactly the same way as in the previous question, except that we now have to weight the cases by the variable **Theme: My admiration or jealousy for the athlete**. Select Data Weight Cases...; 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 select the variable in which the number of cases is specified (in this case **Theme: My admiration or jealousy for the athlete**) 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 **Theme: My admiration or jealousy for the athlete.**

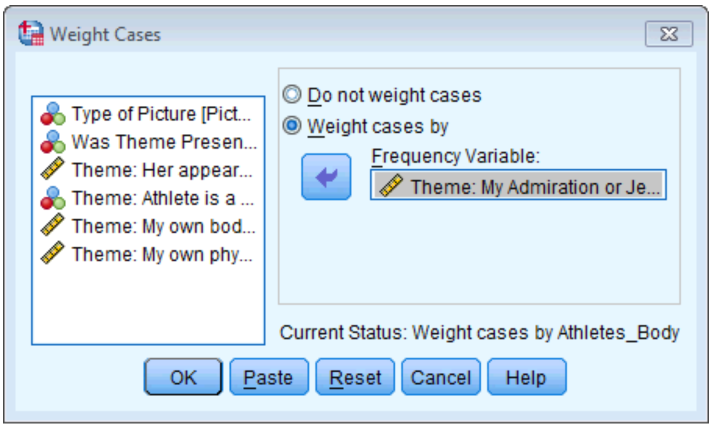

Figure 5

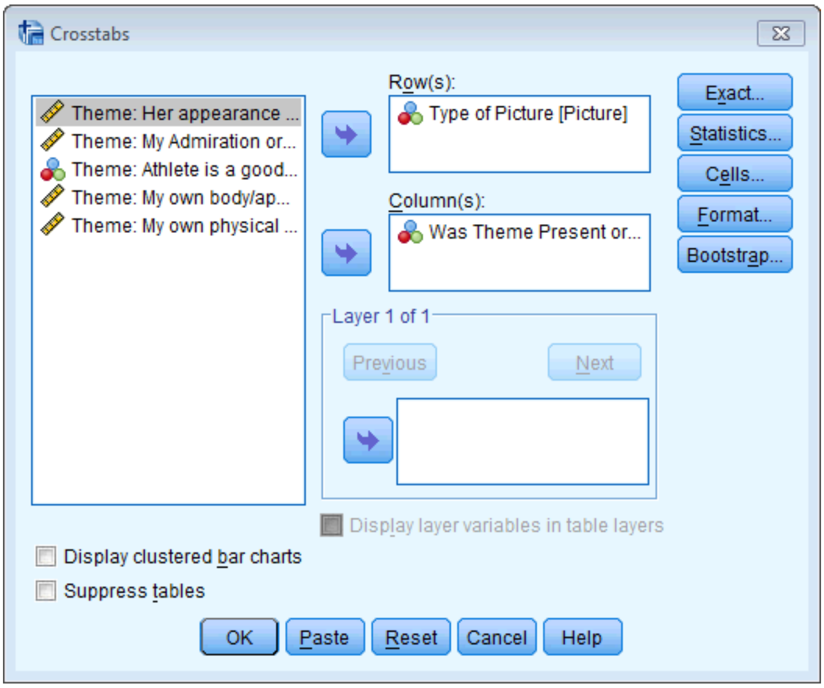

Figure 6

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	105a	12b	117
		Expected Count	86.0	31.0	117.0
		% within Type of Picture	89.7%	10.3%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	55.6%	17.6%	45.5%
		Std. Residual	2.0	-3.4	
	Sexualized Athletes	Count	84a	56b	140
		Expected Count	103.0	37.0	140.0
		% within Type of Picture	60.0%	40.0%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	44.4%	82.4%	54.5%
		Std. Residual	-1.9	3.1	
Total		Count	189	68	257
		Expected Count	189.0	68.0	257.0
		% within Type of Picture	73.5%	26.5%	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 = Admiration

Output 5

Chi-Square Tests[a]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	28.978[b]	1	.000		
Continuity Correction[c]	27.469	1	.000		
Likelihood Ratio	31.169	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	28.865	1	.000		
N of Valid Cases	257				

a. Type of Theme = Admiration

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 30.96.

c. Computed only for a 2x2 table

Output 6

The chi-square test is highly significant, $\chi^2(1) = 28.98$, $p < .001$. This indicates that the profile of theme present vs. theme absent differed across different pictures. Looking at the standardized residuals, they are significant for both pictures of performance athletes and sexualized pictures of athletes. If we look at the direction of these residuals (i.e., whether they are positive or negative), we can see that for pictures of performance athletes, the residual for 'theme absent' was positive but for 'theme present' was negative; this indicates that in this condition, more people than we would expect did not include the theme *My admiration or jealousy for the athlete* and fewer people than we would expect did include this theme in what they wrote. In the sexualized picture condition, on the other hand, the opposite was true: the residual for 'theme absent' was negative and for 'theme present was positive'. This indicates that in the sexualized picture condition, more people than we would expect included the theme *My admiration or jealousy for the athlete* in what they wrote and fewer people than we would expect did not include this theme in what they wrote.

Daniels says:

> *My feelings about her.* Participants who saw the sexualized athletes (40.0%) made more *admiration/jealousy* statements than did participants who saw the performance athletes (10.3%), χ^2 (1, n = 257) = 28.98, $p < .001$, Cramer's V = .34. An *admiration/jealousy* statement about a sexualized athlete was, "seeing her body, as nice as it is, makes me wish my body was as nice as hers" (19-year-old, European American). The opposite pattern was found for *role model/ inspiration* statements. Participants who saw the performance

Task 4

Using the same data, carry out a chi-square test to see whether the type of picture viewed was associated with indicating that the athlete was a role model or motivating.

We run this analysis in exactly the same way as in the previous question, except that we now have to weight the cases by the variable **Theme: My admiration or jealousy for the athlete**. Select Data Weight Cases...; 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 [arrow button] to move it back to the variable list and clear the box. Then select the variable in which the number of cases is specified (in this case **Theme: Athlete is a good role model**) and drag it to the box labelled *Frequency Variable* (or click on [arrow button]). This process tells the computer that it should weight each category combination by the number in the column labelled **Theme: Athlete is a good role model.**

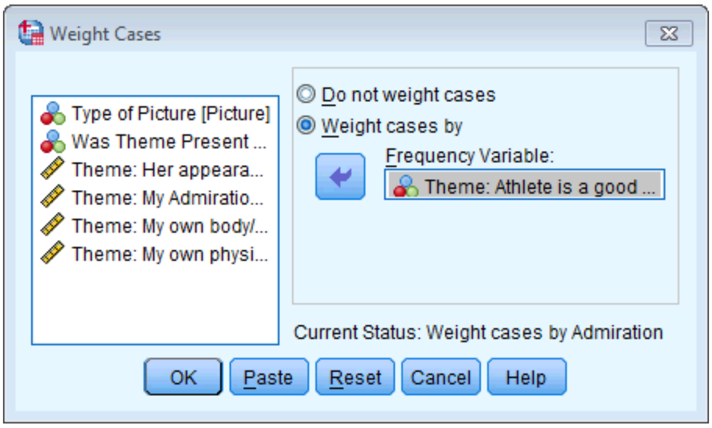

Figure 7

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	70a	47b	117
		Expected Count	92.4	24.6	117.0
		% within Type of Picture	59.8%	40.2%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	34.5%	87.0%	45.5%
		Std. Residual	-2.3	4.5	
	Sexualized Athletes	Count	133a	7b	140
		Expected Count	110.6	29.4	140.0
		% within Type of Picture	95.0%	5.0%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	65.5%	13.0%	54.5%
		Std. Residual	2.1	-4.1	
Total		Count	203	54	257
		Expected Count	203.0	54.0	257.0
		% within Type of Picture	79.0%	21.0%	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 = Role Model

Output 7

Chi-Square Tests[a]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	47.503[b]	1	.000		
Continuity Correction[c]	45.408	1	.000		
Likelihood Ratio	51.023	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	47.319	1	.000		
N of Valid Cases	257				

a. Type of Theme = Role Model

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 24.58.

c. Computed only for a 2x2 table

Output 8

The chi-square test is highly significant, $\chi^2(1) = 47.50$, $p < .001$. This indicates that the profile of theme present vs. theme absent differed across different pictures. Looking at the standardized residuals, they are significant for both types of pictures. If we look at the direction of these residuals (i.e., whether they are positive or negative), we can see that for pictures of performance athletes, the residual for 'theme absent' was negative but was positive for 'theme present'. This indicates that when looking at pictures of performance athletes, more people than we would expect included the theme *Athlete is a good role model* and fewer people than we would expect did not include this theme in what they wrote. In the sexualized picture condition on the other hand, the opposite was true: the residual for 'theme absent' was positive and for 'theme present' it was negative. This indicates that in the sexualized picture condition, more people than we would expect did not include the theme *Athlete is a good role model* in what they wrote and fewer people than we would expect did include this theme in what they wrote.

Daniels says:

> *inspiration* statements. Participants who saw the performance athletes (40.2%) made more *role model/inspiration* statements than did participants who saw the sexualized athletes (5.0%), χ^2 (1, n = 257) = 47.50, $p < .001$, Cramer's V = .43. A *role model/inspiration* statement about a performance athlete was, "in this photo, Mia Hamm runs her heart out for the love of one game. Although I'm not a soccer player, this gives me a sense of determination to achieve my goal even if it doesn't involve a soccer ball. This photo represents woman [sic] who are strong…" (15-year-old, European American).

Task 5

Using the same data, carry out a chi-square test to see whether the type of picture viewed was associated with the participant commenting on their own physical activity.

