

STATISTICS FOR PSYCHOLOGISTS

PART IV: STATISTICS IN APA STYLE

Section Abstract: This section describes basic rules for presenting statistical results in APA style. All rules come from the newest APA style manual. Specific examples of mini Results summaries are provided, using the analyses in the previous section of this project. Sample data tables are provided for more complex examples.

Keywords: APA style, Results sections, statistical interpretation

Section Updated: December 2014

This document is part of an online statistics textbook.
Access to the complete textbook, along with licensing information, is available online:
<http://www4.uwsp.edu/psych/cw/statistics/textbook.htm>

Table of Contents for This Section

GENERAL RULES FOR APA STYLE RESULTS SECTIONS	2
SUMMARY OF PARAMETRIC STATISTICS	3
EXAMPLES OF APA STYLE RESULTS IN THE TEXT	4
EXAMPLES OF APA STYLE RESULTS IN TABLES	6

GENERAL RULES FOR APA STYLE RESULTS SECTIONS

Overview

The APA manual describes appropriate strategies for presenting statistical information. These guidelines were established to provide basic minimal standards and to provide some uniformity across studies.

Using a Sufficient Set of Statistics

Information to Include: Significance testing “is but a starting point and that additional reporting elements such as effect sizes, confidence intervals, and extensive description are needed to convey the most complete meaning of the results” (p. 33).

1. Descriptive statistics are essential and “such a set usually includes at least the following: the per-cell sample sizes; the observed cell means (or frequencies of cases in each category for a categorical variable); and the cell standard deviations” (p. 33).
2. For statistical significance tests, “include the obtained magnitude or value of the test statistics, the degrees of freedom, the probability of obtaining a value as extreme as or more extreme than the one obtained (the exact p value)” (p. 34).
3. When possible, confidence intervals should be emphasized. “The inclusion of confidence intervals (for estimates of parameters, for functions of parameters such as differences in means, and for effect sizes) can be an extremely effective way of reporting results” (p. 34).
4. “For the reader to appreciate the magnitude or importance of a study’s findings, it is almost always necessary to include some measure of effect size” (p. 34). These can be in the original (raw) units or in a standardized metric.

Information in Text versus Data Displays: “Statistical and mathematical copy can be presented in text, in tables, and in figures. . . Select the mode of presentation that optimizes understanding of the data by the reader” (p. 116).

- Generally speaking, the more data you have, the more likely it is that they should be presented in a table or figure. “If you need to present 4 to 20 numbers, first consider a well-prepared table” (p. 116).
- “If you present descriptive statistics in a table or figure, you do not need to repeat them in text, although you should (a) mention the table in which the statistics can be found and (b) emphasize particular data in the narrative when they help in interpretation” (p. 117).
- As a result, it is necessary that the text include a description of the variable(s) under study and a description of the statistical procedures used. The text often includes a description of whether the results support the hypotheses.

