Psychology 205: Research Methods in Psychology
Within and between subject designs

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

NORTHWESTERN UNIVERSITY

April, 2021

Outline

- A research problem
- Theory testing
- Arousal study
  - Arousal and working Memory: II
- Counterbalancing
- Memory study
  - Additional demonstration
- Talking about statistics
- References

Arousal and working memory

1. An investigator was interested in the effect of arousal on short term memory. The hypothesis being tested was that caffeine induced arousal helps short term memory. Subjects were given a list of 20 words to study for 2 minutes, and were then asked to count backwards by 3s from 91. They were then asked to recall as many of the words as possible. The average number of words recalled was 10 (sd=3)
2. After the recall was completed, subjects were given 200 mg of caffeine and allowed to read for 30 minutes while the caffeine took effect. They were then given the same list to study for 2 minutes, followed by counting backwards again from 91. They were then asked to recall as many words as possible from the list. The average this time was now 12 (sd=3). There were 20 subjects in this within subject experiment and the t-test of the correlated differences was 3.6 (d.f. =19, p<.01).
3. The investigator concluded that the hypothesis that caffeine induced arousal helps working memory was supported.

Questions for evaluating research

1. What are the basic constructs being studied?
2. What are the particular operationalizations (observations) associated with the constructs?
3. How much of the variability in a construct is due to the (experimental manipulation) independent variable?
4. What are possible alternative sources of variation?

Arousal and working memory

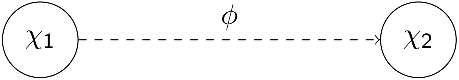
1. An investigator was interested in the effect of arousal on short term memory. The hypothesis being tested was that caffeine induced arousal helps short term memory. Subjects were given a list of 20 words to study for 2 minutes, and were then asked to count backwards by 3s from 91. They were then asked to recall as many of the words as possible. The average number of words recalled was 10 (sd=3)
2. After the recall was completed, subjects were given 200 mg of caffeine and allowed to read for 30 minutes while the caffeine took effect. They were then given the same list to study for 2 minutes, followed by counting backwards again from 91. They were then asked to recall as many words as possible from the list. The average this time was now 12 (sd=3). There were 20 subjects in this within subject experiment and the t-test of the correlated differences was 3.6 (d.f. =19, $p < .01$).
3. From these results, the investigator concluded that the hypothesis that caffeine induced arousal helps working memory was supported.

- Do these results follow?
- Can you think of an alternative explanation for the results?
- How would design a study to control for this alternative explanation?

Theory

Construct 1 Construct 2

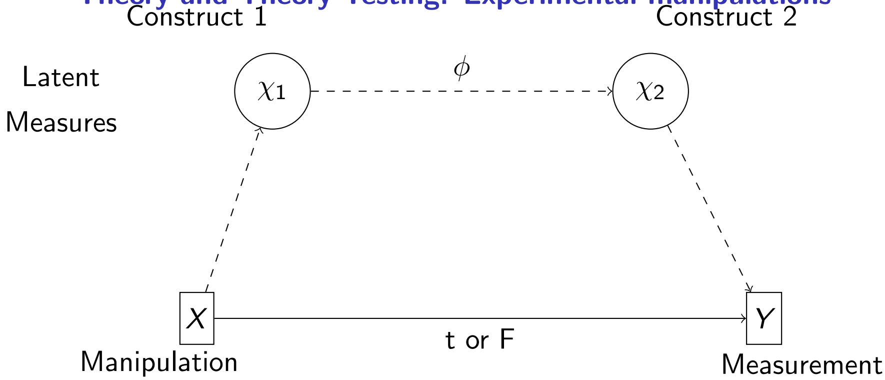
Latent



Measures

What are the constructs of interest and what is the strength of their relationship (ϕ)?

Theory and Theory Testing: Experimental manipulations



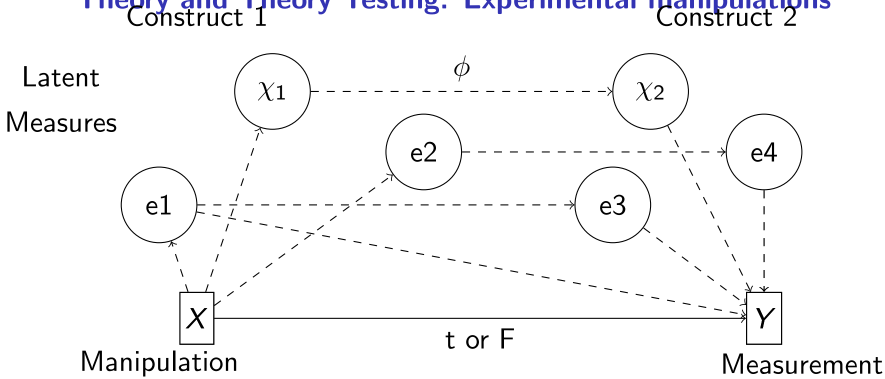
What are the constructs of interest and how can we manipulate or measure them?