We run this analysis in exactly the same way as in the previous question, except that we now have to weight the cases by the variable **Theme: My own physical activity**. Select Data

Weight Cases...; 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 [arrow button] to move it back to the variable list and clear the box. Then select the variable in which the number of cases is specified (in this case **Theme: My own physical activity**) and drag it to the box labelled *Frequency Variable* (or click on [arrow button]). This process tells the computer that it should weight each category combination by the number in the column labelled **Theme: My own physical activity.**

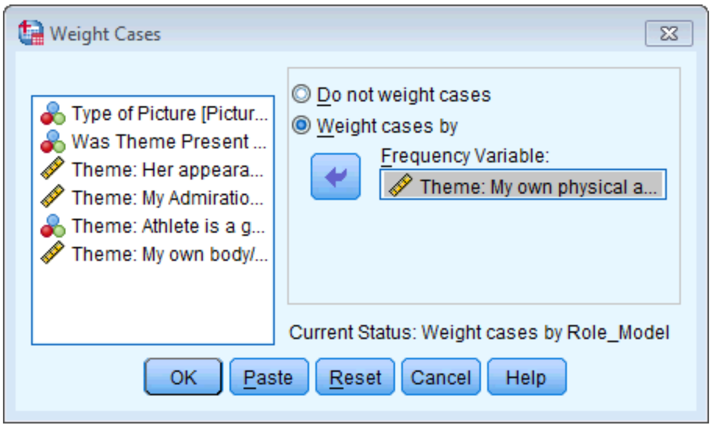

Figure 8

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	84a	33b	117
		Expected Count	92.0	25.0	117.0
		% within Type of Picture	71.8%	28.2%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	41.6%	60.0%	45.5%
		Std. Residual	-.8	1.6	
	Sexualized Athletes	Count	118a	22b	140
		Expected Count	110.0	30.0	140.0
		% within Type of Picture	84.3%	15.7%	100.0%
		% within Was Theme Present or Absent in what participant wrote?	58.4%	40.0%	54.5%
		Std. Residual	.8	-1.5	
Total		Count	202	55	257
		Expected Count	202.0	55.0	257.0
		% within Type of Picture	78.6%	21.4%	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 = My physical activity

Output 9

Chi-Square Tests[a]

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.912[b]	1	.015		
Continuity Correction[c]	5.192	1	.023		
Likelihood Ratio	5.904	1	.015		
Fisher's Exact Test				.021	.011
Linear-by-Linear Association	5.889	1	.015		
N of Valid Cases	257				

a. Type of Theme = My physical activity

b. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 25.04.

c. Computed only for a 2x2 table

Output 2

The chi-square test is significant, $\chi^2(1) = 5.91$, $p = .02$. This indicates that the profile of theme present vs. theme absent differed across different pictures. Looking at the standardized residuals, they are not significant for either type of picture (i.e., they are less than 1.96). If we look at the direction of these residuals (i.e., whether they are positive or negative), we can see that for pictures of performance athletes, the residual for 'theme absent' was negative and for 'theme present' was positive. This indicates that when looking at pictures of performance athletes, more people than we would expect included the theme *My own physical activity* and fewer people than we would expect did not include this theme in what they wrote. In the sexualized picture condition on the other hand, the opposite was true: the residual for 'theme absent' was positive and for 'theme present' it was negative. This indicates that in the sexualized picture condition, more people than we would expect did not include the theme *My own physical activity* in what they wrote and fewer people than we would expect did include this theme in what they wrote.

Daniels says:

> Hypothesis 4 predicted that participants in the performance athlete condition would make statements about their own physical skills more than participants in the sexualized athlete and sexualized model conditions. As expected, participants who saw the performance athletes (28.2%) made more *my physical activity* statements than did participants who saw the sexualized athletes (15.7%), χ^2 (1, n = 257) = 5.91, p = .02, Cramer's V = .15. A *my physical activity* statement after viewing a performance athlete was, "this photograph makes me feel like geting [sic] up and playing some type of sport. It's a very active photo, very aggressive and powerful" (14-year-old, multiple ethnicities). The planned chi-square analysis to investi-

Task 6

I wrote much of the third edition of this book in the Netherlands (I have a soft spot for Holland). I noticed cultural differences to England. The Dutch travel by bike much more than the English. I noticed also that many more Dutch people cycle while steering with only one hand. I pointed this out to one of my friends, Birgit Mayer, and she said that I was being a crazy English fool and that Dutch people did not cycle one-handed. Several weeks of me pointing at one-handed cyclists and her pointing at two-handed cyclists ensued. To put it to the test I counted the number of Dutch and English cyclists who *ride with one or two hands on the handlebars (**Handlebars.sav**). Can you work out which one of us is right?*

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 **Frequency**) 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 **Frequency**.

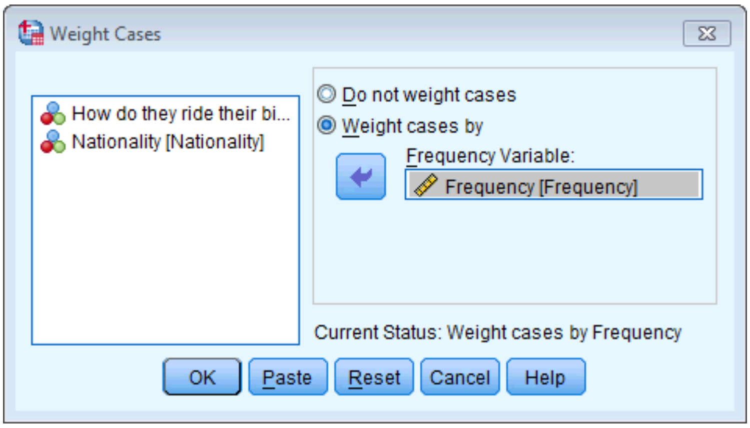

Figure 9

To run the chi-square tests, select Analyze Descriptive Statistics Crosstabs.... First, select one of the variables of interest in the variable list and drag it into the box labelled *Row(s)* (or click on). For this example, I selected **Nationality** to be the rows of the table. Next, select the other variable of interest (**Hands**) and drag it to the box labelled *Column(s)* (or click on). Select the same options as in the book.

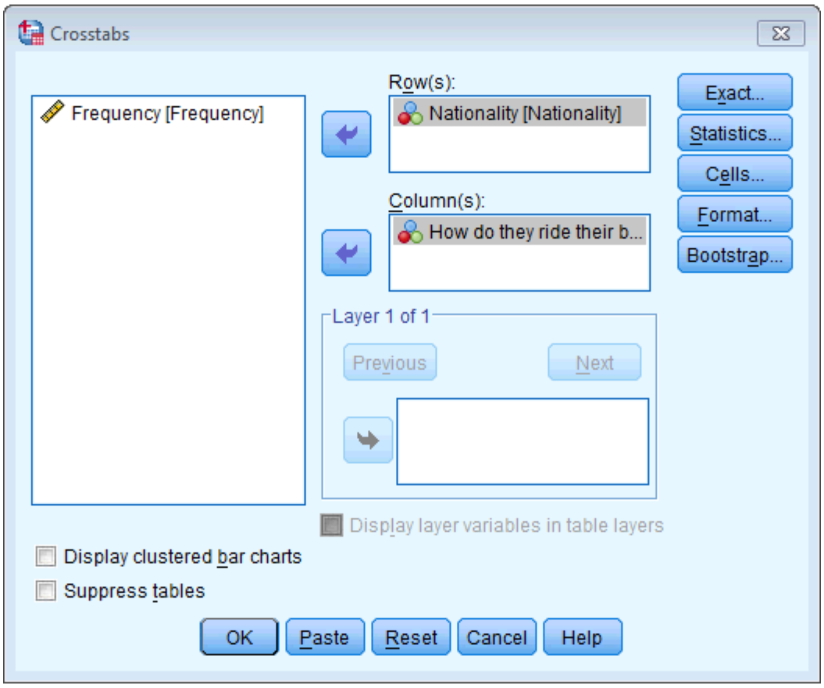

Figure 10

The crosstabulation table produced by SPSS contains the number of cases that fall into each combination of categories. We can see that in total 137 people rode their bike one-handed, of which 120 (87.6%) were Dutch and only 17 (12.4%) were English; 732 people rode their bike two-handed, of which 578 (79%) were Dutch and only 154 (21%) were English.

Nationality * How do they ride their bike? Crosstabulation

			How do they ride their bike?		
			One Handed	Two Handed	Total
Nationality	Dutch	Count	120	578	698
		Expected Count	110.0	588.0	698.0
		% within Nationality	17.2%	82.8%	100.0%
		% within How do they ride their bike?	87.6%	79.0%	80.3%
		Std. Residual	.9	-.4	
	English	Count	17	154	171
		Expected Count	27.0	144.0	171.0
		% within Nationality	9.9%	90.1%	100.0%
		% within How do they ride their bike?	12.4%	21.0%	19.7%
		Std. Residual	-1.9	.8	
	Total	Count	137	732	869
		Expected Count	137.0	732.0	869.0
		% within Nationality	15.8%	84.2%	100.0%
		% within How do they ride their bike?	100.0%	100.0%	100.0%

Output 11

Before moving on to look at the test statistic itself, it is vital that we check that the assumption for chi-square has been met. The assumption is that in 2 × 2 tables (which is what we have here), all expected frequencies should be greater than 5. If you look at the expected counts in the crosstabulation table, it should be clear that the smallest expected count is 27 (for English people who ride their bike one-handed). This value exceeds 5 and so the assumption has been met.

The value of the chi-square statistic is 5.44. This value has a two-tailed significance of .020, which is smaller than .05 (hence significant). This suggests that the pattern of bike riding (i.e., relative numbers of one- and two-handed riders) significantly differs in English and Dutch people.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	5.437[a]	1	.020	.026	.011	
Continuity Correction[b]	4.905	1	.027			
Likelihood Ratio	5.958	1	.015	.019	.011	
Fisher's Exact Test				.019	.011	
Linear-by-Linear Association	5.431[c]	1	.020	.026	.011	.006
N of Valid Cases	869					

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 26.96.

b. Computed only for a 2x2 table

c. The standardized statistic is 2.330.

