2-Way Completely Randomized Design

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

2-Way Completely Randomized Design

Introduction

In this module, we introduce the two-factor completely randomized analysis of variance design.

We introduce the ANOVA concepts of *main effects*, *simple main effects*, and *interactions*.

We discuss the measurement of effect size in the context of the design.

An Introductory Example

Maxwell and Delaney (2003) present the following introductory example of the effect of biofeedback and drug on blood pressure. There were 4 groups in the study, in a classic 2×2 design. The 4 groups were of the 4 possible combinations of Drug-No Drug, and Biofeedback-No Biofeedback.

Since there are 4 groups, it is possible to analyze the data as a 1-Way ANOVA.

Loading in the data file, we can see that there is a *Group* factor, that describes the combination of conditions and has 4 levels. There are also two other factor grouping variables, each of which records the "Present-Absent" status of *Drug* and *Biofeedback*.

An Introductory Example

```
> bp <- read.table("bp.txt", header = T, sep = ",")
> bp
                 Group Blood.Pressure Biofeedback
1  biofeedback and drug             158     Present
2  biofeedback and drug             163     Present
3  biofeedback and drug             173     Present
4  biofeedback and drug             178     Present
5  biofeedback and drug             168     Present
6     biofeedback alone             188     Present
7     biofeedback alone             183     Present
8     biofeedback alone             198     Present
9     biofeedback alone             178     Present
10    biofeedback alone             193     Present
11           drug alone             186      Absent
12           drug alone             191      Absent
13           drug alone             196      Absent
14           drug alone             181      Absent
15           drug alone             176      Absent
16              neither             185      Absent
17              neither             190      Absent
18              neither             195      Absent
19              neither             200      Absent
20              neither             180      Absent
      Drug
1  Present
2  Present
3  Present
4  Present
5  Present
6   Absent
7   Absent
8   Absent
9   Absent
10  Absent
11 Present
12 Present
13 Present
14 Present
15 Present
16  Absent
17  Absent
18  Absent
19  Absent
20  Absent
```

An Introductory Example

We can analyze these data with the tools we have developed so far, as follows.

1. First, we can test whether there are *any* differences between means with an overall ANOVA F test.
2. Next, we can employ our generalized t statistic to test several specific hypotheses of interest which we have already encountered.