SUMMARY OF PARAMETRIC STATISTICS

Statistic	Purpose	APA Style	Description
Descriptive Statistics			
Mean	To provide an estimate of the population from which the sample was selected.	$M = \underline{\hspace{1cm}}$	Indicates the center point of the distribution and serves as the reference point for nearly all other statistics.
Standard Deviation	To provide an estimate of the amount of variability/dispersion in the distribution of population scores.	$SD = \underline{\hspace{1cm}}$	Indicates the variability of scores around their respective mean. Zero indicates no variability.
Measures of Effect Size			
Cohen's d	To provide a standardized measure of an effect (defined as the difference between two means).	$d = \underline{\hspace{1cm}}$.	Indicates the size of the treatment effect relative to the within-group variability of scores.
Correlation	To provide a measure of the association between two variables measured in a sample.	$r(df) = \underline{\hspace{1cm}}$	Indicates the strength of the relationship between two variables and can range from -1 to +1.
Eta-Squared	To provide a standardized measure of an effect (defined as the relationship between two variables).	$\eta^2 = \underline{\hspace{1cm}}$.	Indicates the proportion of variance in the dependent variable accounted for by the independent variable.
Confidence Intervals			
CI for a Mean	To provide an interval estimate of the population mean. Can be derived from both the z and t distributions.	$\underline{\hspace{1cm}}\% \text{ CI } [\underline{\hspace{1cm}}, \underline{\hspace{1cm}}]$	Indicates that there is the given probability that the interval specified covers the true population mean.
CI for a Mean Difference	To provide an interval estimate of the population mean difference. Can be derived from both the z and t distributions.	$\underline{\hspace{1cm}}\% \text{ CI } [\underline{\hspace{1cm}}, \underline{\hspace{1cm}}]$	Indicates that there is the given probability that the interval specified covers the true population mean difference.
Significance Tests			
One Sample t Test	To compare a single sample mean to a population mean when the population standard deviation is not known	$t(df) = \underline{\hspace{1cm}}, p = \underline{\hspace{1cm}}$.	A small probability is obtained when the statistic is sufficiently large, indicating that the two means significantly differ from each other.
Independent Samples t Test	To compare two sample means when the samples are from a single-factor between-subjects design.		
Related Samples t Test	To compare two sample means when the samples are from a single-factor within-subjects design.		
One-Way ANOVA	To compare two or more sample means when the means are from a single-factor between-subjects design.	$F(df_1, df_2) = \underline{\hspace{1cm}}, p = \underline{\hspace{1cm}}$.	A small probability is obtained when the statistic is sufficiently large, indicating that the set of means differ significantly from each other.
Repeated Measures ANOVA	To compare two or more sample means when the means are from a single-factor within-subjects design.		
Factorial ANOVA	To compare four or more groups defined by a multiple variables in a factorial research design.		

Note. Many of the statistics from each of the categories are frequently and perhaps often appropriately presented in tables or figures rather than in the text.

EXAMPLES OF APA STYLE RESULTS IN THE TEXT

Descriptive Statistics: The purpose of the descriptive statistics is to provide the reader with an idea about the basic elements of the group(s) being studied. Note that this also forms the basis of the in-text presentation of descriptive statistics for the inferential analyses below.

On the quiz, the nine students had a mean score of 7.00 ($SD = 1.23$). Scores of 6.00, 7.00, and 8.00 represented the 25th, 50th, and 75th percentiles, respectively.

Correlations: Correlations provide a measure of statistical relationship between two variables. Note that correlations can be tested for statistical significance (and that this information should be summarized if it is available and of interest).

For the nine students, the scores on the first quiz ($M = 7.00$, $SD = 1.23$) and the first exam ($M = 80.89$, $SD = 6.90$) were strongly and significantly correlated, $r(8) = .70$, $p = .038$.

One Sample t Test: In this case, a sample mean has been compared to a user-specified test value (or a population mean). Thus, the summary and the inferential statistics focus on that difference.

A one sample t test showed that the difference in quiz scores between the current sample ($N = 9$, $M = 7.00$, $SD = 1.23$) and the hypothesized value (6.000) were statistically significant, $t(8) = 2.45$, $p = .040$, 95% CI [0.06, 1.94], $d = 0.82$.

Independent Samples t Test: For this analysis, the emphasis is on comparing the means from two groups. Here again the summary and the inferential statistics focus on the difference.

An independent sample t test showed that the difference in quiz scores between the control group ($n = 4$, $M = 6.00$, $SD = 0.82$) and the experimental group ($n = 4$, $M = 8.00$, $SD = .82$) were statistically significant, $t(6) = -3.46$, $p = .013$, 95% CI [-3.41, -0.59], $d = -2.45$.

One Way ANOVA: The ANOVA provides an omnibus test of the differences across multiple groups. Because the ANOVA tests the overall differences among the groups, the text discusses the differences only in general.

A one way ANOVA showed that the differences in quiz scores between the control group ($n = 3$, $M = 4.00$, $SD = 1.00$), the first experimental group ($n = 3$, $M = 8.00$, $SD = 1.00$), and the second experimental group ($n = 3$, $M = 9.00$, $SD = 1.00$) were statistically significant, $F(2,6) = 21.00$, $p = .002$, $\eta^2 = .88$.

One Way ANOVA with Post Hoc Tests: Post hoc tests build on the ANOVA results and provide a more focused comparison among the groups. Notice that the post hoc summary duplicates the presentation of the omnibus ANOVA statistics.