Output 12

The significant result indicates that there is an association between whether someone is Dutch or English and whether they ride their bike one- or two-handed. Looking at the frequencies, this finding seems to show that the ratio of one- to two-handed riders differs in Dutch and English people. In Dutch people 17.2% ride their bike one-handed compared to 82.8% who ride two-handed. In England, though, only 9.9% ride their bike one-handed (almost half as many as in Holland), and 90.1% ride two-handed. If we look at the standardized residuals (in the contingency table) we can see that the only cell with a residual approaching significance (a value that lies outside of ±1.96) is the cell for English people riding one-handed (z = –1.9). The fact that this value is negative tells us that *fewer* people than expected fell into this cell.

Task 7

Compute and interpret the odds ratio for Task 6.

The odds of someone riding one-handed if they are Dutch are 120/578 = 0.21, and the odds of someone riding one-handed if they are English are 17/154 = 0.11. Therefore, the odds ratio is 0.21/0.11 = 1.9. In other words, the odds of riding one-handed if you are Dutch are 1.9 times higher than if you are English (or the odds of riding one-handed if you are English are about half that of a Dutch person). We could report as follows:

- ✓ There was a significant association between nationality and whether the Dutch or English rode their bike one- or two-handed, $\chi^2(1) = 5.44$, $p < .05$. This represents the fact that, based on the odds ratio, the odds of riding a bike one-handed were 1.9 time higher for Dutch people than for English people. This supports Field's argument that there are more one-handed bike riders in the Netherlands than in England and utterly refutes Mayer's competing theory. These data are in no way made up.

Task 8

Certain editors at Sage like to think they're a bit of a whiz at football (soccer if you prefer). To see whether they are better than Sussex lecturers and postgraduates we invited various employees of Sage to join in our football matches. Every player was allowed only to play in one match. Over many matches, we counted the number of players that scored goals. The data are in the file ***Sage Editors Can't Play Football.sav****. Do a chi-square test to see whether more publishers or academics scored goals. We predict that Sussex people will score more than Sage people.*

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 **Frequency**) 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 **Frequency**.

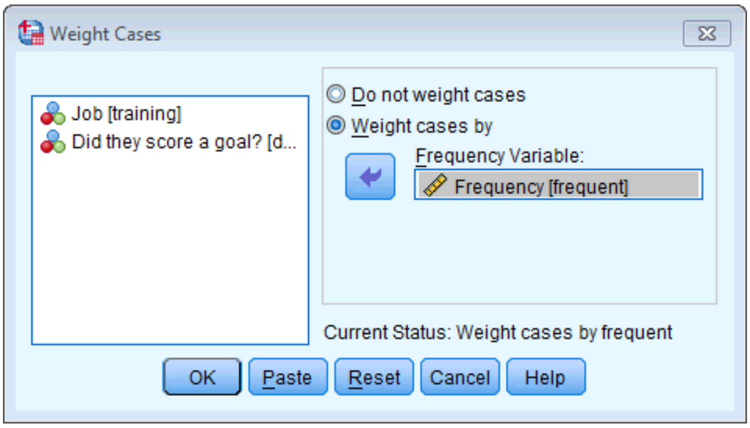

Figure 11

To run the chi-square tests, select Analyze Descriptive Statistics Crosstabs... . First, select one of the variables of interest in the variable list and drag it into the box labelled *Row(s)* (or click on). For this example, I selected **Job** to be the rows of the table. Next, select the other variable of interest (**Score**) and drag it to the box labelled *Column(s)* (or click on). Select the same options as in the book.

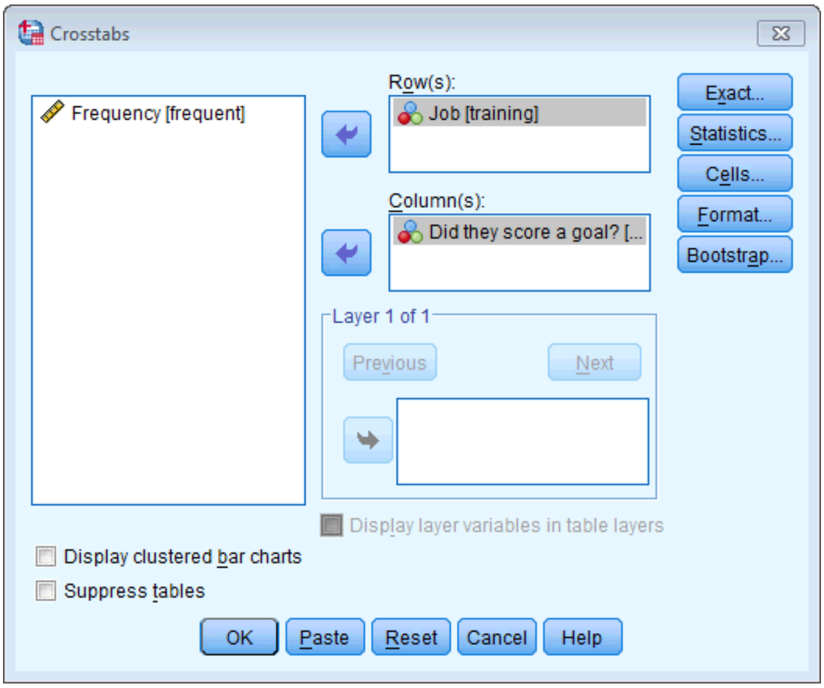

Figure 12

The crosstabulation table produced by SPSS contains the number of cases that fall into each combination of categories. We can see that in total 28 people scored goals (36.4% of the total) and of these 5 were from Sage Publications (17.9% of the total that scored) and 23 were from Sussex (82.1% of the total that scored); 49 people didn't score at all (63.6% of the total) and, of those, 19 worked for Sage (38.8% of the total that didn't score) and 30 were from Sussex (61.2% of the total that didn't score).

Job * Did they score a goal? Crosstabulation

			Did they score a goal?		
			Yes	No	Total
Job	Sage Publications	Count	5	19	24
		Expected Count	8.7	15.3	24.0
		% within Job	20.8%	79.2%	100.0%
		% within Did they score a goal?	17.9%	38.8%	31.2%
		% of Total	6.5%	24.7%	31.2%
	University of Sussex	Count	23	30	53
		Expected Count	19.3	33.7	53.0
		% within Job	43.4%	56.6%	100.0%
		% within Did they score a goal?	82.1%	61.2%	68.8%
		% of Total	29.9%	39.0%	68.8%
Total		Count	28	49	77
		Expected Count	28.0	49.0	77.0
		% within Job	36.4%	63.6%	100.0%
		% within Did they score a goal?	100.0%	100.0%	100.0%
		% of Total	36.4%	63.6%	100.0%

Output 13

Before moving on to look at the test statistic itself it is vital that we check that the assumption for chi-square has been met. The assumption is that in 2 × 2 tables (which is what we have here), all expected frequencies should be greater than 5. If you look at the expected counts in the crosstabulation table, it should be clear that the smallest expected count is 8.7 (for Sage editors who scored). This value exceeds 5 and so the assumption has been met.

Pearson's chi-square test examines whether there is an association between two categorical variables (in this case the job and whether the person scored or not). As part of the *Crosstabs* procedure SPSS produces a table that includes the chi-square statistic and its significance value. The Pearson chi-square statistic tests whether the two variables are independent. If the significance value is small enough (conventionally *Sig.* must be less than .05) then we reject the hypothesis that the variables are independent and accept the hypothesis that they are in some way related. The value of the chi-square statistic is given in the table (and the degrees of freedom) as is the significance value. The value of the chi-square statistic is 3.63. This value has a two-tailed significance of .057, which is bigger than .05 (hence non-significant). However, we made a specific prediction (that Sussex people would score more than Sage people), hence we can halve this value. Therefore, the chi-square is significant (one-tailed) because $p = .0285$, which is less than .05. The one-tailed significance values of the other statistics are also less than .05, so we have consistent results.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	3.634[b]	1	.057		
Continuity Correction[a]	2.725	1	.099		
Likelihood Ratio	3.834	1	.050		
Fisher's Exact Test				.075	.047
Linear-by-Linear Association	3.587	1	.058		
N of Valid Cases	77				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 8.73.

Output 14

The highly significant result indicates that there is an association between the type of job someone does and whether they score goals. This significant finding reflects the fact that for Sussex employees there is about a 50% split of those that scored and those that didn't, but for Sage employees there is about a 20–80 split with only 20% scoring and 80% not scoring. This supports our hypothesis that people from Sage, despite their delusions, are crap at football!

Task 9

Compute and interpret the odds ratio for Task 8.

The odds of someone scoring given that they were employed by Sage are 5/19 = 0.26, and the odds of someone scoring given that they were employed by Sussex University are 23/30 = 0.77. Therefore, the odds ratio is 0.26/0.77 = 0.34. In other words, the odds of scoring if you work for Sage are 0.34 times as high as if you work for Sussex; another way to express this is that if you work for Sage, the odds of scoring are 1/0.34 = 2.95 times lower than if you work for Sussex! We could report this as follows:

- ✓ There was a significant association between the type of job and whether or not a person scored a goal, $\chi^2(1) = 3.63$, $p < .05$ (one-tailed). This represents the fact that, based on the odds ratio, Sage employees were 2.95 times less likely to score than Sussex employees.