We infer ϕ (the relationship of the constructs) from the t or F.

Independent and Dependent Variables

1. We refer to those variables we can manipulate as **Independent Variables**.
2. And those variables that depend upon our manipulations as **Dependent Variables**.
3. We manipulate the IVs (at least two levels) and then observe how the DVs vary as a function of the IV.
4. We infer that the IVs are causing the DVs, but we need to make sure that there are not other sources of variance that are causing the DVs.
5. We attempt to exclude these alternative explanations.
6. This is the essence of experimental design.

Theory and Theory Testing: Experimental manipulations

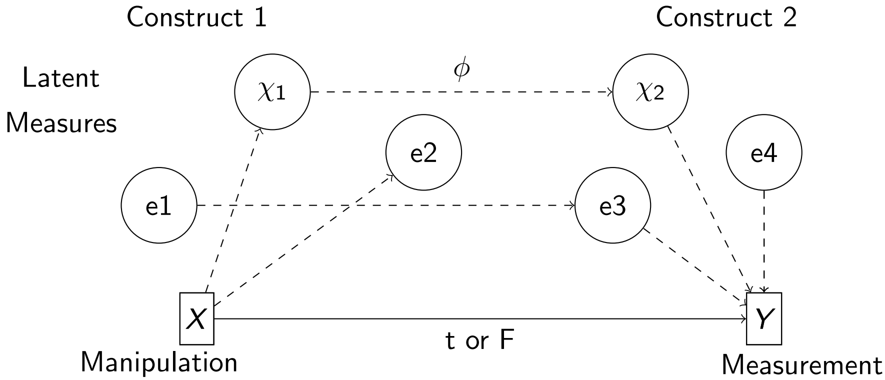


What are the constructs of interest and how can we manipulate or measure them?

We infer ϕ (the relationship of the constructs) from the t or F.

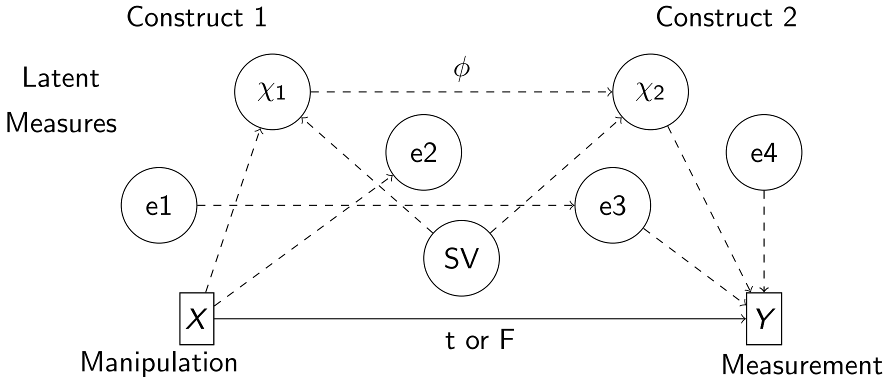
How can we eliminate alternative paths?

Theory and Theory Testing: Experimental manipulations



We infer ϕ (the relationship of the constructs) from the t or F.
We eliminate confounding paths by experimental design.

The challenge of subject variables



We infer ϕ (the relationship of the constructs) from the t or F.
We eliminate confounding paths by experimental design.

How to control for variability

1. Between subject variability
   - People differ because of
   - ability
   - motivation
   - practice
   - Need some way to control for these differences.
   - Particularly problematic if people differ on when they decide to participate (order effects)
2. Within subject variability
   - controls for order effects
   - Uses each person as their own control
   - Fatigue
   - Learning
   - Use counterbalancing across trials

