Chapter 13: Answers

Task 1

A psychologist was interested in the cross species differences between men and dogs. She observed a group of dogs and a group of men in a naturalistic setting (20 of each). She classified several behaviours as being dog-like (urinating against tress and lampposts, attempts to copulate with anything that moved, and attempts to lick their own genitals). For each man and dog she counted the number of dog-like behaviours displayed in a 24 hour period. It was hypothesized that dogs would display more dog-like behaviours than men. The data are in the file **MenLikeDogs.sav**, analyse them with a Mann-Whitney test.

SPSS Output

Ranks

	Species	N	Mean Rank	Sum of Ranks
Dog-Like Behaviour	Dog	20	20.77	415.50
	Man	20	20.23	404.50
	Total	40		

Test Statistics[b]

	Dog-Like Behaviour
Mann-Whitney U	194.500
Wilcoxon W	404.500
Z	-.150
Asymp. Sig. (2-tailed)	.881
Exact Sig. [2*(1-tailed Sig.)]	.883[a]

a. Not corrected for ties.

b. Grouping Variable: Species

Calculating an Effect-Size

The output tells us that z is –.15, and we had 20 men and 20 dogs so the total number of observations was 40. The effect size is, therefore:

$$r = \frac{-0.15}{\sqrt{40}}$$
$$= -0.02$$

This represents a tiny effect (it is close to zero), which tells us that there truly isn't much difference between dogs and men.

Writing and Interpreting the Results

We could report something like:

- ✓ Men (*Mdn* = 27) did not seem to differ from dogs (*Mdn* = 24) in the amount of dog-like behaviour they displayed (U = 194.5, *ns*).

Note that I've reported the median for each condition. Of course, we really ought to include the effect size as well. We could do two things. The first is to report the z-score associated with the test statistic. This value would enable the reader to determine both the exact significance of the test, and to calculate the effect size r:

- ✓ Men (*Mdn* = 27) and dogs (*Mdn* = 24) did not significantly differ in the extent to which they displayed dog-like behaviours (U = 194.5, *ns*, z = -.15).

The alternative is to just report the effect size (because readers can convert back to the z-score if they need to for any reason). This approach is better because the effect size will probably be most useful to the reader.

- ✓ Men (*Mdn* = 27) and dogs (*Mdn* = 24) did not significantly differ in the extent to which they displayed dog-like behaviours (*U* = 194.5, *ns*, *r* = -.02).

Task 2

There's been much speculation over the years about the influence of subliminal messages on records. To name a few cases, both Ozzy Osbourne and Judas Priest have been accused of putting backward masked messages on their albums that subliminally influence poor unsuspecting teenagers into dong things like blowing their heads off with shotguns. A psychologist was interested in whether backward masked messages really did have an effect. He took the master tapes of Britney Spears' 'Baby one more time' and created a second version that had the masked message 'deliver your soul to the dark lord' repeated in the chorus. He took this version, and the original and played one version (randomly) to a group of 32 people. He took the same group 6 months later and played them whatever version they hadn't heard the time before. So each person heard both the original, and the version with the masked message, but at different points in time. The psychologist measured the number of goats that were sacrificed in the week after listening to each version. It was hypothesized that the backward message would lead to more goats being sacrificed. The data are in the file **DarkLord.sav**, analyse them with a Wilcoxon signed-rank test.

SPSS Output

Ranks

		N	Mean Rank	Sum of Ranks
No Message - Message	Negative Ranks	11[a]	10.14	111.50
	Positive Ranks	17[b]	17.32	294.50
	Ties	4[c]		
	Total	32		

a. No Message < Message

b. No Message > Message

c. Message = No Message

Test Statistics[b]

	No Message - Message
Z	-2.094[a]
Asymp. Sig. (2-tailed)	.036

a. Based on negative ranks.

b. Wilcoxon Signed Ranks Test

Calculating an Effect-Size

The output tells us that *z* is –2.094, and we had 64 observations (although we only used 32 people and tested them twice it is the number of observations, not the number of people, that is important here). The effect size is, therefore:

$$r = \frac{-2.094}{\sqrt{64}}$$
$$= -0.26$$

This represents a medium effect (it is close to Cohen's benchmark of 0.3), which tells us that the effect of whether or a subliminal message was present was a substantive effect.

