Writing Statistical Copy in APA Style

Presented below are some examples of correct presentation of statistical results in APA style. Clearly, there are other correct ways of presenting this material in terms of descriptions of the experimental design and basic style of expression. Some general points to note are:

1. Exact *p* levels should be reported where possible. When your SPSS output provides you with a significance level that consists of a string of zeroes (e.g., *p* = .000), the rule is: Drop the last zero and change it to a 1, and write *p* < [whatever]. For example, if your significance level on a correlation is sig=.000, you would write *p* < .001. It is quite common for SPSS output to provide you with significance levels that consist of strings of zeroes. You should *never* write *p* = .000, because this is statistically impossible.

2. With regard to decimal places, some general rules are: Report correlations to two decimal places (e.g. *r*(N = 120) = .89). Report *F* ratios and *t* scores to two decimal places. With significance levels, report to two decimal places if the first digit is not a zero (e.g. *p* = .34); if the first digit is a zero (e.g. *p* = .037) or if the first and second digits are zeroes (e.g. *p* = .003), report to three decimal places; only report to four decimal places if the first three digits are zeroes (e.g. *p* = .0006), or if *p* < .0001, although most SPSS procedures will not provide significance levels to four decimal places.

3. Only include leading zeroes (e.g. 0.56) if the statistic can take on a value greater than 1. Correlations and significance levels never take leading zeroes. *t* and *F* values do.

4. With many assumption tests the norm is to present significance levels only, although there is nothing wrong with providing complete information, and many SPSS procedures provide you with complete information for assumption tests. The Levene test, for example, is an *F* ratio, and can be reported in full. Similarly, for post-hoc procedures, there is no need to provide detailed statistical information, and most SPSS procedures will not provide you will detailed results anyway; depending on which version of SPSS you are using, all you might get are significance levels.

5. Provide correct Greek letters where possible (e.g., α, β, η, χ); it looks more professional. Greek letters are found under the Symbol font in MS Word (any version), or you can use Insert Symbol.

6. Different types of correlation coefficient are represented by the use of sub-scripts. For example, Pearson’s Product-Moment correlation is represented simply as *r*, but a point-biserial correlation is presented as *r*pb; a Kendall’s tau as *r*τ. In all other respects, correlations are presented identically.

7. Although it is not an APA requirement, you should provide some measure of effect size where possible. Many journals are now requiring the reporting of effect sizes each time a test of significance is reported.

8. Pages 20-27 and 136-146 of the APA Publication Manual (5th edition) provide clear information on correct presentation of statistical copy. Other information relevant to the reporting of statistical material is scattered throughout the APA Publication Manual in related sections; use the index to locate relevant material.

The APA symbol for a mean is *M*, and for a standard deviation, *SD*. If you want to be really smart, you can present these as small capitals, which is the preferred method of printing capital letters
when they are acronyms; for example, $M$, and, $SD$. Also, means and standard deviations are two of the few statistical results that are presented in parentheses:

Scores for males ($M = 11.45$, $SD = 4.56$) were higher than those for females ($M = 9.88$, $SD = 2.78$).

Finally, please note that this guide is provided for students and at many levels of data analysis expertise, ranging from undergraduate to higher degree by research. It’s important to understand that not all of the material provided here is going to be necessary for you to provide in the context of any assignment that you might be doing. For example, in the analysis of covariance, the test of the assumption of homogeneity of slopes is quite an advanced topic that is only taught at higher levels, as is testing for simple main effects in factorial ANOVA, and the various measures of effect size. When using this guide in the context of an assignment, you should only feel obliged to provide the results that you have been taught to use.

**A Chi-Square Contingency Table**

A contingency table analysis of sex with voting preference revealed a significant relationship between these two variables, $\chi^2 (3, N = 101) = 35.15$, $p = .019$, $V = .25$. Examination of standardised residuals indicated that the high proportion of women voting labour (standardised residual = 2.4) contributed to the significant result.

*Notes*: The “$V$” is a measure of strength of association/effect size (Cramer’s $V$).

**An Independent Samples $t$-Test With Associated Assumption Test**

A Levene test found that the assumption of homogeneity of variance was met, $p = .71$; therefore a two-tailed independent samples $t$-test based on equal variances was carried out. No significant sex difference in sequential processing ability was found, $t(99) = 1.53$, $p = .13$, $d = 0.12$, 95%CI (0.06, 0.18).

