Rebecca Huselid, Hunter College, Psychology

Fall 13: QR Lesson Plan for Teaching Independent and Related-samples t-test Designs

Statistical Methods in Psychology (Psych248)

A pre-post design was used to compare my students understanding of two basic research designs. After students had read about and heard lecture descriptions of one-sample, independent and related-sample research designs I administered the pre-test worksheet in class – each student completed it independently. I graded them and returned them the next class period with individual feedback and had a class discussion about common mistakes. The intervention also included students working together in small groups to create original designs and to do statistical tests (by hand and using SPSS). The post-test quiz was given one week after the pre-test. Both pre- and post-tests were worth 10 points.

My primary goal of the intervention was for students to explain the differences between two common research designs: **independent-samples** and **related-samples**, recognize examples of both designs and generate original examples. The two quizzes required students to demonstrate the 3 skills listed in goals. They are attached (versions with grading rubrics also).

Results were calculated using data from 31 of 35 students who were in class both days and completed both quizzes. The mean score on the pretest was 4.87 and on the posttest it was 8.61 (out of 10 possible). A paired-sample t-test indicated that the increase in performance was statistically significant (t (30) = 6.80, p <.001). The average increase in performance was 3.7 points. The raw data and statistical analyses are also included in this summary.

I felt that this new approach to teaching these concepts was very effective. I did decide to use the intervention again during the following semester (S14). I was a little surprised and disappointed at how poorly students performed on the pre-test given that I had already lectured on the topic. However being able to define in their own words and create their own examples was a challenge for almost all of them (only 3 of 34 got 10). After the intervention, improvement was dramatic for all and half of them (n=17) scored a perfect 10. Although it was not part of the assessment, I think it helped students do better on subsequent exams covering these concepts throughout the semester. One thing I liked about the intervention was that I had them work together on aspects of the intervention, designing studies together. I think everyone benefited from this. The testing was done independently (students worked alone to complete each quiz after practicing together).

My biggest mistake was not having students complete the consent forms and questionnaires that were part of the NICHE assessment design. For this I sincerely apologize. I hope to continue to use and improve on this lesson plan and assessment strategy as I continue to teach statistics.

PSYCH248 Dr. Huselid: Brief PreTest –Independent and Related-Sample Designs

1. In your own words define both Related-Sample and Independent-Samples Designs – also called Within-Subjects and Between-Subject designs.
2. Imagine that you have been asked to design two studies comparing the reading scores of children in 2 third-grade classrooms: Class J-taught by Ms. J and Class K is taught by Mr. K.
3. How would you design it to use an Independent-t test to compare them?
4. How could you redesign it to use a Related-sample t test to compare reading scores?

3. Identifying Research Designs. Listen as your instructor reads descriptions of studies. Identify the type of test statistic appropriate for each one: independent t, related-samples t, or one-sample t. Briefly explain why it is the appropriate test.

Type of test Why?

Study 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

PSYCH248 Dr. Huselid: Post-Test: Independent and Related-Sample Designs

1. In your own words explain the key differences between a Repeated-Measures Design and an Independent-Samples Designs.

1b. Explain the key differences between Matched Pairs and Independent-Samples Designs.

2. Imagine that your job is to design a study to determine if playing violent video games makes kids more aggressive in a new game situation.

1. How would you design it to use an Independent-t test to compare them?
2. How could you design it to use a Related-sample t test to compare them?

3. Identifying Research Designs. Listen to the study descriptions. Identify the test statistic appropriate for each one: independent t, related-samples t, or one-sample t. Briefly explain why it is the appropriate test. Type of test Why?

Study 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 2 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

PSYCH248 Huselid: Revised Brief Assessment Plan - Limited to 3 key questions only

TIME TWO - PRETEST

From Computer Lab Worksheet Week 8: Review of Independent & Related t-test Designs

1. In your own words define both Related-Sample and Independent-Samples Designs – also called Within-Subjects and Between-Subject designs. (2 pts total – 0-1 for both)

GRADING RUBRIC:

Independent design: 1 = mentions random assignment or separate groups/different people

0 = includes neither

Related designs: 1 = mentions they a) same people tested twice, or b) are paired based on some pretest

Or they have some natural relationship to each other (married, siblings, friends)

0 = mentions nothing that clearly explains why they are related.

2. Imagine that you were asked to design a study comparing the reading scores of children in two third-grade classrooms: Class J is taught by Ms. J, and Class K is taught by Ms. K. (2 pts total - 1 or 0 both)

a. How could you design this so you could use an Independent-t test to compare them?