Hypothetical Data with subject and practice effects

R code

```
#Make up some data for 12 subjects and 16 time points
data <- matrix(rep(1:12,16),ncol=16)
data[,2:16] <- data[,2:16] + col(data[,2:16])
colnames(data) <- paste0("T",1:ncol(data))
rownames(data) <- paste0("S",1:nrow(data))
data  #show them
rowMeans(data)  #do the subjects differ?
colMeans(data)  #is there change over trials
```

```
> data
    T1 T2 T3 T4 T5 T6 T7 T8 T9 T10 T11 T12 T13 T14 T15 T16
S1   1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16
S2   2  3  4  5  6  7  8  9 10  11  12  13  14  15  16  17
S3   3  4  5  6  7  8  9 10 11  12  13  14  15  16  17  18
S4   4  5  6  7  8  9 10 11 12  13  14  15  16  17  18  19
S5   5  6  7  8  9 10 11 12 13  14  15  16  17  18  19  20
S6   6  7  8  9 10 11 12 13 14  15  16  17  18  19  20  21
S7   7  8  9 10 11 12 13 14 15  16  17  18  19  20  21  22
S8   8  9 10 11 12 13 14 15 16  17  18  19  20  21  22  23
S9   9 10 11 12 13 14 15 16 17  18  19  20  21  22  23  24
S10 10 11 12 13 14 15 16 17 18  19  20  21  22  23  24  25
S11 11 12 13 14 15 16 17 18 19  20  21  22  23  24  25  26
S12 12 13 14 15 16 17 18 19 20  21  22  23  24  25  26  27

rowMeans(data)  #do the subjects differ?
  S1   S2   S3   S4   S5   S6   S7   S8   S9  S10  S11  S12
 8.5  9.5 10.5 11.5 12.5 13.5 14.5 15.5 16.5 17.5 18.5 19.5
> colMeans(data)  #is there change over trials
  T1   T2   T3   T4   T5   T6   T7   T8   T9  T10  T11  T12  T13  T14  T15  T16
 6.5  7.5  8.5  9.5 10.5 11.5 12.5 13.5 14.5 15.5 16.5 17.5 18.5 19.5 20.5 21.5
```

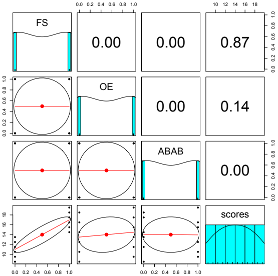
Alternative designs to control for subject or stimuli effects

1. First half versus second half – Very Bad
2. Odd versus Even – Better but still not perfect
3. ABBA counterbalanced - No linear order effects
4. Add these coding variables to our data to show subjects

R code

```r
FS = c(rep(0,6),rep(1,6))  #first second
OE <-  rep(0:1,6) %% 2  #use modular aritthmetic for odd even
ABAB <- c(0,1,1,0,1,0,0,1,1,0,0,1)  #ABBA BAAB BAAB
data.df <- data.frame(FS,OE, ABAB,scores=rowMeans(data))
data.df
```

```
data.df
    FS OE ABAB scores
S1   0  0    0    8.5
S2   0  1    1    9.5
S3   0  0    1   10.5
S4   0  1    0   11.5
S5   0  0    1   12.5
S6   0  1    0   13.5
S7   1  0    0   14.5
S8   1  1    1   15.5
S9   1  0    1   16.5
S10  1  1    0   17.5
S11  1  0    0   18.5
S12  1  1    1   19.5
```



- Note how FS and OE correlate with score because of confounding.
- ABBA counterbalancing does not.

3 ways of ordering stimuli – how to counter balance

R code

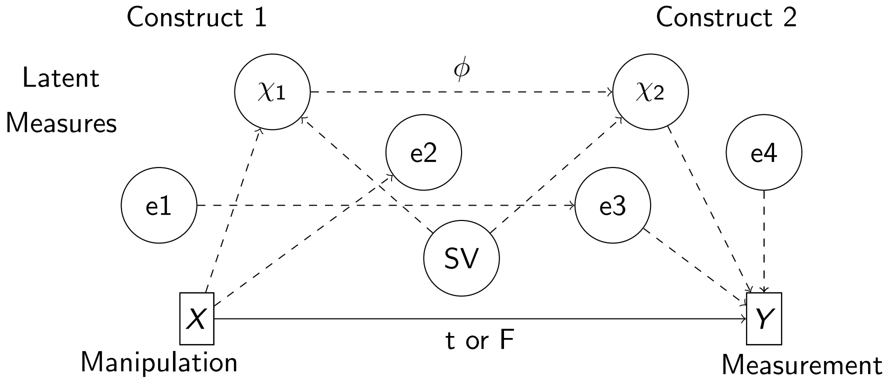
```
FS=c(rowMeans(data[,1:8]),rowMeans(data[,9:16]))
OE <- c(rowMeans(data[,1:15,2]),rowMeans(data[,2:16,2]))
ABBA <- c(rowMeans(data[,c(1,4,5,8,10,11,13,16)]),
      rowMeans(data[,c(2,3,6,7,9,12,14,15)]))
trial.df <-  data.frame(first=FS[1:12],second=FS[13:24],odd=OE[1:12],
      even=OE[13:24],A = ABBA[1:12],B=ABBA[13:24])
t(trial.df)  #Flip it on its side to see it

describe(trial.df)
```

```
t(trial.df)  #Flip it on its side to see it

         S1   S2   S3   S4   S5   S6   S7   S8   S9  S10  S11  S12
first   4.5  5.5  6.5  7.5  8.5  9.5 10.5 11.5 12.5 13.5 14.5 15.5
second 12.5 13.5 14.5 15.5 16.5 17.5 18.5 19.5 20.5 21.5 22.5 23.5
odd     8.0  9.0 10.0 11.0 12.0 13.0 14.0 15.0 16.0 17.0 18.0 19.0
even    9.0 10.0 11.0 12.0 13.0 14.0 15.0 16.0 17.0 18.0 19.0 20.0
A       8.5  9.5 10.5 11.5 12.5 13.5 14.5 15.5 16.5 17.5 18.5 19.5
B       8.5  9.5 10.5 11.5 12.5 13.5 14.5 15.5 16.5 17.5 18.5 19.5


     vars  n mean   sd median trimmed  mad  min  max range skew kurtosis   se
first     1 12 10.0 3.61   10.0    10.0 4.45  4.5 15.5    11    0     -1.5 1.04
second    2 12 18.0 3.61   18.0    18.0 4.45 12.5 23.5    11    0     -1.5 1.04
odd       3 12 13.5 3.61   13.5    13.5 4.45  8.0 19.0    11    0     -1.5 1.04
even      4 12 14.5 3.61   14.5    14.5 4.45  9.0 20.0    11    0     -1.5 1.04
A         5 12 14.0 3.61   14.0    14.0 4.45  8.5 19.5    11    0     -1.5 1.04
B         6 12 14.0 3.61   14.0    14.0 4.45  8.5 19.5    11    0     -1.5 1.04
```