An Introductory Example

First we run the ANOVA

```
> attach(bp)
> fit.1 <- aov(Blood.Pressure ~ factor(Group))
> summary(fit.1)

             Df Sum Sq Mean Sq F value  Pr(>F)
factor(Group)  3   1540   513.3   8.213 0.00155 **
Residuals     16   1000    62.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We see that there is definitely a significant difference between the groups, in the sense that we can reject the null hypothesis that all group means are equal.

An Introductory Example

Interaction Plot

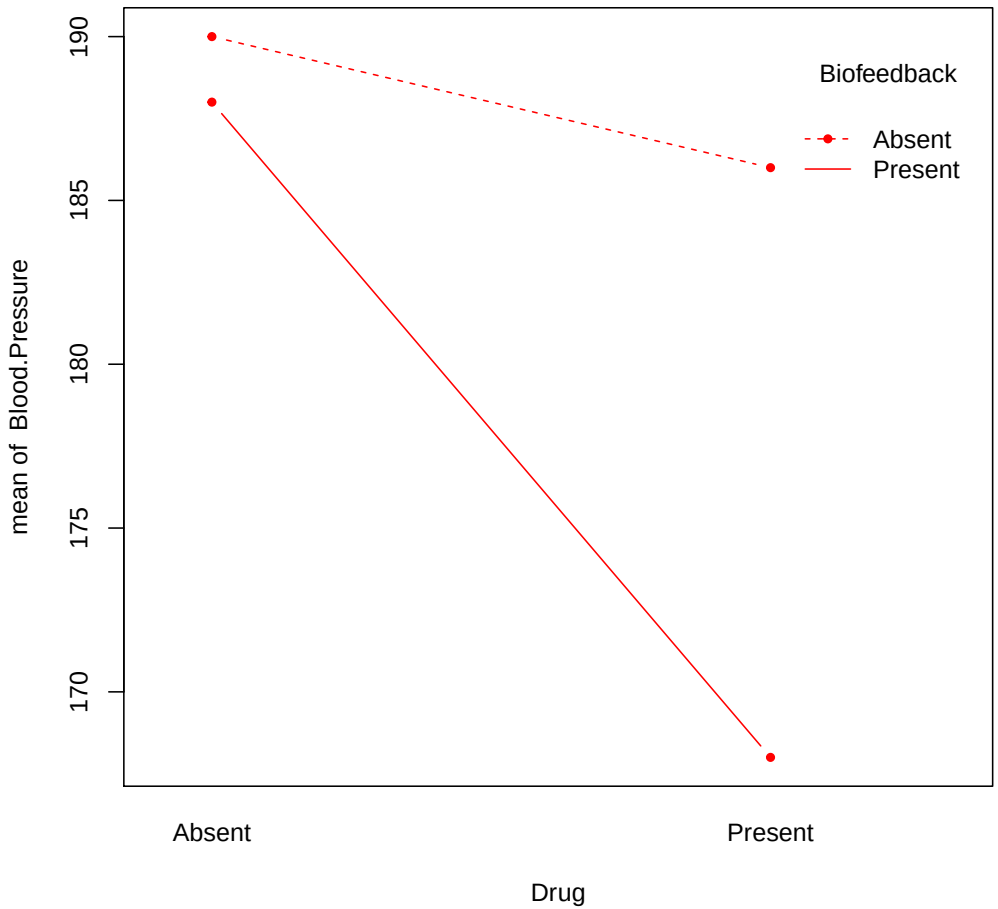
Next, I am going to jump into something that normally is presented somewhat later, i.e., an *interaction plot* of the cell means for the 4 cells of the design.

They say a picture can be worth a thousand words, and that is essentially true here.

An Introductory Example

Interaction Plot