Writing and Interpreting the Results

We could report something like:

- ✓ The number of goats sacrificed after hearing the message (*Mdn* = 9) was significantly less than after hearing the normal version of the song (*Mdn* = 11) (T = 111.50, $p < .05$).

As with the Mann-Whitney test we should report either the z-score, or the effect size. The effect size is most useful:

- ✓ The number of goats sacrificed after hearing the message (*Mdn* = 9) was significantly less than after hearing the normal version of the song (*Mdn* = 11) (T = 111.50, $p < .05$, r = -.26).

Task 3

A psychologist was interested in the effects of television programs on domestic life. She hypothesised that through 'learning by watching', certain programs might actually encourage people to behave like the characters within them. This in turn could affect the viewer's own relationships (depending on whether the program depicted harmonious or dysfunctional relationships). She took episodes of three popular TV shows, and showed them to 54 couples after which the couple were left alone in the room for an hour. The experimenter measured the number of times the couple argued. Each couple viewed all three of the TV programs at different points in time (a week apart) and the order in which the programs were viewed was counterbalanced over couples. The TV programs selected were Eastenders (which typically portrays the lives of extremely miserable, argumentative, London folk who like nothing more than to beat each other up, lie to each other, sleep with each others wives and generally show no evidence of any consideration to their fellow humans!), Friends (which portrays a group of unrealistically considerate and nice people who love each other oh so very much—but for some reason I love it anyway!), and a National Geographic program about whales (this was supposed to act as a control). The data are in the file **Eastenders.sav**; access them and conduct Friedman's ANOVA on the data.

SPSS Output

Ranks

	Mean Rank
Eastenders	2.29
Friends	1.81
National Geographic	1.91

The first table shows the mean rank in each condition. These mean ranks are important later for interpreting any effects; they show that the ranks were highest after watching Eastenders.

Test Statistics[a]

N	54
Chi-Square	7.586
df	2
Asymp. Sig.	.023

a. Friedman Test

The next table shows the chi-square test statistic and its associated degrees of freedom (in this case we had 3 groups so the degrees of freedom are 3–1, or 2), and the significance. Therefore, we could conclude that the type of program watched significantly affected the subsequent number of arguments (because the significance value is less than 0.05). However, this result doesn't tell us exactly where the differences lie.

A nice succinct set of comparisons would be to compare each group against the control:

- ✓ Test 1: Eastenders compared to control
- ✓ Test 2: Friends compared to control

This gives rise to only two tests, so rather than use 0.05 as our critical level of significance, we'd use 0.05/2 = 0.025.

Ranks

		N	Mean Rank	Sum of Ranks
National Geographic - Eastenders	Negative Ranks	31[a]	28.85	894.50
	Positive Ranks	18[b]	18.36	330.50
	Ties	5[c]		
	Total	54		
National Geographic - Friends	Negative Ranks	21[d]	22.00	462.00
	Positive Ranks	24[e]	23.88	573.00
	Ties	9[f]		
	Total	54		

a. National Geographic < Eastenders

b. National Geographic > Eastenders

c. Eastenders = National Geographic

d. National Geographic < Friends

e. National Geographic > Friends

f. Friends = National Geographic

Test Statistics[c]

	National Geographic - Eastenders	National Geographic - Friends
Z	-2.813[a]	-.629[b]
Asymp. Sig. (2-tailed)	.005	.530

a. Based on positive ranks.

b. Based on negative ranks.

c. Wilcoxon Signed Ranks Test

The next tables show the test statistics from doing Wilcoxon tests on the two comparisons that I suggested. Remember that we are now using a critical value of 0.025, so we compare the significance of both test statistics against this critical value. The test comparing Eastenders to the National Geographic program about whales has a significance value of 0.005, which is well below our criterion of 0.025, therefore, we can conclude that Eastenders led to significantly more arguments than the program about whales. The second comparison compares the number of arguments after Friends with the number after the program about whales. This contrast is non-significant (the significance of the test statistic is 0.530, which is bigger than our critical value of 0.025), so we can conclude that there was no difference in the number of arguments after watching friends compared to after watching the whales. The effect we got seems to mainly reflect the fact that Eastenders makes people argue more.