Experimental Designs

1. Within Subjects
   - Controls for subject variability
   - Sensitive to within subject changes such as fatigue, learning, differential transfer
2. Between subjects
   - Controls for within subject changes
   - Sensitive to between subject variability
   - Effects due to subject selection, attrition, randomization
3. Mixed designs
   - Controls for materials effect (i.e., are some word lists easier to learn)

Experimental design (revisited)



We infer ϕ (the relationship of the constructs) from the t or F.
We eliminate confounding paths by experimental design.

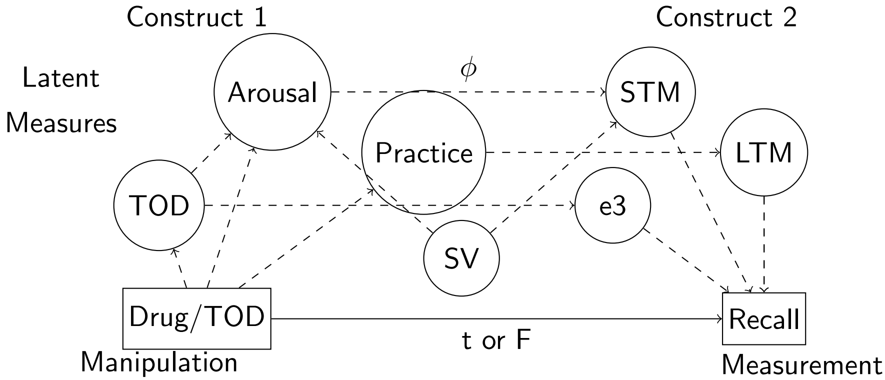
Analysis of any study

1. What are the constructs of interest?
2. How are they measured/manipulated?
3. What are possible alternative sources of variation?
    - Within subjects threats
    - Between subject threats
4. How strong is the relationship between the manipulation/observation of the IV and the measurement of the DV?

Arousal and Working Memory: revisited

1. Hypothesis – Alertness (arousal) facilitates short term memory
2. Constructs
   - Arousal
   - Short Term Memory (memory for very recent events)
3. Manipulations/Observables
   - Caffeine increases arousal
   - Study list - Filled Delay interval (why) – Immediate List recall
4. Alternative Explanations

Experimental design (revisited)



We infer ϕ (the relationship of the constructs) from the t or F.
We eliminate confounding paths by experimental design.

Arousal and working Memory: II

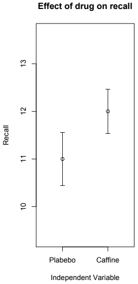
1. Another investigator was interested in the effect of caffeine induced arousal on short term memory. The hypothesis being tested was that caffeine induced arousal helps short term memory. To control for time of day effects, all subjects were run at 8 am. Subjects were given a list of 20 words to study for 2 minutes and were then asked to count backwards from 91 by 3s. They were then asked to recall as many of the word as possible. The average number of words recalled was 11 (sd=3.)
2. After the recall was completed, subjects were allowed to read quietly for an hour in order to minimize any possible carry over from the previous trial. Then the participants were given 200 mg. of caffeine and then allowed to read for 30 minutes while the caffeine took effect. They were then given a new list of words to study for 2 minutes, followed by counting forwards by 7s from 17. they were then asked to recall as many words as possible from the list. The average this time was now 12 (sd=2.5). With 30 subjects, this difference had a t- test of correlated differences of 2.8, df=29, $p<.01$.
3. From the results of this within subject study, the investigator concluded that the hypothesis that caffeine induced arousal helps working memory as supported.