```
> interaction.plot(Drug, Biofeedback, Blood.Pressure, type = "b",
+     col = "red", pch = 20)
```



An Introductory Example

Simple Main Effects

Let's go at the analysis in a somewhat non-standard order.

We'll discuss *simple main effects* first.

The question of whether there is a simple main effect for *Drug* at a particular *level* of *Biofeedback* addresses the question, "Within a particular level of *Biofeedback*, does *Drug* have any effect on *Blood.Pressure*?

Let's ignore sampling variability for the moment, and imagine that the cell means plotted on the graph are actually the true population means.

Then we shall assess whether there is *a simple main effect for Drug at level "Present" of Biofeedback*.

An Introductory Example

Simple Main Effects

We look only at the line representing when *Biofeedback* is present. It is the more steeply sloped line.

Then, we examine the values of *Blood.Pressure* at the two levels (Present,Absent) of the drug.

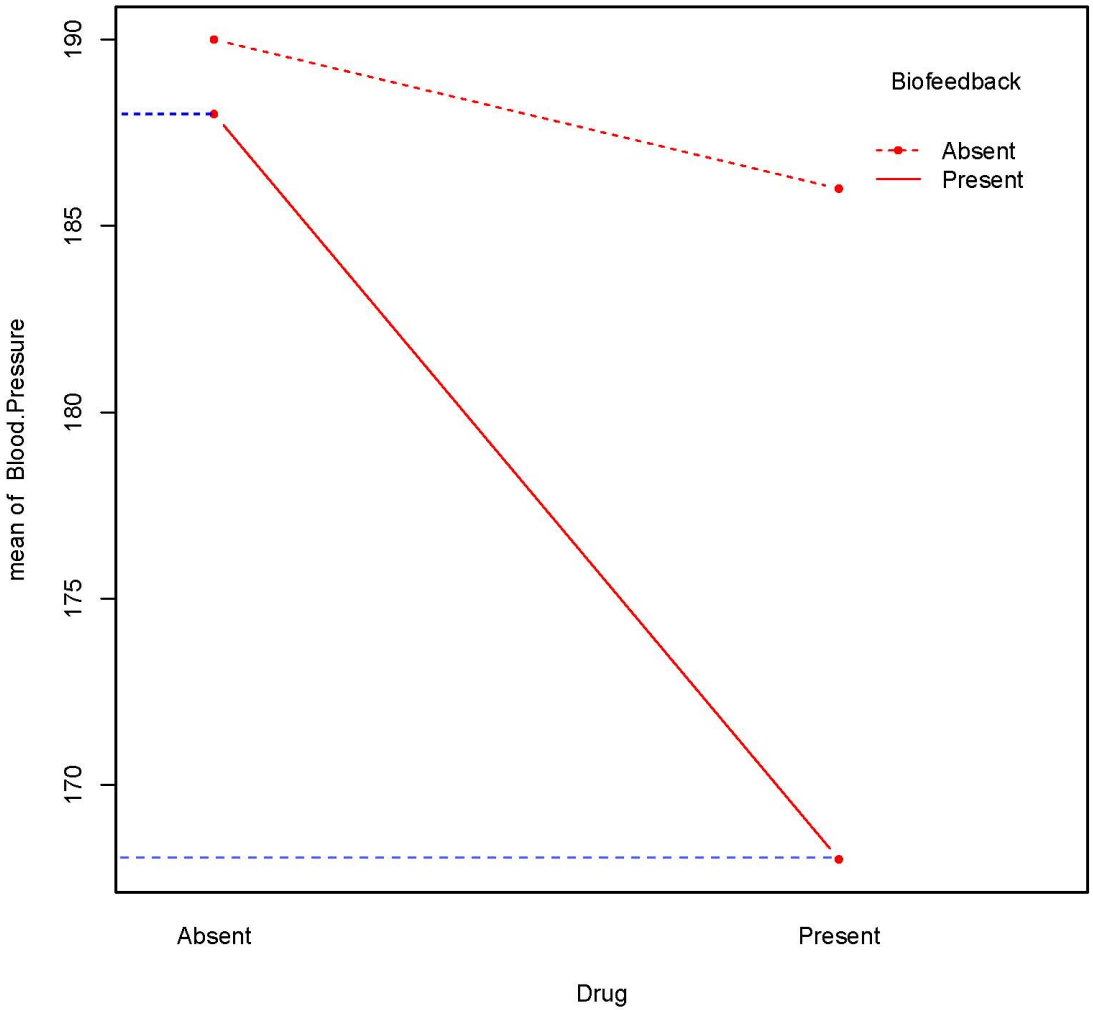
This is facilitated in the next graph by drawing horizontal blue dotted lines to the vertical axis from the points on line marked "Biofeedback Present."

You can see that the presence of the *Drug* is coincidental with a 20 point reduction in *Blood.Pressure*, as the two points are at 168 and 188.

If these were population means, we would say that the simple main effect of *Drug* on *Blood.Pressure* when *Biofeedback* is present is -20.