*Notes*: This result includes the effect size measure, Cohen’s $d$, and a 95% confidence interval around that effect size measure. Also note, that the SPSS $t$-test procedure doesn’t provide much info on the Levene test, so all you are able to report is the $p$ level. The EXPLORE procedure provides full ANOVA output for the Levene test.

**A Matched Samples $t$-Test**

A two-tailed paired samples $t$-test found no significant difference between left- and right-hand reaction time, $t(100) = 1.47$, $p = .14$, $d = 0.18$, 95%CI (0.08, 0.28).
Notes: This result includes the effect size measure, Cohen’s $d$, and a 95% confidence interval around that effect size measure.

A Correlation Coefficient

There was a significant positive correlation between State and Trait Anxiety, $r(N = 125) = .68$, $p < .001$, $r^2 = .46$.

Notes: APA style does not require an indicator of sample size to be included with the result, but I prefer it. It is acceptable to leave the “$N =$” out of the parentheses (e.g., $r(125)$). Note that the symbol for sample size for a complete sample is $N$. Different types of correlation coefficient are identified by subscripts, which are not presented in italics; for example, a point biserial correlation is $r_{pb}$; an intra-class correlation is $r_{ic}$, and so on. I have also included $r^2$ with this result. This is not common, but it is informative, because it shows the proportion of shared variability associated with the correlation.

A One-Way Independent Groups (or Single-Factor Between-Subjects) ANOVA With Post-Hoc Tests

A single-factor between-subjects ANOVA was used to analyse the relationship between the four memory enhancement methods and memory test performance; a significant overall treatment effect was found, $F(3, 36) = 63.12$, $p < .001$, $\eta^2 = .65$. Subsequent post-hoc tests using Tukey’s HSD procedure ($\alpha = .05$) revealed significant differences at $p < .001$ between all possible pairwise comparisons among the treatment groups.

Notes: In the above example, all of the post-hoc tests were significant at $p < .001$, and I was able to report the results of the post-hoc tests with a single summary statement. However, this will often not be the case, and I would be obliged to report individual post-hoc test results. This result incorporates a measure of the effect size, eta-squared, $\eta^2$.

Planned Comparisons With Associated Assumption Test

Results were analysed using three a priori between-subjects planned comparisons. Because a Levene Test found that the homogeneity of variance assumption had been violated, $p = .001$, hypothesis tests were based on unequal variances. A significant effect was found for the first comparison, which contrasted the control group with the combined effect of the three treatment groups, $t(10.3) = 8.40$, $p < .0001$, $d = 1.46$, 95%CI (0.86, 2.06). The second test compared the two cognitive-based methods, imagery and mnemonics; this comparison was also significant, $t(12.4) = 4.75$, $p < .0001$, $d$
Finally, the drug-based method was compared against the average effect of the two cognitive-based methods; again, a significant effect was found, \( t(21.4) = 9.79, p < .0001, d = 1.30, 95\%CI (0.95, 1.65) \).

*Notes:* This result includes the effect size measure, Cohen’s \( d \), and a 95% confidence interval around that effect size measure. Note the fractional degrees of freedom for the main tests; this is common when reporting the results based on unequal variances.

### A Between-Subjects Factorial ANOVA

Data were analysed using a 4 x 2 between-subjects factorial ANOVA. The first factor consisted of the four memory enhancement methods (control, drug, imagery, mnemonics). The second factor was sex (male, female). A significant interaction was found between the two factors, \( F(3, 32) = 4.88, p = .007, \) partial \( \eta^2 = .24 \). Because the interaction term was found to be significant, main effects were not considered. Follow-up testing of the interaction effect using simple main effects found a significant difference among the four treatments for both males, \( F(3, 33) = 41.25, p < .001, \) partial \( \eta^2 = .79 \), and females, \( F(3, 33) = 27.88, p < .001, \) partial \( \eta^2 = .72 \). There were no significant sex differences at any of the four levels of memory program, although the strongest effects were seen within the control subjects, partial \( \eta^2 = .08 \), and the mnemonic subjects, partial \( \eta^2 = .03 \).