Task 10

I was interested in whether horoscopes are just tosh. Therefore, I took 2201 people, made a note of their star sign (this variable, obviously, has 12 categories: Capricorn, Aquarius, Pisces, Aries, Taurus, Gemini, Cancer, Leo, Virgo, Libra, Scorpio and Sagittarius) and whether they believed in horoscopes (this variable has two categories: believer or unbeliever). I then sent them a horoscope in the post of what would happen over the next month. Everybody, regardless of their star sign, received the same horoscope, which read: 'August is an exciting month for you. You will make friends with a tramp in the first week of the month and cook him a cheese omelette. Curiosity is your greatest virtue, and in the second week, you'll discover knowledge of a subject that you previously thought was boring, statistics perhaps. You might purchase a book around this time that guides you towards this knowledge. Your new wisdom leads to a change in career around the third week, when you ditch your current job and become an accountant. By the final week you find yourself free from the constraints of having friends, your boy/girlfriend has left you for a Russian ballet dancer with a glass eye, and you now spend your weekends doing loglinear analysis by hand with a pigeon called Hephzibah for company.' At the end of August I interviewed all of these people and I classified the horoscope as having come true, or not, based on how closely their lives had matched the fictitious horoscope. The data are in the file ***Horoscope.sav****. Conduct a loglinear analysis to see whether there is a relationship between the person's star sign, whether they believe in horoscopes and whether the horoscope came true.*

Running the analysis

Data are entered for this example as frequency values for each combination of categories, so before you begin you must weight the cases by the variable **Frequency**. If you don't do this the entire output will be wrong! 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 **Frequency**) 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 **Frequency**.

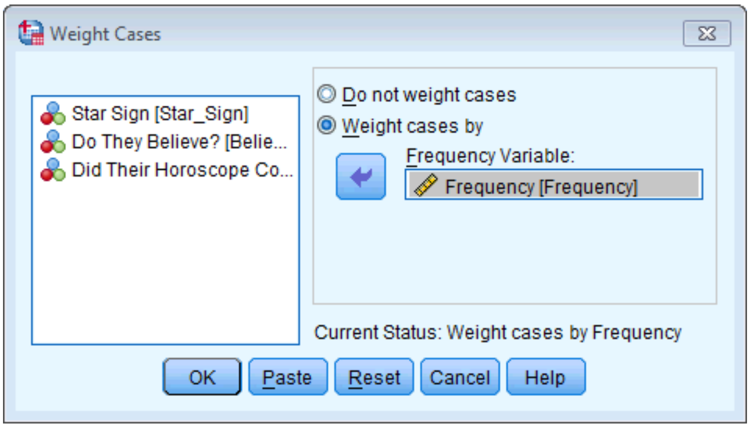

Figure 13

To get a crosstabulation table, select Analyze Descriptive Statistics ▸ Crosstabs... . We have three variables in our crosstabulation table: whether someone believes in star signs or not (**Believe**), the star sign of the person (**Star_Sign**) and whether the horoscope came true or not (**True**). Select **Believe** and drag it into the box labelled *Row(s)* (or click on). Next, select **True** and drag it to the box labelled *Column(s)* (or click on). We need to define our third variable as a layer. Select **Star_Sign** and drag it to the box labelled *Layer 1 of 1* (or click on). Then click on Cells... and select the options required.

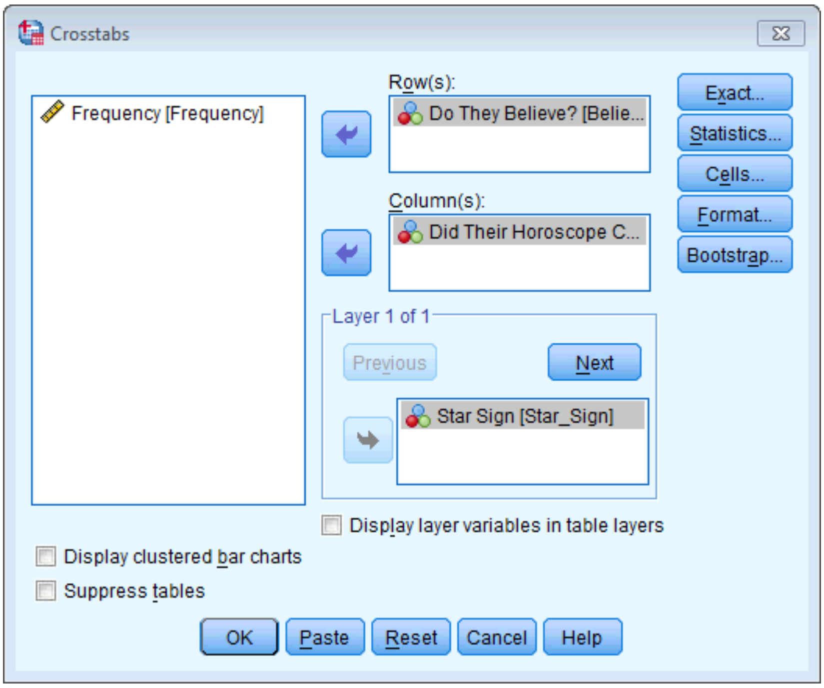

Figure 14

The crosstabulation table produced by SPSS contains the number of cases that fall into each combination of categories. Although this table is quite complicated, you should be able to see that there are roughly the same number of believers and non-believers and similar numbers of those whose horoscopes came true or didn't. These proportions are fairly consistent also across the different star signs! Also there are no expected counts less than 5, so our assumptions are met.

Do They Believe? * Did Their Horoscope Come True? * Star Sign Crosstabulation

Star Sign				Did Their Horoscope Come True? Horoscope Didn't Come True	Horoscope Came True	Total
Capricorn	Do They Believe?	Unbeliever	Count	56	46	102
			Expected Count	51.0	51.0	102.0
			Std. Residual	.7	-.7	
		Believer	Count	50	60	110
			Expected Count	55.0	55.0	110.0
			Std. Residual	-.7	.7	
		Total	Count	106	106	212
			Expected Count	106.0	106.0	212.0
Aquarius	Do They Believe?	Unbeliever	Count	26	20	46
			Expected Count	22.8	23.2	46.0
			Std. Residual	.7	-.7	
		Believer	Count	22	29	51
			Expected Count	25.2	25.8	51.0
			Std. Residual	-.6	.6	
		Total	Count	48	49	97
			Expected Count	48.0	49.0	97.0
Pisces	Do They Believe?	Unbeliever	Count	55	51	106
			Expected Count	52.6	53.4	106.0
			Std. Residual	.3	-.3	
		Believer	Count	64	70	134
			Expected Count	66.4	67.6	134.0
			Std. Residual	-.3	.3	
		Total	Count	119	121	240
			Expected Count	119.0	121.0	240.0
Aries	Do They Believe?	Unbeliever	Count	42	36	78
			Expected Count	43.2	34.8	78.0
			Std. Residual	-.2	.2	
		Believer	Count	70	54	124
			Expected Count	68.8	55.2	124.0
			Std. Residual	.2	-.2	
		Total	Count	112	90	202
			Expected Count	112.0	90.0	202.0
Taurus	Do They Believe?	Unbeliever	Count	56	42	98
			Expected Count	50.3	47.7	98.0
			Std. Residual	.8	-.8	
		Believer	Count	41	50	91
			Expected Count	46.7	44.3	91.0
			Std. Residual	-.8	.9	
		Total	Count	97	92	189
			Expected Count	97.0	92.0	189.0
Gemini	Do They Believe?	Unbeliever	Count	65	53	118
			Expected Count	60.1	57.9	118.0
			Std. Residual	.6	-.6	
		Believer	Count	40	48	88
			Expected Count	44.9	43.1	88.0
			Std. Residual	-.7	.7	
		Total	Count	105	101	206
			Expected Count	105.0	101.0	206.0
Cancer	Do They Believe?	Unbeliever	Count	84	76	160
			Expected Count	85.0	75.0	160.0
			Std. Residual	-.1	.1	
		Believer	Count	96	83	179
			Expected Count	95.0	84.0	179.0
			Std. Residual	.1	-.1	
		Total	Count	180	159	339
			Expected Count	180.0	159.0	339.0
Leo	Do They Believe?	Unbeliever	Count	14	23	37
			Expected Count	13.9	23.1	37.0
			Std. Residual	.0	.0	
		Believer	Count	12	20	32
			Expected Count	12.1	19.9	32.0
			Std. Residual	.0	.0	
		Total	Count	26	43	69
			Expected Count	26.0	43.0	69.0
Virgo	Do They Believe?	Unbeliever	Count	69	55	124
			Expected Count	61.2	62.8	124.0
			Std. Residual	1.0	-1.0	
		Believer	Count	49	66	115
			Expected Count	56.8	58.2	115.0
			Std. Residual	-1.0	1.0	
		Total	Count	118	121	239
			Expected Count	118.0	121.0	239.0
Libra	Do They Believe?	Unbeliever	Count	27	26	53
			Expected Count	23.4	29.6	53.0
			Std. Residual	.7	-.7	
		Believer	Count	22	36	58
			Expected Count	25.6	32.4	58.0
			Std. Residual	-.7	.6	
		Total	Count	49	62	111
			Expected Count	49.0	62.0	111.0
Scorpio	Do They Believe?	Unbeliever	Count	32	20	52
			Expected Count	27.0	25.0	52.0
			Std. Residual	1.0	-1.0	
		Believer	Count	24	32	56
			Expected Count	29.0	27.0	56.0
			Std. Residual	-.9	1.0	
		Total	Count	56	52	108
			Expected Count	56.0	52.0	108.0
Sagittarius	Do They Believe?	Unbeliever	Count	56	41	97
			Expected Count	50.3	46.7	97.0
			Std. Residual	.8	-.8	
		Believer	Count	42	50	92
			Expected Count	47.7	44.3	92.0
			Std. Residual	-.8	.9	
		Total	Count	98	91	189
			Expected Count	98.0	91.0	189.0

Output 15

The loglinear analysis

Then run the main analysis. The way to run loglinear analysis that is consistent with my section on the theory of the analysis is to select Analyze Loglinear ▸ Model Selection... to access the dialog box. Select any variables that you want to include in the analysis by clicking on them with the mouse (remember that you can select several at the same time by holding down the *Ctrl* key) and then dragging them to the box labelled *Factor(s)* (or click on ➔). When there is a variable in this box the Define Range... button becomes active. We have to tell SPSS the codes that we've used to define our categorical variables. Select a variable in the *Factor(s)* box and then click on Define Range... to activate a dialog box that allows you to specify the value of the minimum and maximum code that you've used for that variable. When you've done this, click on Continue to return to main dialog box.