Calculating an Effect-Size

We can calculate effect sizes for the Wilcoxon tests that we used to follow up the main analysis.

For the first comparison (Eastenders vs. control) z is –2.813, and because this based on comparing two groups each containing 54 observations, we had 108 observations in total (remember it isn't important that the observations come from the same people). The effect size is, therefore:

$$r_{Eastenders-Control} = \frac{-2.813}{\sqrt{108}} = -0.27$$

This represents a medium effect (it is close to Cohen's benchmark of 0.3), which tells us that the effect of Eastenders relative to the control was a substantive effect: Eastenders produced substantially more arguments.

For the second comparison (Friends vs. control) z is –0.629, and this was again based on 108 observations. The effect size is, therefore:

$$r_{Story-Control} = \frac{-0.629}{\sqrt{108}} = -0.06$$

This represents virtually no effect (it is close to zero). Therefore, friends had very little effect in creating arguments compared to the control.

Writing and Interpreting the Results

For Friedman's ANOVA we need only report the test statistic (which we saw earlier is denoted by χ^2), its degrees of freedom and its significance. So, we could report something like:

- ✓ The number of arguments that couples had was significantly affected by the program they had just watched ($\chi^2(2) = 7.59$, $p < .05$).

We need to report the follow up tests as well (including their effect sizes):

- ✓ The number of arguments that couples had was significantly affected by the program they had just watched ($\chi^2(2) = 7.59$, $p < .05$). Wilcoxon tests were used to follow-up this finding. A Bonferroni correction was applied and so all effects are reported at a .025 level of significance. It appeared that watching Eastenders significantly affected the number of arguments compared to the program about whales ($T = 330.50$, $r = -.27$). However, the number of arguments was not significantly different after Friends compared to after the program about whales ($T = 462$, *ns*, $r = -.06$). We can conclude that watching Eastenders did produce significantly more arguments compared to watching a program about whales, and this effect was medium in size. However, Friends didn't produce any substantial reduction in the number of arguments relative to the control program.

Task 4

A researcher was interested in trying to prevent coulrophobia (fear of clowns) in children. She decided to do an experiment in which different groups of children (15 in each) were exposed to different forms of positive information about clowns. The first group watched some adverts for McDonald's in which their mascot Ronald McDonald is seen cavorting about with children going on about how they should love their mum. A second group was told a story about a clown who helped some children when they got lost in a forest (although what on earth a clown was doing in a forest remains a mystery). A third group was entertained by a real clown, who came into the classroom and made balloon animals for the children[1]. A final group acted as a control condition and they had nothing done to them at all. The researcher took self-report ratings of how much the children liked clowns (rather like the fear-beliefs questionnaire in chapter 2) resulting in a score for each child that could range from 0 (not scared of clowns at all) to 5 (very scared of clowns). The data are in the file **coulrophobia.sav**; access the data and conduct a Kruskal-Wallis test.

SPSS Output

[1] Unfortunately, the first time they attempted the study the clown accidentally burst one of the balloons. The noise frightened the children and they associated that fear response with the clown. All 15 children are currently in therapy for coulrophobia!

Ranks

	Format of Information	N	Mean Rank
Fear beliefs	Advert	15	45.03
	Story	15	21.87
	Exposure	15	23.77
	None	15	31.33
	Total	60	

This table tells us the mean rank in each condition. These mean ranks are important later for interpreting any effects.

Test Statistics[a,b]

	Fear beliefs
Chi-Square	17.058
df	3
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: Format of Information

This table shows this test statistic (SPSS labels it chi-square rather than *H*) and its associated degrees of freedom (in this case we had 4 groups so the degrees of freedom are 4–1, or 3), and the significance (which is less than the critical value of 0.05). Therefore, we could conclude that the type of information presented to the children about clowns significantly affected their fear ratings of clowns.