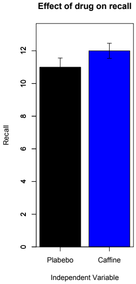
Data and graphics from Caffeine study

R code

```
study2.df <- data.frame(mean=c(11,12),sd=c(3,2.5),n=c(30,30))
 rownames(study2.df) <- c("Plabebo","Caffine")
 study2.df
 error.bars(stats=study2.df,eyes=FALSE,ylab="Recall",main="Effect of drug on re
 error.bars(stats=study2.df,eyes=FALSE,ylab="Recall",main="Effect of drug on re
```

```
        mean  sd  n
Plabebo   11 3.0 30
Caffine   12 2.5 30
```





Arousal and working Memory: II

1. Another investigator was interested in the effect of caffeine induced arousal on short term memory. The hypothesis being tested was that caffeine induced arousal helps short term memory. To control for time of day effects, all subjects were run at 8 am. Subjects were given a list of 20 words to study for 2 minutes and were then asked to count backwards from 91 by 3s. They were then asked to recall as many of the word as possible. The average number of words recalled was 11 (sd=3.)
2. After the recall was completed, subjects were allowed to read quietly for an hour in order to minimize any possible carry over from the previous trial. Then the participants were given 200 mg. of caffeine and then allowed to read for 30 minutes while the caffeine took effect. They were then given a new list of words to study for 2 minutes, followed by counting forwards by 7s from 17. they were then asked to recall as many words as possible from the list. The average this time was now 12 (sd=2.5). With 30 subjects, this difference had a t- test of correlated differences of 2.8, df=29, $p < .01$.
3. From the results of this within subject study, the investigator concluded that the hypothesis that caffeine induced arousal helps working memory as supported.
4. Do these results follow?
   - Can you think of an alternative explanation for the effects?
   - How would design a study to control for this alternative explanation?

Arousal and Working Memory: III

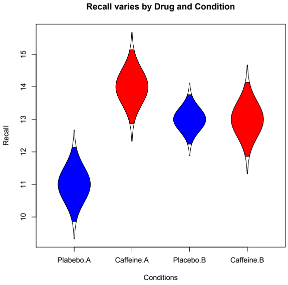
1. Yet another investigator was interested in the effect of caffeine induced arousal on short term memory The hypothesis being tested was that caffeine induced arousal helps short term memory. To control for time of day effects, all subjects were run at 8 am.
2. However, to control for possible order effects, 1/2 of the participants were run in one within subject condition, the other half in the other condition.
3. That is, half were given a list of 20 words to study for 2 minutes and were then asked to count backwards from 91 by 3s. They were then asked to recall as many of the word as possible. The average number of words recalled for this group was 11 (sd=3.) Then the participants were given 200 mg. of caffeine and then allowed to read for 30 minutes while the caffeine took effect. They were then given a new list of words to study for 2 minutes, followed by counting forwards by 7s from 17. they were then asked to recall as many words as possible from the list. The average this time was now 14 (sd=2.5). With 30 subjects, this difference had a t-test of correlated differences of 2.8, df=29, $p<.01$.
4. The other half of the participants were given the caffeine on trial one and not given anything on trial 2. Their performance on trial 1 was 13 (sd=2) and on trial 2 was 12.8 (sd=2). This difference was not reliably different from a chance difference (t=.4 ns.)
5. Although the one order showed the effect and the other did not, the investigator then pooled the data from the two orders and found that the caffeine condition in general led to better performance. (mean caffeine = 13, mean control = 11.9). From these results the investigator concluded that the hypothesis that caffeine induced arousal helps working memory as supported.

Showing these results

R code

```
study2a.df <- data.frame(drug=c("Placebo","Caffeine","Placebo","Caffeine")
,order=c(1,1,2,2),mean=c(11,14,13,13),sd=c(3,3,2,3),n=c(30,30,30,30))
rownames(study2a.df) <- cs(Plabebo.A,Caffeine.A,Placebo.B, Caffeine.B)
error.bars(stats=study2a.df,col=c("blue","red","blue","red"),
   main="Recall varies by Drug and Condition",
     ylab="Recall",xlab="Conditions")
```