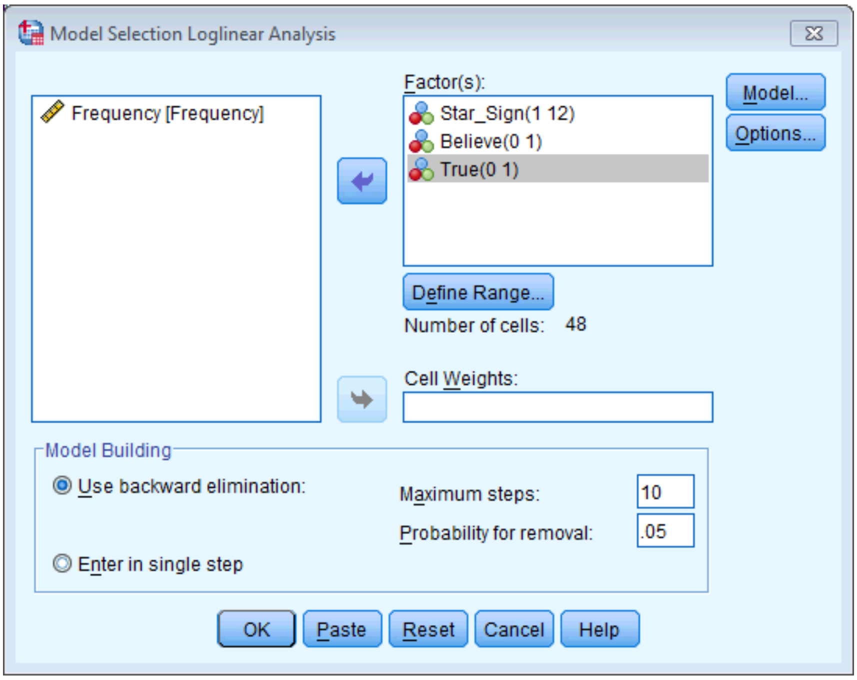

Figure 15

Output from loglinear analysis

The initial output from the loglinear analysis tells us that we have 2201 cases. SPSS then lists all of the factors in the model and the number of levels they have. To begin with, SPSS fits the saturated model (all terms are in the model, including the highest-order interaction, in this case the star sign × believer × true interaction). SPSS then gives us the observed and expected counts for each of the combinations of categories in our model. These values should be the same as the original contingency table, except that each cell has 0.5 added to it. The final bit of this initial output gives us two goodness-of-fit statistics (Pearson's chi-square and the likelihood-ratio statistic, both of which we came across at the beginning of this chapter). In this context these tests are testing the hypothesis that the frequencies predicted by the model (the expected frequencies) are significantly different from the actual frequencies in our data (the observed frequencies). At this stage the model fits the data perfectly, so both statistics are 0 and yield a *p*-value of '.' (i.e., SPSS can't compute the probability).

Data Information

		N
Cases	Valid	48
	Out of Range[a]	0
	Missing	0
	Weighted Valid	2201
Categories	Star Sign	12
	Do They Believe?	2
	Did Their Horoscope Come True?	2

a. Cases rejected because of out of range factor values.

Output 16

Cell Counts and Residuals

Star Sign	Do They Believe?	Did Their Horoscope Come True?	Observed Count[a]	Observed %	Expected Count	Expected %	Residuals	Std. Residuals
Capricorn	Unbeliever	Horoscope Didn't Come True	56.500	2.6%	56.500	2.6%	.000	.000
		Horoscope Came True	46.500	2.1%	46.500	2.1%	.000	.000
	Believer	Horoscope Didn't Come True	50.500	2.3%	50.500	2.3%	.000	.000
		Horoscope Came True	60.500	2.7%	60.500	2.7%	.000	.000
Aquarius	Unbeliever	Horoscope Didn't Come True	26.500	1.2%	26.500	1.2%	.000	.000
		Horoscope Came True	20.500	.9%	20.500	.9%	.000	.000
	Believer	Horoscope Didn't Come True	22.500	1.0%	22.500	1.0%	.000	.000
		Horoscope Came True	29.500	1.3%	29.500	1.3%	.000	.000
Pisces	Unbeliever	Horoscope Didn't Come True	55.500	2.5%	55.500	2.5%	.000	.000
		Horoscope Came True	51.500	2.3%	51.500	2.3%	.000	.000
	Believer	Horoscope Didn't Come True	64.500	2.9%	64.500	2.9%	.000	.000
		Horoscope Came True	70.500	3.2%	70.500	3.2%	.000	.000
Aries	Unbeliever	Horoscope Didn't Come True	42.500	1.9%	42.500	1.9%	.000	.000
		Horoscope Came True	36.500	1.7%	36.500	1.7%	.000	.000
	Believer	Horoscope Didn't Come True	70.500	3.2%	70.500	3.2%	.000	.000
		Horoscope Came True	54.500	2.5%	54.500	2.5%	.000	.000
Taurus	Unbeliever	Horoscope Didn't Come True	56.500	2.6%	56.500	2.6%	.000	.000
		Horoscope Came True	42.500	1.9%	42.500	1.9%	.000	.000
	Believer	Horoscope Didn't Come True	41.500	1.9%	41.500	1.9%	.000	.000
		Horoscope Came True	50.500	2.3%	50.500	2.3%	.000	.000
Gemini	Unbeliever	Horoscope Didn't Come True	65.500	3.0%	65.500	3.0%	.000	.000
		Horoscope Came True	53.500	2.4%	53.500	2.4%	.000	.000
	Believer	Horoscope Didn't Come True	40.500	1.8%	40.500	1.8%	.000	.000
		Horoscope Came True	48.500	2.2%	48.500	2.2%	.000	.000
Cancer	Unbeliever	Horoscope Didn't Come True	84.500	3.8%	84.500	3.8%	.000	.000
		Horoscope Came True	76.500	3.5%	76.500	3.5%	.000	.000
	Believer	Horoscope Didn't Come True	96.500	4.4%	96.500	4.4%	.000	.000
		Horoscope Came True	83.500	3.8%	83.500	3.8%	.000	.000
Leo	Unbeliever	Horoscope Didn't Come True	14.500	.7%	14.500	.7%	.000	.000
		Horoscope Came True	23.500	1.1%	23.500	1.1%	.000	.000
	Believer	Horoscope Didn't Come True	12.500	.6%	12.500	.6%	.000	.000
		Horoscope Came True	20.500	.9%	20.500	.9%	.000	.000
Virgo	Unbeliever	Horoscope Didn't Come True	69.500	3.2%	69.500	3.2%	.000	.000
		Horoscope Came True	55.500	2.5%	55.500	2.5%	.000	.000
	Believer	Horoscope Didn't Come True	49.500	2.2%	49.500	2.2%	.000	.000
		Horoscope Came True	66.500	3.0%	66.500	3.0%	.000	.000
Libra	Unbeliever	Horoscope Didn't Come True	27.500	1.2%	27.500	1.2%	.000	.000
		Horoscope Came True	26.500	1.2%	26.500	1.2%	.000	.000
	Believer	Horoscope Didn't Come True	22.500	1.0%	22.500	1.0%	.000	.000
		Horoscope Came True	36.500	1.7%	36.500	1.7%	.000	.000
Scorpio	Unbeliever	Horoscope Didn't Come True	32.500	1.5%	32.500	1.5%	.000	.000
		Horoscope Came True	20.500	.9%	20.500	.9%	.000	.000
	Believer	Horoscope Didn't Come True	24.500	1.1%	24.500	1.1%	.000	.000
		Horoscope Came True	32.500	1.5%	32.500	1.5%	.000	.000
Sagittarius	Unbeliever	Horoscope Didn't Come True	56.500	2.6%	56.500	2.6%	.000	.000
		Horoscope Came True	41.500	1.9%	41.500	1.9%	.000	.000
	Believer	Horoscope Didn't Come True	42.500	1.9%	42.500	1.9%	.000	.000
		Horoscope Came True	50.500	2.3%	50.500	2.3%	.000	.000

a. For saturated models, .500 has been added to all observed cells.

Output 17

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	.000	0	.
Pearson	.000	0	.

Output 18

The next part of the output tells us something about which components of the model can be removed. The first bit of the output is labelled *K-Way and Higher-Order Effects*, and underneath there is a table showing likelihood-ratio and chi-square statistics when K = 1, 2 and 3 (as we go down the rows of the table).

The first row (K = 1) tells us whether removing the one-way effects (i.e., the main effects of star sign, believer and true) and any higher-order effects will significantly affect the fit of the model. There are lots of higher-order effects here – there are the two-way interactions and the three-way interaction – and so this is basically testing whether if we remove everything from the model there will be a significant effect on the fit of the model. This is highly significant because the p-value is given as .000, which is less than .05. The next row of the table (K = 2) tells us whether removing the two-way interactions (i.e., the star sign × believer, star sign × true, and believer × true interactions) and any higher-order effects will affect the model. In this case there is a higher-order effect (the three-way interaction) so this is testing whether removing the two-way interactions and the three-way interaction would affect the fit of the model. This is significant (p = .03, which is less than .05) indicating that if we removed the two-way interactions and the three-way interaction then this would have a significant detrimental effect on the model. The final row (K = 3) is testing whether removing the three-way effect *and* higher-order effects will significantly affect the fit of the model. The three-way interaction is of course the highest-order effect that we have. so this is simply testing whether removal of the three-way interaction (star sign × believer × true) will significantly affect the fit of the model. If you look at the two columns labelled *Sig.* then you can see that both chi-square and likelihood ratio tests agree that removing this interaction will not significantly affect the fit of the model (because $p > .05$).