+1 : randomly assign the children to the two classrooms before class starts then do the reading test

Or

- compare the two separate groups of children – this is quasi-experimental if they are existing groups not randomly assigned

0 = not a design including two separate groups

b. How could you redesign this so you could use a Related-t test to compare them? (hint, don’t think repeated measures. What else can make them related or paired in some way?)

* Pair or match the children based on some initial reading readiness or IQ test (or any background variable) so a child in K is matched to a child in J with a similar score. Then use related t to test reading.
* If they are twins you could randomly assign one twin to each classroom – (wild but give credit for this) The analysis would still be related because they are natural pairs.
* If they took Mrs K for nine weeks – then they were taught by Mrs J for nine weeks – you could compare scores across time (repeated design) – but this has a confound – the passage of time and maturity of child.

3. Identifying Research Designs. Listen as your instructor reads descriptions of studies. Identify the type of test statistic appropriate for each one: independent t, related-samples t, or one-sample t. Briefly explain why it is appropriate and whether or not the hypothesis is two-tailed or one-tailed (directional).

(6 points total – Type of test (1 or 0) Why also 1 or 0.

Type of test Why?

Study 1 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 2 not included in assessment – covers a third t-test.

Study 3 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Study 4 \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Descriptions used for #3 Identify Research Designs.

1. A research study was conducted to examine the differences between older and younger adults on perceived life satisfaction. Ten older adults (over the age of 70) and ten younger adults (between 20 and 30) were given a life satisfaction test. What type of t-test should you use to compare older and younger adults? Now briefly explain why.

(Independent-sample t. Because they are two separate groups – not same people in both.)

***This one will be included in the lab but not in this assessment.***

1. You wish to determine whether New York City public school students score higher on average than the rest of the country on a standardized math exam. How would you compare the average math score of 5th graders nationally (which is known to be 85) with the average score of 5th graders in NYC? (which this past year was 86.5)? Now briefly explain why.

(One-sample t). Comparing one sample of kids to a known population value (all USA).

1. To improve the study habits of students, nine students were randomly selected to attend a Study skills seminar to teach them better study skills. To test whether it worked, a researcher recorded the hours per week each student studied both before and after the seminar. What type of test should you use to determine if the Study Skills Course changed study time? Now briefly explain why.

(Related-samples t). Because it is repeated-measures – sample people are reporting study time before and after the course. This also has an obvious design flaw – possible confound – demand characteristic that you could discuss with them.)

1. A researcher believes that sleep deprivation will lower cognitive performance. Participants spend the night in his lab and are randomly assigned to get either 3 hours of sleep or 8 hours of sleep. The next morning all take the Cognitive Ability Test with scores ranging from 1-30. How does he test whether the low sleep group did worse than normal sleep group? Now briefly explain why.

(Independent-samples t. These people were randomly assigned to condition – high or low sleep – two separate groups.)

SUMMARY – 12 points possible

#1 – both definitions are graded 0,1,2 so 4 points possible

#2 – both a and b are graded 0 or 1 - correct/incorrect 2 points possible

#3 – type of test and why both graded 0 or 1 6 points possible

TIME TWO – POST TEST – part of Lab 8

From Computer Lab Worksheet Week 8: Review of Independent & Related t-test Designs

1a. In your own words explain the key difference between a Repeated Measures Design and Independent-Samples Designs. 1b. Next please explain the key differences between Matched Pairs and Independent-Samples Design. (4 pts total – 0-2 for both)

GRADING RUBRIC: KEY DIFFERENCES

1a. Repeated-Measures/Independent:

2 = Independent designs include 2 separate groups of people – Repeated-measures designs include the same people tested twice.

1 = mentions one or the other but not both

0 = includes neither

1b. Matched Pairs/Independent:

2 = two separate groups of people in both designs. –Independent designs people in both groups are not related in any way - but in matched pairs they have been paired based on some pretest score (IQ) or they have some natural relationship to each other (married, siblings, friends)

1 = mentions 1 or 2 of the above aspects – but only partially correct

0 = mentions nothing that clearly explains why they are different.

2. Imagine that you were asked to design a study to determine if exposure to violent video games makes kids more aggressive in a new game situation – playing Wac-A-Mole (hitting a fake mole). (2 pts total - 1 or 0 both)

a. How could you design a study to test this using an Independent-t test to compare 2 groups

+1 for: randomly assign the children to two conditions – playing violent video game or no video game play then have them all do the Wac-A-Mole test

Or - compare a group of children who normally play lots of violent video games with a group who doesn’t play violent video games – this is quasi-experimental -they are existing groups not randomly assigned – then have them do the Wac-A-Mole test

0 = not a design including two separate groups

b. How could you design this study to use a Related-t test to compare them? (any related design is OK)

+1 Pair or match the children based on amount of video-game play they report – then compare scores on Wac-A-Mole aggresion.