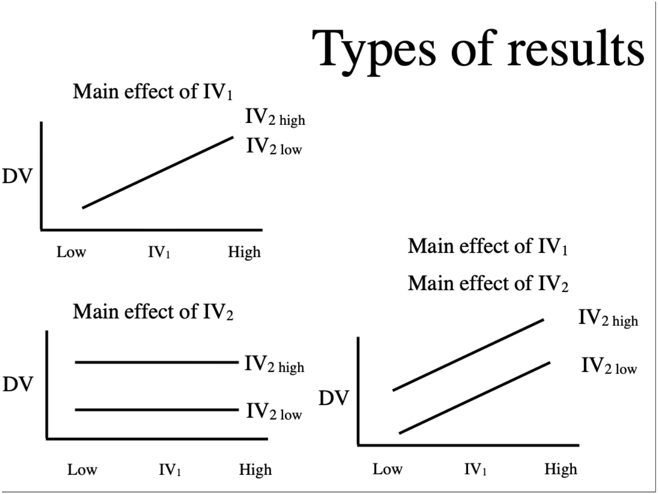
```
               drug order mean sd  n
Plabebo.A   Placebo     1   11  3 30
Caffeine.A Caffeine     1   14  3 30
Placebo.B   Placebo     2   13  2 30
Caffeine.B Caffeine     2   13  3 30
```



Two variables – Three answers

1. When we study two variables at the same time, we can ask three different questions:
   - Is there an effect of Variable 1?
   - Is there an effect of Variable 2?
   - Does the effect of Variable 1 depend upon Variable 2 (do they interact)?
2. Typically discussed in terms of analysis of variances, but can also be done in terms of regressions
3. The question is do the slopes differ from 0 and from each other?

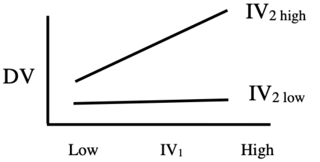
Types of results



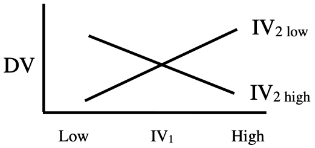
Types of interactions

Types of interactions

Effect of IV_2 depends upon that of IV_1



Fan Fold



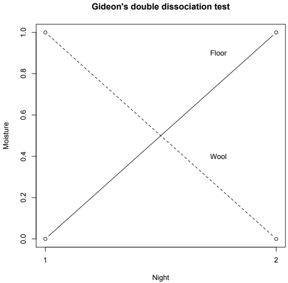
Cross over

Inferential power of an interaction

1. Main effect of a variable shows that there is a relationship between IV and DV.
2. Interaction of two IVs with DV means that the effect of one IV depends upon the other IV.
3. By having an interaction, we are able to specify the limits of our effects.
4. Interactions allow more powerful inference, for they can exclude more alternative hypotheses

Earliest known example of a within subject study with a cross over interaction (double dissociation)

1. Gideon was an early methodologist who understood principles of good design (Judges 6:36-40)
2. Day 1: Make the wool wet, keep the floor dry – alternative explanations for effect
3. Day 2: reverse conditions: keep the wool dry, make the floor wet
4. By having a reversal, it is harder to explain effect



Experimental Designs

1. Within Subjects – Every subject is own control
    - Every subject is a complete experiment
    - Controls for subject variability
        - Ability
        - Motivation
2. Sensitive to within subject changes
    - Fatigue
    - Learning
3. Counterbalancing controls for some transient effects but is open to threats of
    - Differential transfer

How to study several within subject variables at the same time

1. Counterbalancing to avoid confounding – IV1 and IV2 are experimentally independent
2. Conditions crossed with conditions
3. All conditions for IV1 occur with all conditions of IV2 – no systematic relationship between IV1 and IV2
4. Conditions balanced across orders of presentation

Purpose of counterbalancing

1. Conditions are independent of order and of each other
2. This allows us to determine effect of each variable independently of the other variables.
3. If conditions are related to order or to each other, we are unable to determine which variable is having an effect

Results: a selective summary

1. Do not need to report every analysis, just the ones that tell the important story.
2. Think about how to aggregate the data to best summarize it
3. Transforms of data to make more understandable – e.g., percent correct rather than raw number
4. Story must be truth
5. Don't hide "inconvenient data"
6. Assume someone else will want to analyze your data

Data= Model + Residual

1. The process of science is improve the model and reduce the error
2. Models are progressively more complicated
3. Consider the recall data:
   - Model 0: Data
   - Model 1: Data = Mean + Residual
   - Model 2: Data = Position_i + Residual
   - Model 3: Data = Type of presentation + Residual – ...