A nice succinct set of comparisons would be to compare each group against the control:

- *Test 1: Advert compared to control*
- *Test 2: Story compared to control*
- *Test 3: Exposure compared to control*

This results in three tests, so rather than use 0.05 as our critical level of significance, we'd use 0.05/3 = 0.0167. The following tables show the test statistics from doing Mann-Whitney tests on the three focussed comparisons that I suggested.

Advert vs. control:

Test Statistics[b]

	Fear beliefs
Mann-Whitney U	37.500
Wilcoxon W	157.500
Z	-3.261
Asymp. Sig. (2-tailed)	.001
Exact Sig. [2*(1-tailed Sig.)]	.001[a]

a. Not corrected for ties.

b. Grouping Variable: Format of Information

Story vs. control:

Test Statistics[b]

	Fear beliefs
Mann-Whitney U	65.000
Wilcoxon W	185.000
Z	-2.091
Asymp. Sig. (2-tailed)	.037
Exact Sig. [2*(1-tailed Sig.)]	.050[a]

a. Not corrected for ties.

b. Grouping Variable: Format of Information

Exposure vs. control:

Test Statistics[b]

	Fear beliefs
Mann-Whitney U	72.500
Wilcoxon W	192.500
Z	-1.743
Asymp. Sig. (2-tailed)	.081
Exact Sig. [2*(1-tailed Sig.)]	.098[a]

a. Not corrected for ties.

b. Grouping Variable: Format of Information

Remember that we are now using a critical value of 0.0167, so the only comparison that is significant is when comparing the advert to the control group (because the observed significance value of 0.001 is less than 0.0167). The other two comparisons produce significance values that are greater than 0.0167 so we'd have to say they're non-significant.

So the effect we got seems to mainly reflect the fact that McDonald's adverts significantly increased fear beliefs about clowns relative to controls (which is no surprise given what a creepy weirdo Ronald McDonald is!).

Calculating an Effect-Size

We can calculate effect sizes for the Mann-Whitney tests that we used to follow up the main analysis. For the first comparison (adverts vs. control) the z is –3.261, and because this based on comparing two groups each containing 15 observations, we had 30 observations in total. The effect size is, therefore:

$$r_{Advert-Control} = \frac{-3.261}{\sqrt{30}} = -0.60$$

This represents a large effect, which tells us that the effect of adverts relative to the control was a substantive effect.

For the second comparison (story vs. control) z is –2.091, and this was again based on 30 observations. The effect size is, therefore:

$$r_{Story-Control} = \frac{-2.091}{\sqrt{30}} = -0.38$$

This represents a medium to large effect. Therefore, although non-significant the effect of stories relative to the control was a substantive effect.

For the final comparison (exposure vs. control) z is –1.743, and this was again based on 30 observations. The effect size is, therefore:

$$r_{Exposure-Control} = \frac{-1.743}{\sqrt{30}} = -0.32$$

This represents a medium effect. Therefore, although non-significant the effect of exposure relative to the control was a substantive effect.

Writing and Interpreting the Results

For the Kruskal-Wallis test, we need only report the test statistic (which we saw earlier is denoted by H), its degrees of freedom and its significance. So, we could report something like:

- ✓ Children's fear beliefs about clowns was significantly affected the format of information given to them ($H(3) = 17.06$, $p < .01$).

However, we need to report the follow up tests as well (including their effect sizes):

- ✓ Children's fear beliefs about clowns was significantly affected the format of information given to them ($H(3) = 17.06$, $p < .01$). Mann-Whitney tests were used to follow-up this finding. A Bonferroni correction was applied and so all effects are reported at a .0167 level of significance. It appeared that fear beliefs were significantly higher after the adverts compared to the control ($U = 37.50$, $r = -.60$). However, fear beliefs were not significantly different after the stories ($U = 65.00$, *ns*, $r = -.38$) or exposure ($U = 72.5$, *ns*, $r = -.32$) relative to the control. We can conclude that clown information through stories and exposure did produce medium size effects in reducing fear beliefs about clowns, but not significantly so (future work with larger samples might be appropriate), but that Ronald McDonald was sufficient to significantly increase fear beliefs about clowns.