An Introductory Example

Simple Main Effects



An Introductory Example

Simple Main Effects

Notice that, if we knew population means, assessment of whether there is a simple main effect conditional on the variable that changes across lines would be easy.

We simply examine the plot and see if the line is flat.

If it is, there is a simple main effect which is easy to characterize if there are only two levels.

Unfortunately, our task is significantly complicated by the presence of sampling error.

An Introductory Example

Simple Main Effects

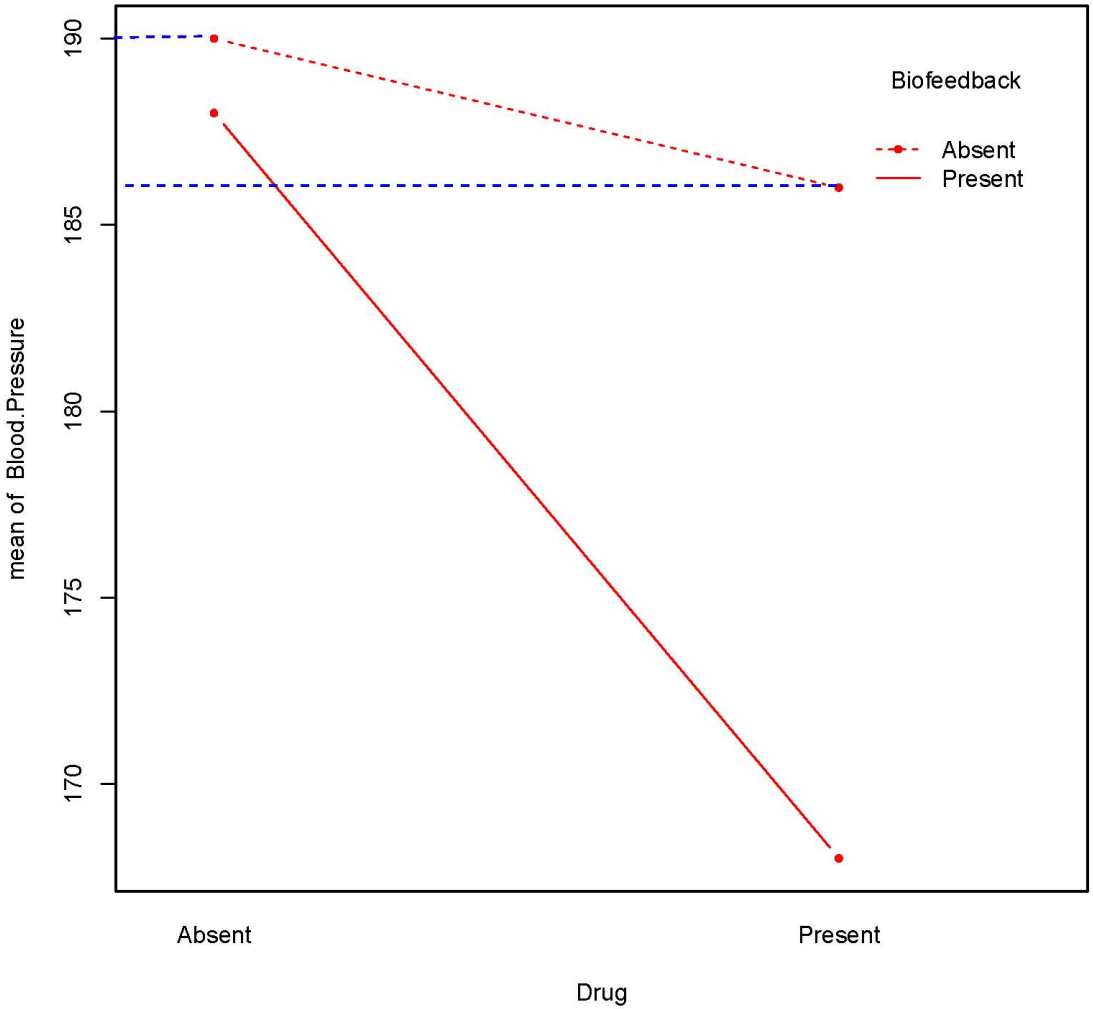
To analyze whether there is a simple main effect of *Drug* at level "Absent" of *Biofeedback*, we look only at the line representing when *Biofeedback* is absent. It is the less steeply sloped line.

You can see from the dotted lines that the presence of the *Drug* is coincidental with a 4 point reduction in *Blood.Pressure*, as the two points are at 186 and 190.

So the simple main effect in this case is -4

An Introductory Example

Simple Main Effects



An Introductory Example

Interaction Effects

When simple main effects of factor A given B are different for different levels of B, we say there is an *interaction* between A and B.

This means that A behaves differently, depending on the value of B

Needless to say, this can be very important in practice.