+1 Do a repeated measures design – have child play a non violent game before Wac-A-Mole one day and have the same child play a violent game before Wac-A-Mole the next day.

0 - not a design including paired or repeated measures – not relevant

Study Descriptions for Posttest: Each test chosen is correct or incorrect. Why answers also graded 0 or 1.

Study 1: A real-estate agent claims that average rental prices for two-bedroom apartments are about the same in two neighborhoods of the city (NoHo vs SoHo). To test this claim, students randomly sample apartment complexes in each neighborhood and collect data on monthly rent from 15 apartments in each area. What type of test should they use to determine if the two neighborhoods differ in rental price?

(A: independent t. Two separate neighborhoods compared -– quasi-experimental – average price of rent is the DV – comparing 2 means.)

Study 2: At a swim meet each swimmer competes first in a “heat” and later in a “final” for the same event (such as 100 freestyle). A high school swim coach is interested in testing her hypothesis that times are faster for the second race. What test should she use to compare times in the heat and the final race? Why?

(A: repeated-measures t. It’s the same swimmer so compare the difference between their two scores.

REVIEW THIS - This will reduce variability due to how good each swimmer is. They are competing against themselves. If most swimmers end up faster then average time will go down for the team as a whole.)

Study 3: Do men and women have different priorities when choosing a career? Data were collected from a community sample of men and women using a questionnaire about Life Values. Assume the DV is normally distributed in the population. How would you test to see if ratings are different for men and women? Why choose this t-test?

(A: independent t. Two separate groups – men and women – quasi-experimental. REVIEW THIS Compare means of men and women. )

Study 4: Do Psychology majors have different personalities than other students? Data were collected from a sample of Psychology majors at Hunter and were compared to the known population average on a personality scale. Assume the DV is normally distributed in the population. How would you test to see if Psych Majors at Hunter are different from the average college student? Why choose this t-test?

(A: one-sample t. One sample compared with the population mean on a personality test.

SUMMARY – 10 points possible

#1 – both definitions are graded 0 or 1 2 points possible

#2 – both a and b are graded 0 or 1 - correct/incorrect 2 points possible

#3 – type of test and why both graded 0 or 1 6 points possible

Raw Data – Huselid F13

|  |  |
| --- | --- |
| PreTest | PostTest |
| 5 | 9 |
| 9 | 10 |
| 3 | 10 |
| 5 | 7 |
| 0 | 8 |
| 9 | 8 |
| 2 | 10 |
| 10 | 10 |
| 5 | 9 |
| 7 | 10 |
| 3 | 10 |
| 0 | 10 |
| 8 | 10 |
| 4 | 8 |
| 3 | 7 |
| 8 | 9 |
| 3 | 9 |
| 5 | 8 |
| 3 | 6 |
| 9 | 9 |
| 0 | 2 |
| 6 | 7 |
| 10 | 10 |
| NA | 10 |
| 0 | 6 |
| NA | 10 |
| 7 | 10 |
| 6 | 8 |
| 2 | 10 |
| 10 | 10 |
| 3 | 10 |
| 5 | 7 |
| 1 | 10 |
| NA | 10 |

**T-Test – SPSS Results F13 QR Assessment of Indep/Related t-tests**

[DataSet1] C:\Users\Rebecca Huselid Appl\Dropbox\P248\Niche Assessment\IndepRelated Assessment.sav

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Paired Samples Statistics** | | | | | |
|  | | Mean | N | Std. Deviation | Std. Error Mean |
| Pair 1 | PostTest | 8.61 | 31 | 1.801 | .324 |
| PreTest | 4.87 | 31 | 3.212 | .577 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Paired Samples Correlations** | | | | |
|  | | N | Correlation | Sig. |
| Pair 1 | PostTest & PreTest | 31 | .360 | .047 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Paired Samples Test** | | | | | | |
|  | | Paired Differences | | | | |
| Mean | Std. Deviation | Std. Error Mean | 95% Confidence Interval of the Difference | |
| Lower | Upper |
| Pair 1 | PostTest - PreTest | 3.742 | 3.066 | .551 | 2.617 | 4.866 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Paired Samples Test** | | | | |
|  | | t | df | Sig. (2-tailed) |
|
|
| Pair 1 | PostTest – PreTest | 6.796 | 30 | .000 |