*Notes:* For these results, the effect size statistic, partial \( \eta^2 \), has been provided. Partial \( \eta^2 \) is easily obtained from SPSS. There is a difference between partial \( \eta^2 \) and \( \eta^2 \), but for functional purposes, they mean the same thing. Also, note that full statistical information has not been provided for the non-significant simple main effects. It is acceptable to do this, although were I writing this for a report, I would probably provide full statistical information. Finally, note the practice of considering the interaction test first and, if that is significant, ignoring the main effects and focusing exclusively on follow-up procedures related to the interaction. Although this approach is statistically "pure", it is not incorrect to consider the overall main effects as well as the interaction. However, if the interaction is significant, it should be the focus of any follow-up analyses.

### A Single-Factor Between-Subjects Analysis of Covariance (ANCOVA)

Data from the four memory enhancement conditions were analysed using a single-factor between-subjects analysis of covariance (ANCOVA), with IQ test scores serving as a covariate. A test of the assumption of homogeneity of slopes revealed no significant interaction between IQ and the four treatment groups, \( F(1,35) = 1.32, p = .68, \) partial \( \eta^2 = .05 \). Results showed that IQ test scores covaried significantly with the
dependent variable, \( F(1,35) = 295.85, p < .001, \) partial \( \eta^2 = .75. \) After partialling out the variance associated with IQ test scores, there was a significant difference among the four treatment groups, \( F(3, 35) = 4.28, p = .011, \) partial \( \eta^2 = .28 \) Post-hoc testing using pairwise comparisons of the estimated marginal means with Bonferroni adjusted \( \alpha \) levels revealed a significant difference between Group 1 and Groups 3 and 4 (both \( p < .001 \)), Group 1 and Group 2, \( p = .041 \), and Group 2 and Group 4, \( p = .005 \). The comparisons between Group 3 and Group 4, \( p = .08 \), and Group 2 and Group 3, \( p = .14 \), were non-significant.

Notes: The effect size partial \( \eta^2 \) has been provided. Also, note the assumption test for the important assumption of homogeneity of slopes. Finally, note the complete detail on all relevant significance levels provided for the post-hoc tests.

A Single-Factor Within-Subjects (aka One-Way Repeated Measures) ANOVA With Assumption Test and Post-Hoc Tests

In order to test for any differences among the four essay-production methods, a single-factor within-subjects analysis of variance was performed on the essay quality scores. The multivariate approach to the analysis of within-subjects data was used. A significant difference was found among the four essay production conditions, Wilks’ \( \Lambda = .003, F(3, 37) = 3573.05, p < .001, \) multivariate \( \eta^2 = .88. \) Post-hoc testing was carried out using pairwise comparisons of estimated marginal means with Bonferroni adjusted \( \alpha \) levels. Significant differences at \( p < .001 \) were found for all but one of the six possible pairwise comparisons. There was no significant difference between written essays and those produced using speech-recognition, \( p = .73. \)

Notes: Note that the effect size statistic here is “multivariate \( \eta^2 \)” not “partial \( \eta^2 \”).

A Single-Factor Between-Subjects Multivariate Analysis of Variance (MANOVA)

To investigate differences among the four memory enhancement conditions, the following four dependent variables were entered into a single-factor between-subjects multivariate analysis of variance (MANOVA): scores on the memory test, Stanford-Binet IQ scores, scores on the experimental IQ test, and scores on the test of learning anxiety. A significant multivariate effect was found, Wilks’ \( \Lambda = .052, F(12, 87.6) = 15.01, p < .001, \) multivariate \( \eta^2 = .35. \) Follow-up univariate
Analyses of each dependent variable found a significant difference among the four memory enhancement methods for only one of the dependent variables—Stanford-Binet IQ test scores, $F(3, 96) = 14.07, p < .001, \eta^2 = .28$. Post-hoc testing amongst the four groups on this dependent variable using Tukey’s HSD procedure revealed significant differences at $p < .001$ for all six post-hoc comparisons.

Notes: For the follow-up univariate tests, significance levels only have been provided for the significant result, although were I writing this up for a thesis or a paper I would certainly provide complete statistical information for the non-significant results as well. Finally, note the steps that have taken place in this analysis: 1. The multivariate analysis is carried out. 2. If this is significant, you then move onto the analysis of each of the dependent variables separately. 3. If any of these are significant, you report the appropriate post-hoc tests.

The following books also provide excellent examples of writing up the results of analyses in APA style:

