Research Methods II

One-Way ANOVA

Stacey Coleman

Outline:

- Brief recap on design that use two groups/conditions
- Introduction to basic ANOVA designs
  - Within- and between-subjects
- ANOVA analysis

TERMS TO REMEMBER

FACTORS

These are your independent variables (IVs)

LEVELS (of factors)

Similar to conditions/groups

BETWEEN-SUBJECTS FACTOR

Factors whose levels vary between participants

Each participant only experiences one level of a factor (e.g., male/female)

WITHIN-SUBJECTS FACTOR

Factors whose various levels vary within a participant

Each participant experiences all levels of the factor (e.g., testing mood at 9am, 12pm and 3pm)

Levels of Variables

- As researchers, we are interested in asking questions and making choices…

For example..

How do two variables of interest relate/associate with each other?

- Does ‘task difficulty’ effect the number of errors made?
- Are ‘males or females’ better at multi-tasking?

Levels of Variables

- In many cases (as with sex group) we can create designs where the IV is:
    - Controlled for
    - Collated (recorded)
    - Grouped (to provide us with our levels of the IV)

- In order to make a **binary comparison** of DV's
    - Male score (mean) versus Female score (mean)
    - Difference testing (between-subjects t-test!)

Levels of Variables

- Many IV's can conform to this 'binary assumption', but some distinction between examples such as:

  - **IV – gender**
    - Natural variables category of two levels: male/female

  - **IV – age**
    - Controlled category of two levels: teen/adult

Levels of Variables

- Many studies deliberately design and control for a binary comparison of conditions:

  - E.g., age in a test of memory
    - Participants may range from 18 to 75 (not age ‘groups’)
    - May collate into groups: ‘old’ and ‘young’
    - Binary split method: all **above** mean = ‘old’; all **below** mean = ‘young’

      - **ISSUES?!**

- Different size groups
- Ages close to the mean not a good example of old or young…

Levels of Variables

- Deliberate sampling
  - Take participants between 18-25, and between 60-75; controlled group sizes
  - Natural variable (age) **re-coded** into two categories/levels of an IV

Levels of Variables

- Many studies deliberately control **stimuli** or **tasks** to create a comparison of two conditions:
  - In **experimentally controlled** conditions, for example
    - Similarity of items in working memory (similar/dissimilar)
    - Amount of time given (long/short)

Levels of Variables

- May also used **scores/ratings** produced by a participant to group in this way:
  - E.g., wish to analyse **amount of experience** of punishment (on a scale measurement) as an IV for it's effect on current attitudes (DV) towards punishment
  - Binary split could be used to group into 'high' and 'low' based on previous experience
    - Design issues previously considered should be assessed here though;
    - Difficult to 'deliberately' sample for such IV grouping

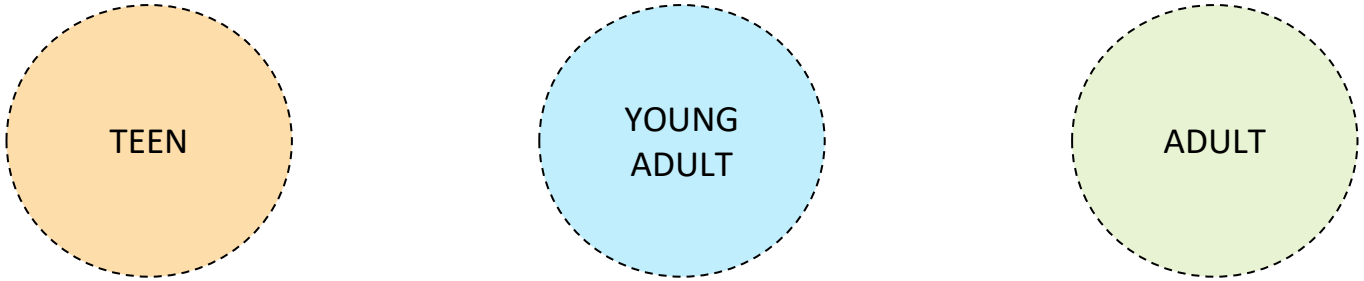
These are examples of **SIMPLE DESIGNS**

- These examples contain one IV (with two levels) and one DV

- The effects are judged or analysed in isolation using, for example, a t-test
  - Already familiar with t-test
    - Calculate statistic (t) to compare mean scores in two levels of an IV
    - The significance value of this comparison can then be judged and interpreted

Levels of Variables

- Issue of ‘ecology’
- Are binary categories reflecting the ecology (what is real in the environment/person):
  - Often variables may form more naturally into **more than two categories**
  - Often researchers wants to deliberately dissect a variable into **more than two levels**
  - For this, we need to build more complex models/theory, and thus use more complex analyses



Example: IV = Age

60 participants aged between 12 and 75

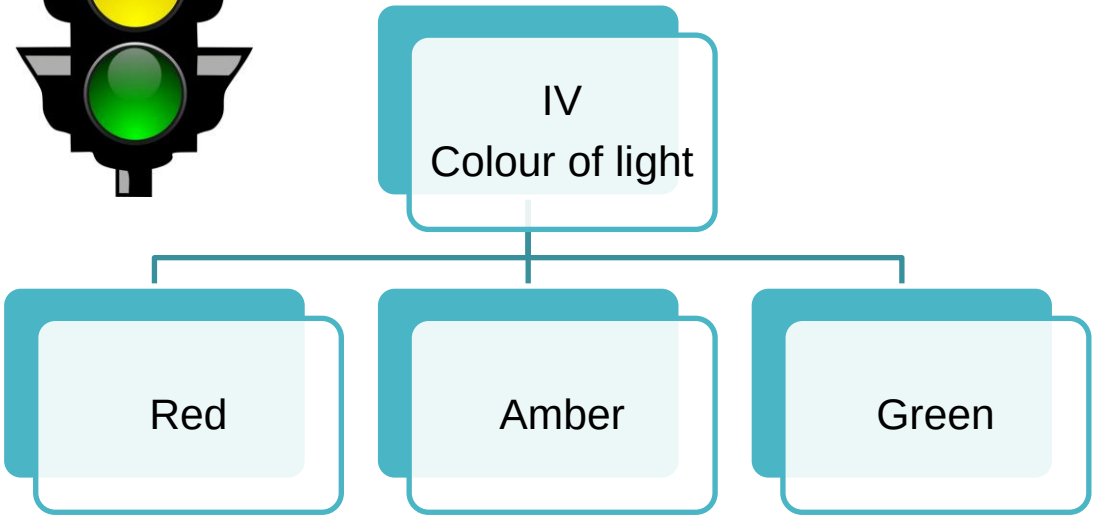
Consider splitting these participants into **three age groups** (so between-subjects)

Analysis of Variance (ANOVA)

- If an IV has more than two groups (between-subjects) or more than two conditions (within-subjects) then the Analysis of Variance (ANOVA) statistical technique is used to analyse the data.

- E.g., we want to explore how rapidly people respond to different coloured traffic lights: red, amber, and green

Analysis of Variance (ANOVA)



- Can either use between-subjects (each P takes part in only one of the colour conditions) or within-subjects (all P's respond to all coloured lights)

Analysis

- Looking to compare mean scores between **all three groups/conditions**

- ***Between-subjects one-way ANOVA***
- ***Within-subjects one-way ANOVA***

ANOVA

- We could analyse the difference between two conditions (e.g., red and green) using a paired samples t-test
- So why don't we just conduct **three t-tests** when there are **three conditions**?

Red vs. amber
Red vs. green
Amber vs. green

ANOVA

1. Too many multiple comparisons – increased likelihood of **Type I error**
   - False positive; accepting a difference as significant ($p < .05$) in error

ANOVA corrects for this problem:

- More stringent test of significant than the t-test
- Accounts for multiple comparisons in significance level

ANOVA

2. Does not tell us about the GENERAL effect of the IV

   - Does colour in general have an effect upon reaction time?

One-way ANOVA

- Used when we have one IV with more than two levels (conditions)
  - In this case, we have three; red, amber, and green
- May also be called Multi-level or Multi-factorial analysis

One-way ANOVA: Assumptions

- ANOVA is a parametric test, and there are a number of assumptions that must be met in order to use the test:

1. The DV must be at interval or ratio level
2. Data is from a population that is normally distributed
3. There is homogeneity of variance (population variances are all equal)
4. For independent group designs, independent random samples must have been taken from each population

Between-Subjects or Repeated Measures?

Why use repeated measures?

- Between participants designs have weaknesses:
  - Individual differences
  - Other possible confounds (e.g., time of day)

- Repeated measures (within) designs reduce this problem
  - Individual differences are consistent across conditions

Considerations when using within-subjects designs

- As all participants see all conditions of IV, the conditions completed first may influence later ones
  - Potential ‘order’ effects needs consideration

- COUNTERBALANCING or randomising the order of conditions
- Some participants receive: Red, amber, green
- Others receive: Red, green, amber
- Others receive: Green, amber, red
- Others receive: Green, red, amber
- Others receive: Amber, red, green
- Others receive: Amber, green, red

ANOVA

- Human performance varies both between individuals and within the same individual

- If the variance between conditions is larger than the variance within conditions, then we can say that the IV is having a larger effect on scores than individual differences are

- So, ANOVA needs to calculate both types of variances.

ANOVA

- ANALYSIS OF VARIANCE
  - Compares the VARIANCE (variability in scores) **between** different groups (believed to be due to the IV) with the variability **within** each of the groups (believed to be due to chance)

ANOVA

$$F = \frac{\text{variance between groups}}{\textit{variance within groups}}$$

Larger F value = more variability between groups caused by the IV

- If there are differences between groups, the between group variance will be **larger** than the within group variance

- Calculated as the *F ratio*
  - The F-ratio looks at whether variances are due to our manipulation of the IV or from other nuisance factors (also known as the error variance).

Interpreting the ANVOA output

- Use your SPSS statistics book
  - This is the only way you will understand it for yourself!

- Tells you exactly what to interpret and how to report it

One-Way ANOVA: Main stats generated

F-value

- Compares directly to the t-value in t-test
- Represents amount of differences between ALL conditions

Degrees of freedom

- No. of values in the final calculation that are free to vary
- Always **TWO** values for *df* with ANOVA – relates sample size and number of conditions

Significance value

- Consulted to say whether effect/differences are significant
- Used same way as the t-test
- $p < .05$

Good practice in SPSS

- If you don’t need decimal places, reduce to 0 in the Variable view
- Always generate descriptive statistics
- In the ‘Options’ dialogue box, select to ask SPSS for:
  - ‘Estimates of effects size’ which tells you the proportion of the total variance accounted for by the factor (ANOVA).
  - ‘Homogeneity tests’ – you can proceed and assume homogeneity if Levene’s test is **not** significant

Reporting ANOVA statistics

- Formula is the same for **all** types of ANOVA

F (df1, df2) = F-value, *p* </> .05

Good practice to start reporting exact p-values – as long as confident with correct significant/non-significant interpretation

One-Way ANOVA practice on Moodle

- SPSS Panopto Recordings:
  - Between-subjects one-way ANOVA
  - Within-subjects one-way ANOVA

- SPSS Worksheets:
  - Between-subjects one-way ANOVA
  - Within-subjects one-way ANOVA