K-Way and Higher-Order Effects

	K	df	Likelihood Ratio Chi-Square	Likelihood Ratio Sig.	Pearson Chi-Square	Pearson Sig.	Number of Iterations
K-way and Higher Order Effects[a]	1	47	411.393	.000	400.923	.000	0
	2	34	50.930	.031	51.094	.030	2
	3	11	8.841	.637	8.850	.636	3
K-way Effects[b]	1	13	360.463	.000	349.829	.000	0
	2	23	42.089	.009	42.244	.009	0
	3	11	8.841	.637	8.850	.636	0

a. Tests that k-way and higher order effects are zero.

b. Tests that k-way effects are zero.

Output 19

The next part of the table expresses the same thing but without including the higher-order effects. It's labelled *K-Way Effects* and lists tests for when K = 1, 2 and 3. The first row (K = 1), therefore, tests whether removing the main effects (the one-way effects) has a significant detrimental effect on the model. The p-values are less than .05, indicating that if we removed the main effects of star sign, believer and true from our model it would significantly affect the fit of the model (in other words, one or more of these effects is a significant predictor of the data). The second row (K = 2) tests whether removing the two-way interactions has a significant detrimental effect on the model. The p-values are less than .05, indicating that if we removed the star sign × believer, star sign × true and believer × true interactions then this would significantly reduce how well the model fits the data. In other words, one or more of these two-way interactions is a significant predictor of the data. The final row (K = 3) tests whether removing the three-way interaction has a significant detrimental effect on the model. The p-values are greater than .05, indicating that if we removed the star sign × believer × true interaction then this would not significantly reduce how well the model fits the data. In other words, this three-way interaction is not a significant predictor of the data. This row should be identical to the final row of the upper part of the table (the *K-Way and Higher-Order Effects*) because it is the highest-order effect and so in the previous table there were no higher-order effects to include in the test (look at the output and you'll see the results are identical).

What this is actually telling us is that the three-way interaction is not significant: removing it from the model does not have a significant effect on how well the model fits the data. We also know that removing all two-way interactions does have a significant effect on the model, as does removing the main effects, but you have to remember that loglinear analysis should be done hierarchically and so these two-way interactions are more important than the main effects.

The *Partial Associations* table simply breaks down the table that we've just looked at into its component parts. So, for example, although we know from the previous output that removing all of the two-way interactions significantly affects the model, we don't know which of the two-way interactions is having the effect. This table tells us. We get a Pearson chi-square test for each of the two-way interactions and the main effects, and the column labelled *Sig.* tells us which of these effects is significant (values less than .05 are significant). We can tell from this that the star sign × believe and believe × true interactions are significant, but the star sign × true interaction is not. Likewise, we saw in the previous output that removing the one-way effects also significantly affects the fit of the model, and these findings are confirmed here because the main effect of star sign is highly significant (although this just means that we collected different amounts of data for each of the star signs!).

Partial Associations

Effect	df	Partial Chi-Square	Sig.	Number of Iterations
Star_Sign*Believe	11	20.666	.037	2
Star_Sign*True	11	10.740	.465	2
Believe*True	1	12.541	.000	2
Star_Sign	11	358.550	.000	2
Believe	1	1.582	.209	2
True	1	.331	.565	2

Output 3

The final bit of output deals with the backward elimination. SPSS begins with the highest-order effect (in this case, the star sign × believe × true interaction), remove it from the model, see what effect this has, and, if this effect is not significant, move on to the next highest effects (in this case the two-way interactions). As we've already seen, removing the three-way interaction does not have a significant effect, and this is confirmed at this stage by the table labelled Step Summary, which confirms that removing the three-way interaction has a non-significant effect on the model. At step 1, the three two-way interactions are then assessed in the bit of the table labelled *Deleted Effect*. From the values of *Sig.* it's clear that the star sign × believe (p = .037) and believe × true (p = .000) interactions are significant but the star sign × true interaction (p = 0. 465) is not. Therefore, at step 2 the non-significant star sign × true interaction is deleted, leaving the remaining two-way interactions in the model. These two interactions are then re-evaluated and both the star sign × believe (p = .049) and believe × true (p = .001) interactions are still significant and so are still retained. Therefore, the final model is the one that retains all main effects and these two interactions. As neither of these interactions can be removed without affecting the model, and these interactions involve all three of the main effects (the variables star sign, true and believe are all involved in at least one of the remaining interactions), the main effects are not examined (because their effect is confounded with the interactions that have been retained). Finally, SPSS evaluates this final model with the likelihood ratio statistic and we're looking for a non-significant test statistic, which indicates that the expected values generated by the model are not significantly different from the observed data (put another way, the model is a good fit of the data). In this case the result is very non-significant, indicating that the model is a good fit of the data.

Step Summary

Step[a]			Effects	Chi-Square[c]	df	Sig.	Number of Iterations
0	Generating Class[b]		Star_Sign*Believe*True	.000	0	.	
	Deleted Effect	1	Star_Sign*Believe*True	8.841	11	.637	3
1	Generating Class[b]		Star_Sign*Believe, Star_Sign*True, Believe*True	8.841	11	.637	
	Deleted Effect	1	Star_Sign*Believe	20.666	11	.037	2
		2	Star_Sign*True	10.740	11	.465	2
		3	Believe*True	12.541	1	.000	2
2	Generating Class[b]		Star_Sign*Believe, Believe*True	19.582	22	.609	
	Deleted Effect	1	Star_Sign*Believe	19.737	11	.049	2
		2	Believe*True	11.612	1	.001	2
3	Generating Class[b]		Star_Sign*Believe, Believe*True	19.582	22	.609	

a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than .050.

b. Statistics are displayed for the best model at each step after step 0.

c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

Output 21

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	19.582	22	.609
Pearson	19.533	22	.612

Output 22

The believe × true interaction

The next step is to try to interpret these interactions. The first useful thing we can do is to collapse the data. Remember from the chapter that there are the following rules for collapsing data: (1) the highest-order interaction should be non-significant; and (2) at least one of the lower-order interaction terms involving the variable to be deleted should be non-significant. We need to look at star sign × believe and believe × true interaction. Let's take the believe × true interaction first. Ideally we want to collapse the data across the star sign variable. To do this the three-way interaction must be non-significant (it was) and at least one lower-order interaction involving star sign must be also (the star sign × true interaction was). So, we can look at this interaction by doing a chi-square on believe and true, ignoring star sign. The results are below:

Did Their Horoscope Come True? * Do They Believe? Crosstabulation

			Do They Believe?		
			Unbeliever	Believer	Total
Did Their Horoscope Come True?	Horoscope Didn't Come True	Count	582	532	1114
		Expected Count	542.1	571.9	1114.0
		% of Total	26.4%	24.2%	50.6%
	Horoscope Came True	Count	489	598	1087
		Expected Count	528.9	558.1	1087.0
		% of Total	22.2%	27.2%	49.4%
Total		Count	1071	1130	2201
		Expected Count	1071.0	1130.0	2201.0
		% of Total	48.7%	51.3%	100.0%

Output 23

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	11.601[b]	1	.001		
Continuity Correction[a]	11.312	1	.001		
Likelihood Ratio	11.612	1	.001		
Fisher's Exact Test				.001	.000
Linear-by-Linear Association	11.596	1	.001		
N of Valid Cases	2201				

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 528.93.

Output 24

This chi-square is highly significant. To interpret this we could consider calculating some odds ratios. First, the odds of the horoscope coming true given that the person was a believer were 598/532. However, the odds of the horoscope coming true given that the person was a non-believer were 489/582. Therefore, the odds ratio is (598/532) ÷ (489/582) = 1.34. We can interpret this by saying that the odds that a horoscope would come true were 1.34 times higher in believers than non-believers. Given that the horoscopes were made-up twaddle this might be evidence that believers behave in ways to make their horoscopes come true!

The star sign × believe interaction

Next, we can look at the star sign × believe interaction. For this interaction we'd like to collapse across the true variable, To do this: (1) the highest-order interaction should be non-significant (which it is); and (2) at least one of the lower-order interaction terms involving the variable to be deleted should be non-significant (the star sign × true interaction was). So, we can look at this interaction by doing a chi-square on star sign and believe, ignoring true. The results are below:

Star Sign * Do They Believe? Crosstabulation

			Do They Believe?		
			Unbeliever	Believer	Total
Star Sign	Capricorn	Count	102	110	212
		Expected Count	103.2	108.8	212.0
		% within Star Sign	48.1%	51.9%	100.0%
	Aquarius	Count	46	51	97
		Expected Count	47.2	49.8	97.0
		% within Star Sign	47.4%	52.6%	100.0%
	Pisces	Count	106	134	240
		Expected Count	116.8	123.2	240.0
		% within Star Sign	44.2%	55.8%	100.0%
	Aries	Count	78	124	202
		Expected Count	98.3	103.7	202.0
		% within Star Sign	38.6%	61.4%	100.0%
	Taurus	Count	98	91	189
		Expected Count	92.0	97.0	189.0
		% within Star Sign	51.9%	48.1%	100.0%
	Gemini	Count	118	88	206
		Expected Count	100.2	105.8	206.0
		% within Star Sign	57.3%	42.7%	100.0%
	Cancer	Count	160	179	339
		Expected Count	165.0	174.0	339.0
		% within Star Sign	47.2%	52.8%	100.0%
	Leo	Count	37	32	69
		Expected Count	33.6	35.4	69.0
		% within Star Sign	53.6%	46.4%	100.0%
	Virgo	Count	124	115	239
		Expected Count	116.3	122.7	239.0
		% within Star Sign	51.9%	48.1%	100.0%
	Libra	Count	53	58	111
		Expected Count	54.0	57.0	111.0
		% within Star Sign	47.7%	52.3%	100.0%
	Scorpio	Count	52	56	108
		Expected Count	52.6	55.4	108.0
		% within Star Sign	48.1%	51.9%	100.0%
	Sagittarius	Count	97	92	189
		Expected Count	92.0	97.0	189.0
		% within Star Sign	51.3%	48.7%	100.0%
Total		Count	1071	1130	2201
		Expected Count	1071.0	1130.0	2201.0
		% within Star Sign	48.7%	51.3%	100.0%

Output 25

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	19.634[a]	11	.051
Likelihood Ratio	19.737	11	.049
Linear-by-Linear Association	2.651	1	.103
N of Valid Cases	2201		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 33.58.

