

Answers to Application Activities in Chapter 2

2.1.2 Application Activity: Practice in Identifying Levels of Measurement

1) Categorical (this starts out continuous but the researcher makes it categorical), continuous; 2) categorical, continuous; 3) continuous, continuous; 4) categorical, categorical, continuous; 5) continuous, categorical, categorical.

2.1.4 Application Activity: Practice Identifying Variables

1	Independent: proficiency level	Dependent: number of language learning strategies used
2	Independent: L1 background	Dependent: comprehensibility ratings
3	Independent: pronunciation training	Dependent: accurate perception of phonemic contrasts
4	This is a “neither” question—the authors wanted to explore correlations between oral language measures and reading comprehension measures	
5	Independent (there are 2): age, status as bilingual/monolingual	Dependent: score on Simon test
6	Independent (there are 2): AOA and proficiency level	Dependent: brain activation patterns
7	Another “neither” question—the author examined the correlation between cultural exposure before study abroad and during study abroad	

8	Independent (there are 2): proficiency level and sounds in phonemic contrasts	Dependent: accuracy in perceiving contrasts
9	Independent: Type of help given during the test—either none (called the transfer items) or some (computer-mediated hints)	Dependent: score on listening comprehension task
10	Independent (there are 2): time tested and group (study abroad or not)	Dependent: writing fluency measured as number of words written per minute

2.2.2 Application Activity: Creating Null Hypotheses

- 1 H₀: There will be no difference in fluency measures between a group that gets explicit instruction in noticing formulaic sequences and another which does not.
- 2 H₀: There will be no relationship between an oral language measure and reading fluency measure in English/Spanish for 4th grade bilinguals.
- 3 H₀: There will be no difference between groups who differ in AOA and proficiency level in which neural brain substrates are activated when performing a certain task.
- 4 H₀: Learners will not improve in their performance on an appropriate judgment questionnaire (AJQ) after receiving conceptual explanations and SCOBAs.
- 5 H₀: Sequence learning ability is not related to morphosyntactic abilities in an L2.

2.2.7 Application Activity: Understanding Statistical Reporting

1 Gurzynski-Weiss and Baralt (2013)

a) The t reports the result of a t -test; b) $df = n - 2$, so $23 = n - 2$, so $n = 23 + 2 = 25$ (going from the explanation in exercise #3(b) in Section 2.2.5); c) $p = .85$ means “the probability of finding a t this large or larger if the null hypothesis were true is 85 in 100.” This is a very high probability so we cannot reject the null hypothesis that there is no difference in the two conditions; d) First of all, it is negative, which is not important. We can basically ignore whether it is negative or positive. But second, it is quite small, much less than 2, so the p -value is large. e) This effect size is negligible. The d effect size would need to be at least $d = .4$ to call it a small effect size, so we may say there was basically no effect for differing conditions.

2 Polio, Fleck, and Leder (1998)

a) The F reports the results of an ANOVA (and it says that it is an ANOVA above the table!); b) It is really small, much smaller than 2; c) We can use the formula $df = k(n - 1)$ where $k = 2$ (# of groups) and $df = 63$. Solving for n , $n = 32.5$, and since there were 2 groups, that means there were overall 65 participants (obviously there couldn't be 23.5 participants in a group, and in reality there were 31 in one group and 34 in the other); d) $p = .935$ means “the probability of finding an F this large or larger if the null hypothesis were true is 93.5%.” This is a very large probability and so we cannot reject the null hypothesis that there is no difference between the two groups. e) This would be a negligible effect size, not even reaching the level of small (which would need to be at least $\eta^2 = .01$).

3 Rose and O'Neill (1999)

a) The χ^2 reports the results of a chi-square test; b) Using the formula given in exercise #4(b) in Section 2.2.5, we can calculate that there were 3 rows and 3 columns; c) $p < .05$ means “the probability of finding a χ^2 this large or larger if the null hypothesis were true is less than 5%.” By the way, it would be nice to know how much smaller this number is than 0.05, because if it were, say, .049 that would be very different from $p = .0001$. Anyway, the inference is that we can reject the null hypothesis and conclude that there is a relationship between formal instruction and type of verbal report.

4 Williams (2005)

a) The r reports the results of correlation tests; b) Williams does not include the N in this sentence so we do not know; c) $p < .05$ means “the probability of finding an r this large or larger if the null hypothesis were true is less than 5%.” We’d like to know the exact probability of each test, as a p -value of .049 would be quite different from one of .001. However, the inference is that the null hypothesis will be rejected in all cases and the alternative hypothesis that there are relationships between these variables will be accepted. Since the largest r (.357) has a p -value of somewhere less than .05, we might assume that Williams did not include the p -values for the previous two r -values (.01 and .263) because they were more than .05, and thus, not statistically significant. The beautiful thing about correlations is that the test statistic *is* the effect size, so we see the effect size in the results without any further work. This leads us to say that . . . d) The effect size for the first correlation is certainly negligible. The effect size for the second correlation should be viewed as small, and the effect size for the third correlation is close to

medium. Thus, even though we may assume that the p -value for the second correlation was greater than .05, it doesn't really matter since we can still see a noticeable effect size for it, and effect size is what we are really interested in.

2.2.9 Application Activity: The Inner Workings of Statistical Testing

- 1 This should be your own words, but you should say something about calculations involving mean scores, standard deviations, and sample sizes. The statistic is nothing more than a calculation performed on these parts.
- 2 This should be your own words, but you should say something about matching up the value of the statistic with a point on a curve (which is formally a continuous probability density curve) and then calculating the area under the curve (again, formally this will be the area under both the right and left sides of the curve). The area under the curve is the probability that you would find a statistic of the size you found, or larger, if the null hypothesis were true.