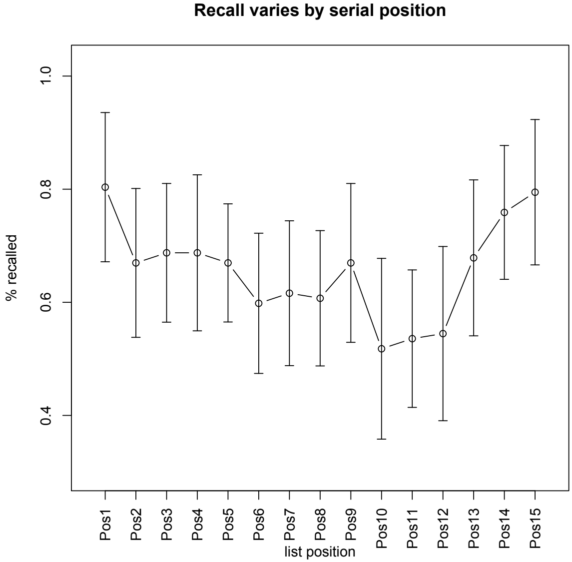
Procedure: overview

1. The memory study had 2 parts.
2. It was inspired and partly based upon the work of Roediger & McDermott (1995)
3. Shortened to make it suitable for a class experiment.
4. Eight lists of 15 words were presented (Lists 1-8 from Roediger & McDermott (1995))
5. Following each list, two minutes were given for recall
6. Recall of the 15 words for each of 8 lists
7. After all lists had been presented, a recognition task was given for a subset of the words.
    - If our data match prior verbal learning data, there should be a serial position of the probability of recall.
    - First and last words from the list should be more frequently recalled.
8. Recognition of the words presented (and not presented)
    - Recognition of 4 words per list (3 presented, one not presented)
    - Recognition of 16 words (8 cued by the lists, 8 not cued) (We reported this part on Monday)

Results

1. Recall (manipulation check)
   - Is there a serial position effect?
   - Primacy
   - Recency (particularly given the instructions)
   - But this was confounded with association strength
2. Recognition (Measured two different ways)
   - Is there a false memory effect?
   - What manipulations affect it?
   - Are these the same manipulations that affect real recognition?

Recall varies by serial position



Why do we examine the serial position effects?

1. We are seeing if we can produce false memories. But if we do, maybe we are not doing a standard memory task.
2. That is, if we show that people have false memories, some people might say that this is a weird task.
3. But the recall was just as we would have expected.
4. The quadradtic trend of the lists was tested and recall varied by the square of postion (t=4.25, df=207 $p < .001$)

```
 setCor(Recall ~ Position + I(Position^2),data=pos.df)
Call: setCor(y = Recall ~ Position + I(Position^2), data = pos.df)

Multiple Regression from raw data

 DV =  Recall
            slope   se     t       p lower.ci upper.ci VIF
(Intercept)  0.00 0.07  0.00 1.0e+00    -0.13     0.13   1
Position    -0.04 0.07 -0.64 5.2e-01    -0.17     0.09   1
Position^2   0.28 0.07  4.25 3.3e-05     0.15     0.41   1

Residual Standard Error =  0.96  with  207  degrees of freedom

 Multiple Regression
         R   R2  Ruw R2uw Shrunken R2 SE of R2 overall F df1 df2        p
Recall 0.29 0.08 0.23 0.05        0.07     0.04      9.22   2 207 0.000146
```

How did I do this?

R code

```
pos.list <-list()
for(i in 1:n.items ) {pos.list[[i]] <- rowSums(recall[items + i-1])}
pos.strn <- unlist(pos.list)
pos.df <- data.frame(Recall=pos.strn,Subject=rep(1:14,15), Position=rep(1:15,ea
setCor(Recall ~ Position + I(Position^2),data=pos.df) #this zero centers
```

```
setCor(Recall ~ Position + I(Position^2),data=pos.df)
Call: setCor(y = Recall ~ Position + I(Position^2), data = pos.df)

Multiple Regression from raw data

 DV =  Recall
            slope   se     t       p lower.ci upper.ci VIF
(Intercept)  0.00 0.07  0.00 1.0e+00    -0.13     0.13   1
Position    -0.04 0.07 -0.64 5.2e-01    -0.17     0.09   1
Position^2   0.28 0.07  4.25 3.3e-05     0.15     0.41   1

Residual Standard Error =  0.96  with  207  degrees of freedom

 Multiple Regression
          R   R2  Ruw R2uw Shrunken R2 SE of R2 overall F df1 df2        p
Recall 0.29 0.08 0.23 0.05        0.07     0.04      9.22   2 207 0.000146
```

The recognition data

1. The recognition words were 5 types:
   1.1 Words that were cued by the lists but not presented
   1.2 The first word from each presented list
   1.3 The 8th word from each list
   1.4 The 10th word from list
   1.5 Control words
2. If there were no false memories, we would expect words that were not present to not be recognized
3. We would expect greater recognition for the first than the 8th and 10th words.

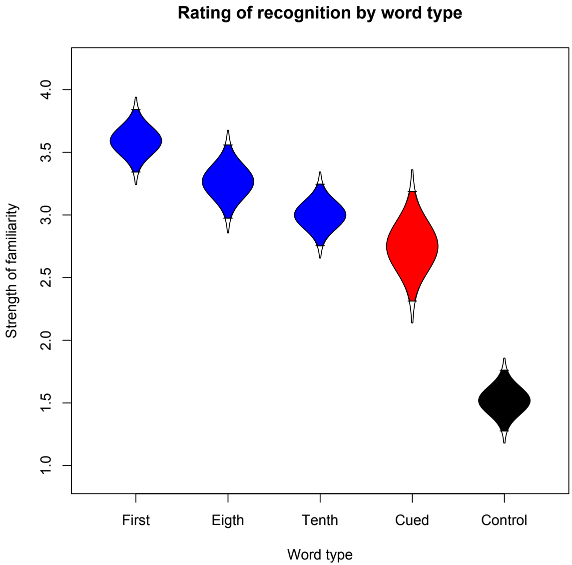
Basic recognition statistics

R code

```
describe(data.df)  #tabular form
#graphic form
error.bars(recog.df,col=c("blue","blue","blue","red","black"),ylab=
"Strength of familiarity",xlab="Word type",
main="Rating of recognition by word type")
```

```
        vars  n mean   sd median trimmed  mad  min  max range  skew kurtosis   se
First      1 15 3.59 0.45   3.88    3.63 0.19 2.62 4.00  1.38 -0.81    -0.90 0.12
Eigth      2 15 3.27 0.53   3.50    3.30 0.37 2.25 3.88  1.62 -0.76    -0.79 0.14
Tenth      3 15 3.00 0.44   3.12    3.06 0.37 1.75 3.50  1.75 -1.33     1.62 0.11
Cued       4 15 2.75 0.79   2.88    2.78 0.93 1.25 3.88  2.62 -0.42    -1.15 0.20
Control    5 15 1.52 0.44   1.41    1.49 0.51 1.03 2.38  1.34  0.53    -1.21 0.11
```

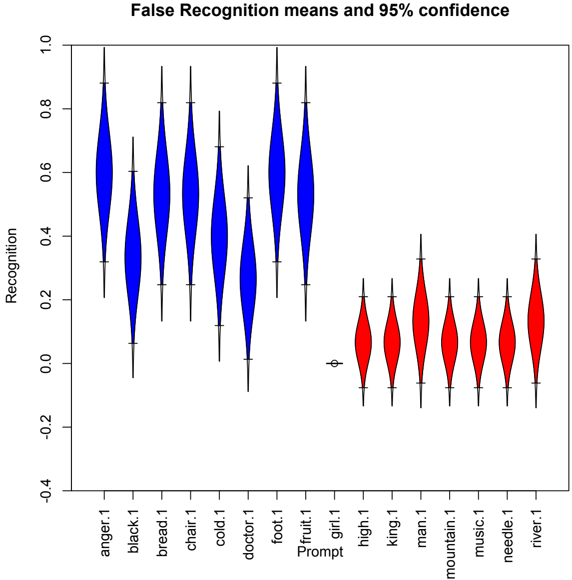
Recognition rating by word type



An additional demonstration of false memory

1. Following the 64 item recognition task, a subsequent task asked for True/False recognition of the 8 Cued words and 8 Control words that were high associated of other (non-presented) lists.
2. These responses were the last 16 in the data file.
3. The recognition data file were read in from the remote file and saved as the `recog` object.

The False Recognition results and their confidence using error.bars



Descriptive statistics of the results

R code

```
describe(data.df)
```

```
        vars  n mean   sd median trimmed  mad min  max range  skew kurtosis   se
group      1 15 1.60 0.51   2.00    1.62 0.00   1 2.00  1.00 -0.37    -1.98 0.13
Cued       2 15 0.48 0.38   0.62    0.47 0.56   0 1.00  1.00 -0.02    -1.71 0.10
Control    3 15 0.07 0.17   0.00    0.04 0.00   0 0.62  0.62  2.31     4.58 0.04
CuedA      4 15 0.50 0.38   0.50    0.50 0.37   0 1.00  1.00 -0.12    -1.52 0.10
CuedB      5 15 0.47 0.40   0.50    0.46 0.74   0 1.00  1.00  0.20    -1.62 0.10
```

Cued versus uncued primes were recalled more

Using R to find the t value: two ways

R code

```
t.test(Cued, Control, data= data.df,paired=TRUE)
```

```
        Paired t-test

data:  Cued and Control
t = 4.2982, df = 14, p-value = 0.0007362
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 0.2003998 0.5996002
sample estimates:
mean of the differences
                    0.4
t2d(4.297,14) = 2.29 (Cohen's d)
```

Descriptive and Inferential Statistics

1. Describe the data
   - Central Tendencies and Dispersion
   - Means, standard deviations
2. Inferential – the Null Hypothesis model
3. How likely are the data given a model of no difference
4. consider the t-test

Roediger, H. L. & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814.