A one way ANOVA showed that the difference in quiz scores between the control group ($n = 3$, $M = 4.00$, $SD = 1.00$), the first experimental group ($n = 3$, $M = 8.00$, $SD = 1.00$), and the second experimental group ($n = 3$, $M = 9.00$, $SD = 1.00$) were statistically significant, $F(2,6) = 21.00$, $p = .002$, $\eta^2 = .88$. Tukey's HSD tests showed that both experimental groups scored statistically significantly higher than the control group. However, the two experimental groups did not differ significantly.

Repeated Measures ANOVA: The RMD ANOVA tests for overall differences across the repeated measures. As such, its summary parallels that of the One Way ANOVA.

A repeated measures ANOVA showed that, for the five people, the difference in quiz scores between the first time point ($M = 6.40$, $SD = 1.14$) and second time point ($M = 7.80$, $SD = 0.84$) were statistically significant, $F(1,4) = 32.67$, $p = .005$, partial $\eta^2 = .88$.

Factorial ANOVA: The factorial ANOVA provides statistics for all of the main effects and interactions in a factorial design. Each effect would be summarized in a style analogous to a One Way ANOVA.

A 2 (Factor A) x 2 (Factor B) ANOVA was conducted on the quiz scores. Neither Factor A, $F(1,8) = 0.00$, $p = 1.00$, partial $\eta^2 = .00$, nor Factor B, $F(1,8) = .75$, $p = .412$, partial $\eta^2 = .09$, had a statistically significant impact on quiz scores. However, the interaction was statistically significant, $F(1,8) = 6.75$, $p = .032$, partial $\eta^2 = .46$. The descriptive statistics for these analyses are presented in Table 1.

EXAMPLES OF APA STYLE RESULTS IN TABLES

In more complex examples – such as those with multiple outcome variables – it is often preferred to place statistical information in tables rather than in the text.

Descriptive Statistics with Correlations: This table provides the reader with both the univariate descriptive statistics (the means and standard deviations) and the bivariate descriptive statistics (the correlations).

Variable	<i>M</i>	<i>SD</i>	1	2	3
Variable 1	7.33	1.58	--		
Variable 2	7.56	1.74	.92	--	
Variable 3	8.11	1.45	.91	.81	--

Descriptive Statistics with Confidence Intervals: This table is useful for removing the basic descriptive statistics (the means and standard deviations) and the confidence intervals from the text.

Variable	Group or Condition 1		Group or Condition 2	
	<i>M (SD)</i>	95% CI	<i>M (SD)</i>	95% CI
Outcome Variable 1	5.67 (.58)	[4.23, 7.10]	7.33 (.58)	[5.90, 8.77]
Outcome Variable 2	5.67 (.58)	[4.23, 7.10]	7.67 (1.15)	[4.80, 10.54]
Outcome Variable 3	6.33 (.58)	[4.90, 7.77]	8.67 (.58)	[7.23, 10.10]

Independent and Dependent Samples *t* Tests: When conducting *t* tests for multiple variables, it is useful to place the majority of the analyses in the table, using the text to interpret the results.

Variable	Group or Time 1		Group or Time 2		<i>t</i> (4)	<i>p</i>	95% CI	Cohen's <i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>				
Outcome 1	5.67	.58	7.33	.58	-3.54	.024	[-2.98, -.36]	-2.86
Outcome 2	5.67	.58	7.67	1.15	-2.68	.055	[-4.07, .07]	-2.20
Outcome 3	6.33	.58	8.67	.58	-4.95	.008	[-3.64, -1.02]	-4.03

One Way and Repeated Measures ANOVAs with Post Hoc Tests: Just like with multiple t tests, it is useful to place the statistical information in a table. Note the use of symbols to indicate which groups were statistically significantly different using the post hoc tests.

Variable	Group or Time 1		Group or Time 2		Group or Time 3		$F(2, 6)$	p	η^2	Tukey's HSD
	M	SD	M	SD	M	SD				
Outcome 1	5.67	.58	7.33	.58	9.00	1.00	15.00	.005	.84	1, 2 < 3
Outcome 2	5.67	.58	7.67	1.15	9.33	.58	15.17	.005	.83	1, 2 < 3
Outcome 3	6.33	.58	8.67	.58	9.33	.58	22.33	.002	.88	1 < 2, 3