Output 26

This chi-square is borderline significant (two-tailed, but then again we had no prediction so we need to look at the two-tailed significance). It doesn't make a lot of sense to compute odds ratios because there are so many star signs (although we could use one star sign as a base category and compute odds ratios for all other signs compared to this category). However, the obvious general interpretation of this effect is that the ratio of believers to unbelievers in certain star signs is different. For example, in most star signs there is a roughly 50–50 split of believers and unbelievers, but for Aries there is a 40–60 split and it is probably this difference that is most contributing to the effect. However, it's important to keep this effect in perspective. It may not be that interesting that we happened to sample a different ratio of believers and unbelievers in certain star signs (unless you believe that certain star signs should have more cynical views of horoscopes than others!). We actually set out to find out something about whether the horoscopes would come true, and it's worth remembering that this interaction ignores the crucial variable that measured whether or not the horoscope came true!

Reporting the results

For this example we could report as follows:

- ✓ The three-way loglinear analysis produced a final model that retained the star sign × believe and believe × true interactions. The likelihood ratio of this model was $\chi^2(22) = 19.58$, $p = .61$. The star sign × believe interaction was significant, $\chi^2(11) = 19.74$, $p < .05$. This interaction indicates that the ratio of believers and unbelievers was different across the 12 star signs. In particular, the ratio in Aries (38.6–62.4 ratio of unbelievers to believers) was quite different from the other groups, which consistently had a roughly 50–50 split. The believe × true interaction was also significant, $\chi^2(1) = 11.61$, $p < .001$. The odds ratio indicated that the odds of the horoscope coming true were 1.34 times more likely in believers than non-believers. Given that the horoscopes were made-up twaddle, this might be evidence that believers behave in ways to make their horoscopes come true.

Task 11

On my statistics course students have weekly SPSS classes in a computer laboratory. These classes are run by postgraduate tutors but I often pop in to help out. I've noticed in these sessions that many students are studying Facebook more than the very *interesting statistics assignments that I have set them. I wanted to see the impact that this behaviour had on their exam performance. I collected data from all 260 students on my course. I checked their **Attendance** and classified them as having attended either more or less than 50% of their lab classes. Next, I classified them as being either someone who looked at **Facebook** during their lab class, or someone who never did. Lastly, after the exam, I classified them as having either passed or failed (**Exam**). The data are in **Facebook.sav**. Do a loglinear analysis on the data to see if there is an association between studying Facebook and failing your exam.*

Running the analysis

Data are entered for this example as frequency values for each combination of categories, so before you begin you must weight the cases by the variable **Frequency**. If you don't do this the entire output will be wrong! 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 **Frequency**) 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 **Frequency**.

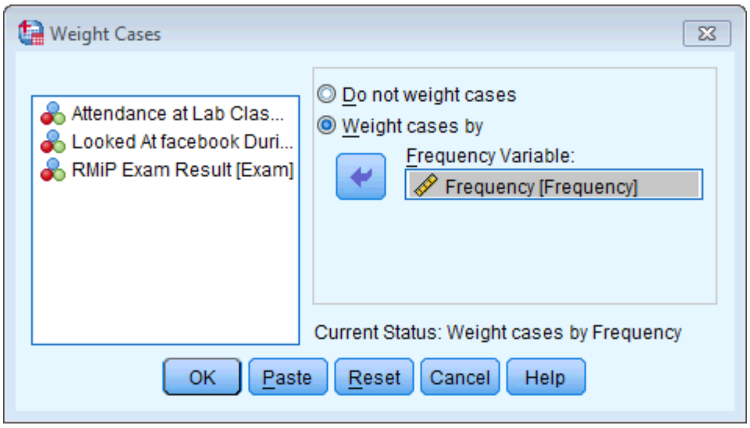

Figure 16

To get a crosstabulation table, select Analyze Descriptive Statistics Crosstabs... . We have three variables in our crosstabulation table: whether someone looked at Facebook during their lab classes (**Facebook**), whether they attended more than 50% of classes (**Attendance**) and whether they passed or failed their RMiP exam (**Exam**). Select **Facebook** and drag it into the box labelled *Row(s)* (or click on). Next, select **Exam** and drag it to the box labelled *Column(s)* (or click on). We need to define our third variable as a layer. Select **Attendance** and drag it to the box labelled *Layer 1 of 1* (or click on). Then click on Cells... and select the options required.

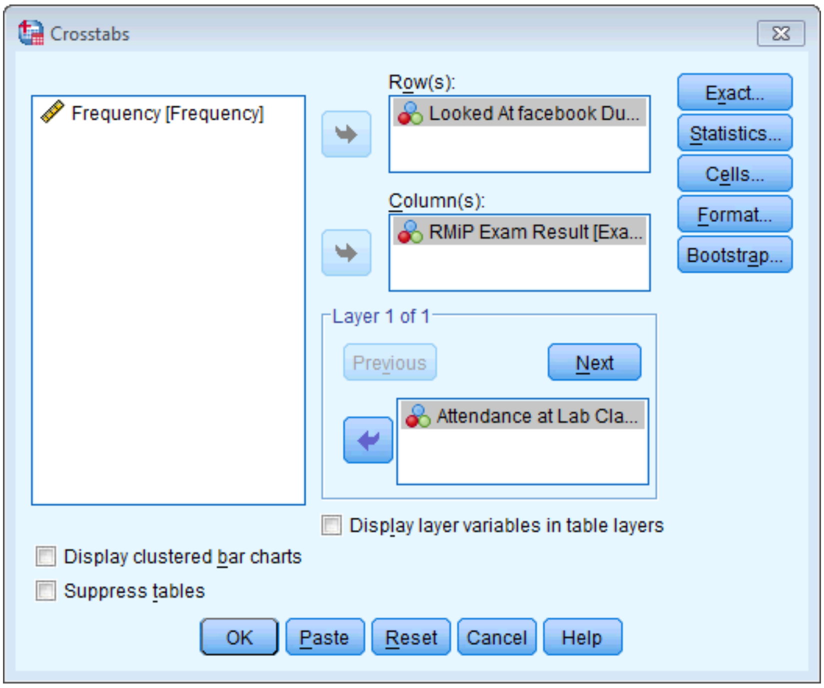

Figure 17

The crosstabulation table produced by SPSS contains the number of cases that fall into each combination of categories. There are no expected counts less than 5, so our assumptions are met.

Looked At facebook During lab Classes * RMiP Exam Result * Attendance at Lab Classes Crosstabulation

Attendance at Lab Classes				RMiP Exam Result: Pass	Fail	Total
More than 50%	Looked At facebook During lab Classes	Looked at Facebook	Count	39	30	69
			Expected Count	55.0	14.0	69.0
			% within Looked At facebook During lab Classes	56.5%	43.5%	100.0%
			% within RMiP Exam Result	28.5%	85.7%	40.1%
			Std. Residual	-2.2	4.3	
		Did Not Look at Facebook	Count	98	5	103
			Expected Count	82.0	21.0	103.0
			% within Looked At facebook During lab Classes	95.1%	4.9%	100.0%
			% within RMiP Exam Result	71.5%	14.3%	59.9%
			Std. Residual	1.8	-3.5	
		Total	Count	137	35	172
			Expected Count	137.0	35.0	172.0
			% within Looked At facebook During lab Classes	79.7%	20.3%	100.0%
			% within RMiP Exam Result	100.0%	100.0%	100.0%
Less than 50%	Looked At facebook During lab Classes	Looked at Facebook	Count	5	30	35
			Expected Count	12.3	22.7	35.0
			% within Looked At facebook During lab Classes	14.3%	85.7%	100.0%
			% within RMiP Exam Result	16.1%	52.6%	39.8%
			Std. Residual	-2.1	1.5	
		Did Not Look at Facebook	Count	26	27	53
			Expected Count	18.7	34.3	53.0
			% within Looked At facebook During lab Classes	49.1%	50.9%	100.0%
			% within RMiP Exam Result	83.9%	47.4%	60.2%
			Std. Residual	1.7	-1.3	
		Total	Count	31	57	88
			Expected Count	31.0	57.0	88.0
			% within Looked At facebook During lab Classes	35.2%	64.8%	100.0%
			% within RMiP Exam Result	100.0%	100.0%	100.0%

Output 27

The loglinear analysis

Then run the main analysis. The way to run loglinear analysis that is consistent with my section on the theory of the analysis is to select Analyze Loglinear ▸ Model Selection... to access the dialog box. Select any variable that you want to include in the analysis by clicking on them with the mouse (remember that you can select several at the same time by holding down the *Ctrl* key) and then dragging them to the box labelled *Factor(s)* (or click on ➡). When there is a variable in this box the Define Range... button becomes active. We have to tell SPSS the codes that we've used to define our categorical variables. Select a variable in the *Factor(s)* box and then click on Define Range... to activate a dialog box that allows you to specify the value of the minimum and maximum code that you've used for that variable. When you've done this, click on Continue to return to main dialog box.

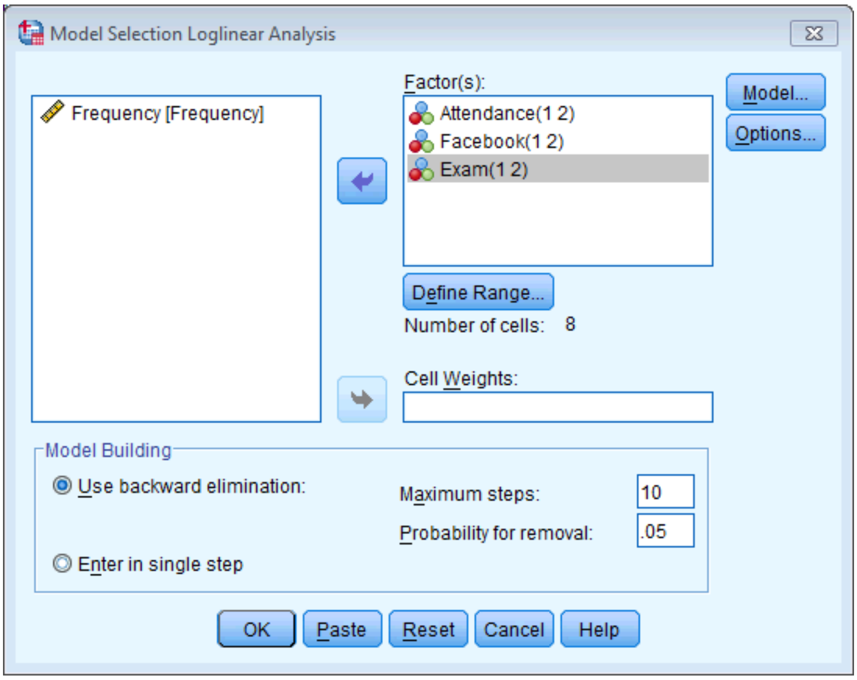

Figure 18

Output from loglinear analysis

K-Way and Higher-Order Effects

	K	df	Likelihood Ratio Chi-Square	Likelihood Ratio Sig.	Pearson Chi-Square	Pearson Sig.	Number of Iterations
K-way and Higher Order Effects[a]	1	7	161.895	.000	182.462	.000	0
	2	4	101.250	.000	88.083	.000	2
	3	1	1.574	.210	1.624	.203	8
K-way Effects[b]	1	3	60.645	.000	94.379	.000	0
	2	3	99.676	.000	86.459	.000	0
	3	1	1.574	.210	1.624	.203	0

a. Tests that k-way and higher order effects are zero.

b. Tests that k-way effects are zero.

Output 28

The first bit of the output labelled *K-Way and Higher-Order Effects* shows likelihood ratio and chi-square statistics when K = 1, 2 and 3 (as we go down the rows of the table). The first row (K = 1) tells us whether removing the one-way effects (i.e., the main effects of attendance, Facebook and exam) and any higher-order effects will significantly affect the fit of the model. There are lots of higher-order effects here – there are the two-way interactions and the three-way interaction – and so this is basically testing whether if we remove everything from the model there will be a significant effect on the fit of the model. This is highly significant because the p-value is given as .000, which is less than .05. The next row of the table (K = 2) tells us whether removing the two-way interactions (i.e., Attendance × Exam, Facebook × Exam and Attendance × Facebook) and any higher-order effects will affect the model. In this case there is a higher-order effect (the three-way interaction) so this is testing whether removing the two-way interactions *and* the three-way interaction would affect the fit of the model. This is significant (the p-value is given as .000, which is less than .05), indicating that if we removed the two-way interactions and the three-way interaction then this would have a significant detrimental effect on the model. The final row (*K* = 3) is testing whether removing the three-way effect *and* higher-order effects will significantly affect the fit of the model. The three-way interaction is of course the highest-order effect that we have, so this is simply testing whether removal of the three-way interaction (Attendance × Facebook × Exam) will significantly affect the fit of the model. If you look at the two columns labelled *Sig.* then you can see that both chi-square and likelihood ratio tests agree that removing this interaction will not significantly affect the fit of the model (because the $p > .05$).

The next part of the table expresses the same thing but without including the higher-order effects. It's labelled *K-Way Effects* and lists tests for when *K* = 1, 2 and 3. The first row (*K* = 1), therefore, tests whether removing the main effects (the one-way effects) has a significant detrimental effect on the model. The *p*-values are less than .05, indicating that if we removed the main effects of Attendance, Facebook and Exam from our model it would significantly affect the fit of the model (in other words, one or more of these effects is a significant predictor of the data). The second row (*K* = 2) tests whether removing the two-way interactions has a significant detrimental effect on the model. The *p*-values are less than .05, indicating that if we removed the two-way interactions then this would significantly reduce how well the model fits the data. In other words, one or more of these two-way interactions is a significant predictor of the data. The final row (*K* = 3) tests whether removing the three-way interaction has a significant detrimental effect on the model. The *p*-values are greater than .05, indicating that if we removed the three-way interaction then this would not significantly reduce how well the model fits the data. In other words, this three-way interaction is not a significant predictor of the data. This row should be identical to the final row of the upper part of the table (the *K-way and Higher-Order Effects*) because it is the highest-order effect and so in the previous table there were no higher-order effects to include in the test (look at the output and you'll see the results are identical).

Partial Associations

Effect	df	Partial Chi-Square	Sig.	Number of Iterations
Attendance*Facebook	1	11.896	.001	2
Attendance*Exam	1	61.801	.000	2
Facebook*Exam	1	49.765	.000	2
Attendance	1	27.631	.000	2
Facebook	1	10.470	.001	2
Exam	1	22.543	.000	2

Output 29

Parameter Estimates

Effect	Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval Lower Bound	95% Confidence Interval Upper Bound
Attendance*Facebook*Exam	1	-.119	.092	-1.295	.195	-.299	.061
Attendance*Facebook	1	.284	.092	3.090	.002	.104	.463
Attendance*Exam	1	.612	.092	6.667	.000	.432	.792
Facebook*Exam	1	-.538	.092	-5.862	.000	-.718	-.358
Attendance	1	.209	.092	2.282	.022	.030	.389
Facebook	1	-.084	.092	-.912	.362	-.264	.096
Exam	1	.174	.092	1.899	.058	-.006	.354

Output 4

Step Summary

Step[a]		Effects	Chi-Square[c]	df	Sig.	Number of Iterations
0	Generating Class[b]	Attendance*Facebook*Exam	.000	0	.	
	Deleted Effect 1	Attendance*Facebook*Exam	1.574	1	.210	8
1	Generating Class[b]	Attendance*Facebook, Attendance*Exam, Facebook*Exam	1.574	1	.210	
	Deleted Effect 1	Attendance*Facebook	11.896	1	.001	2
	2	Attendance*Exam	61.801	1	.000	2
	3	Facebook*Exam	49.765	1	.000	2
2	Generating Class[b]	Attendance*Facebook, Attendance*Exam, Facebook*Exam	1.574	1	.210	

a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than .050.

b. Statistics are displayed for the best model at each step after step 0.

c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

Output 31

The main effect of Attendance was significant, $\chi^2(1) = 27.63$, $p < .001$, indicating (based on the contingency table) that significantly more students attended over 50% of their classes ($N = 39 + 30 + 98 + 5 = 172$) than attended less than 50% ($N = 5 + 30 + 26 + 27 = 88$).

The main effect of Facebook was significant, $\chi^2(1) = 10.47$, $p < .01$, indicating (based on the contingency table) that significantly fewer students looked at Facebook during their classes ($N = 39 + 30 + 5 + 30 = 104$) than did not look at Facebook ($N = 98 + 5 + 26 + 27 = 156$).

The main effect of Exam was significant, $\chi^2(1) = 22.54$, $p < .001$, indicating (based on the contingency table) that significantly more students passed the RMiP exam ($N = 39 + 98 + 5 + 26 = 168$) than failed ($N = 39 + 98 + 5 + 26 = 92$).

The Attendance × Exam interaction was significant, $\chi^2(1) = 61.80$, $p < .01$, indicating that whether you attended more or less than 50% of classes affected exam performance. To illustrate, here's the contingency table:

Attendance at Lab Classes * RMiP Exam Result Crosstabulation

			RMiP Exam Result		
			Pass	Fail	Total
Attendance at Lab Classes	More than 50%	Count	137	35	172
		Expected Count	111.1	60.9	172.0
		% within Attendance at Lab Classes	79.7%	20.3%	100.0%
		Std. Residual	2.5	-3.3	
	Less than 50%	Count	31	57	88
		Expected Count	56.9	31.1	88.0
		% within Attendance at Lab Classes	35.2%	64.8%	100.0%
		Std. Residual	-3.4	4.6	
	Total	Count	168	92	260
		Expected Count	168.0	92.0	260.0
		% within Attendance at Lab Classes	64.6%	35.4%	100.0%

Output 32

This shows that those who attended more than half of their classes had a much better chance of passing their exam (nearly 80% passed) than those attending less than half of their classes (only 35% passed). All of the standardized residuals are significant, indicating that all cells contribute to this overall association.

The Facebook × Exam interaction was significant, $\chi^2(1) = 49.77$, $p < .001$, indicating that whether you looked at Facebook or not affected exam performance. To illustrate, here's the contingency table:

Looked At facebook During lab Classes * RMiP Exam Result Crosstabulation

			RMiP Exam Result		
			Pass	Fail	Total
Looked At facebook During lab Classes	Looked at Facebook	Count	44	60	104
		Expected Count	67.2	36.8	104.0
		% within Looked At facebook During lab Classes	42.3%	57.7%	100.0%
		Std. Residual	-2.8	3.8	
	Did Not Look at Facebook	Count	124	32	156
		Expected Count	100.8	55.2	156.0
		% within Looked At facebook During lab Classes	79.5%	20.5%	100.0%
		Std. Residual	2.3	-3.1	
	Total	Count	168	92	260
		Expected Count	168.0	92.0	260.0
		% within Looked At facebook During lab Classes	64.6%	35.4%	100.0%

Output 33

This shows that those who looked at Facebook had a much lower chance of passing their exam (58% failed) than those who didn't look at Facebook during their lab classes (around 80% passed).

The Facebook × Attendance × Exam interaction was not significant, $\chi^2(1) = 1.57$, $p = .20$. This result indicates that the effect of Facebook (described above) was the same (roughly) in those who attended more than 50% of classes and those that attended less than 50% of classes. In other words, although those attending less than 50% of classes did worse than those attending more than 50%, within that group, those looking at Facebook did relatively worse than those not looking at Facebook.