Two quickly assess whether there are interaction effects from an interaction plot *of population means*, we simply examine the plot to see if the lines are all parallel. If they are not all parallel, then there is an interaction effect.

An Introductory Example

More Simple Main Effects

To analyze whether there is a simple main effect of *Biofeedback* at level "Present" of *Drug*, we look only at the *values* representing when *Drug* is Present.

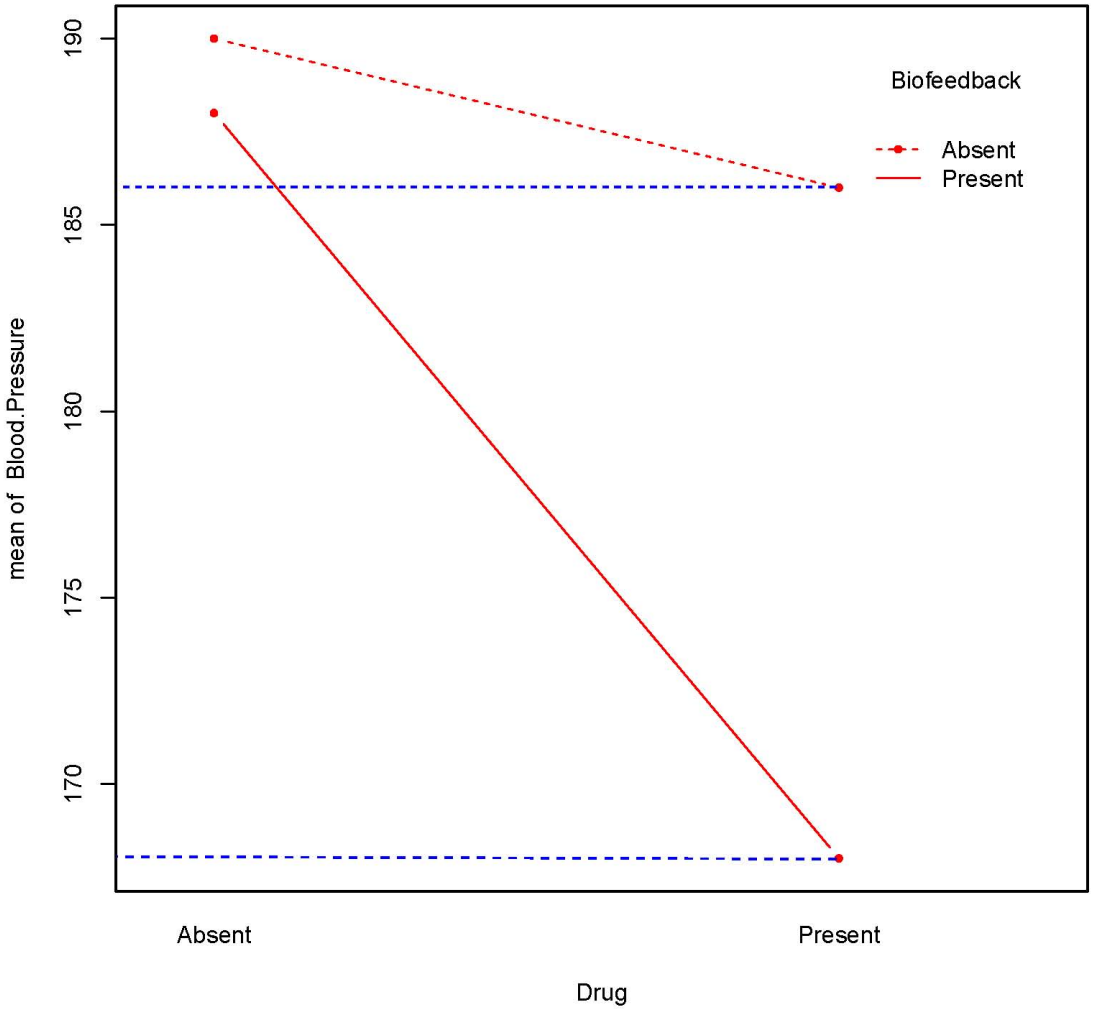
We draw lines from them over to the vertical axis.

The difference between the lines is the simple main effect.

We see that the simple main effect of *Biofeedback* when *Drug* is present is -18, that is, an 18 point reduction in blood pressure.

An Introductory Example

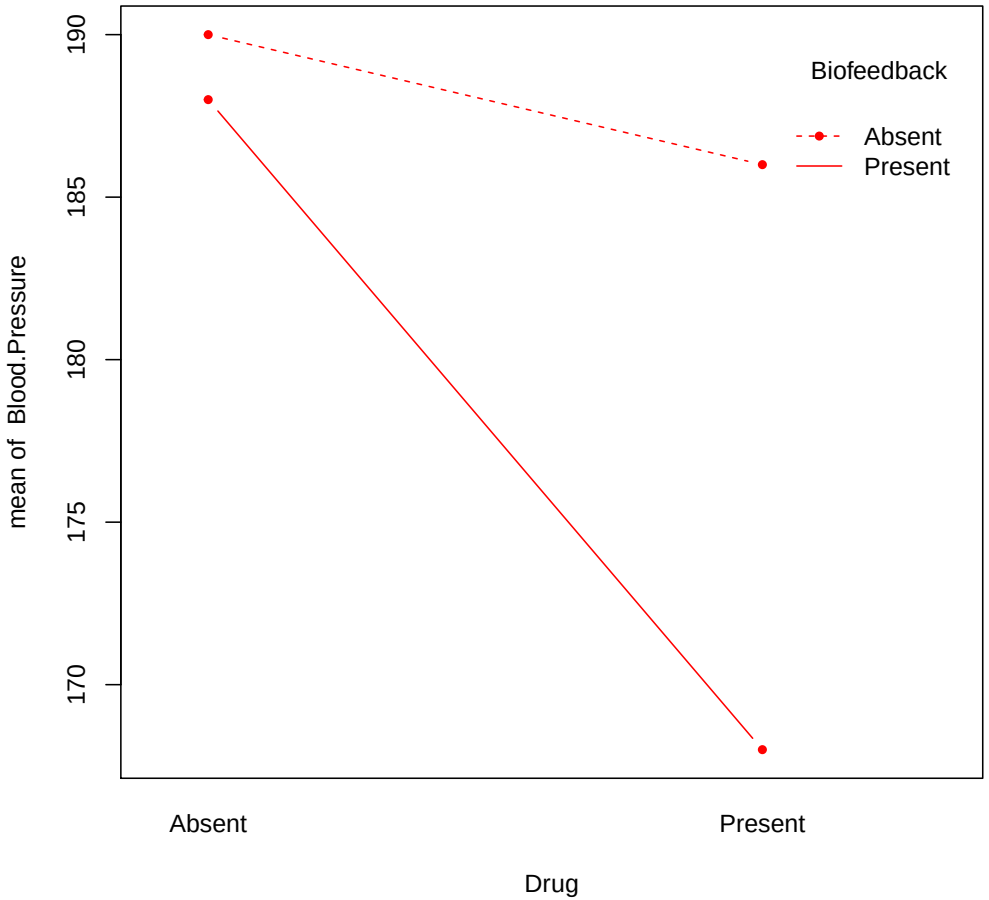
More Simple Main Effects



An Introductory Example

More Simple Main Effects

Now you try it. What is the simple main effect of *Biofeedback* when *Drug* is "Absent"?



An Introductory Example

Testing for Simple Main Effects

From a plot of cell means, it is easy to see if there are simple main effects.

For the factor represented on the x-axis, the following rules apply:

1. We go to each line in the plot and ask, "Is is flat."
2. If the line is flat there is no simple main effect at the level of the second factor represented by that line.
3. If the line is not flat, then there is a simple main effect at the level of the second factor represented by that line.

There is a second factor to consider, which varies across lines in the plot. For this factor, we do the following:

1. Go to each level j represented on the x-axis, and draw a vertical line.
2. If all points on the plot that fall on that line are coincident, then there is no simple main effect for the second factor at level j of the first factor.
3. If all points on the plot that fall on that line are not coincident, then there is a simple main effect for the second factor at level j of the first factor.
4. Equivalently, if the identified points do not all coincide, there is a simple main effect.

An Introductory Example

Main Effects

A *main effect* in factorial ANOVA refers to an observed difference between (or differences among, with more than two levels) means that are calculated across all levels of the other factor(s).

Referring to our interaction plot, and again imagining that the points were population means devoid of sampling error, the question becomes, "Is level of *Drug* related to a difference in *Blood.Pressure* when we collapse across levels of *Biofeedback*?"

Let's see how we evaluate that question graphically.

An Introductory Example

Main Effects

First we go to the point where the *Drug* is absent, and we average across the levels of *Biofeedback*. This gives us the average blood pressure.

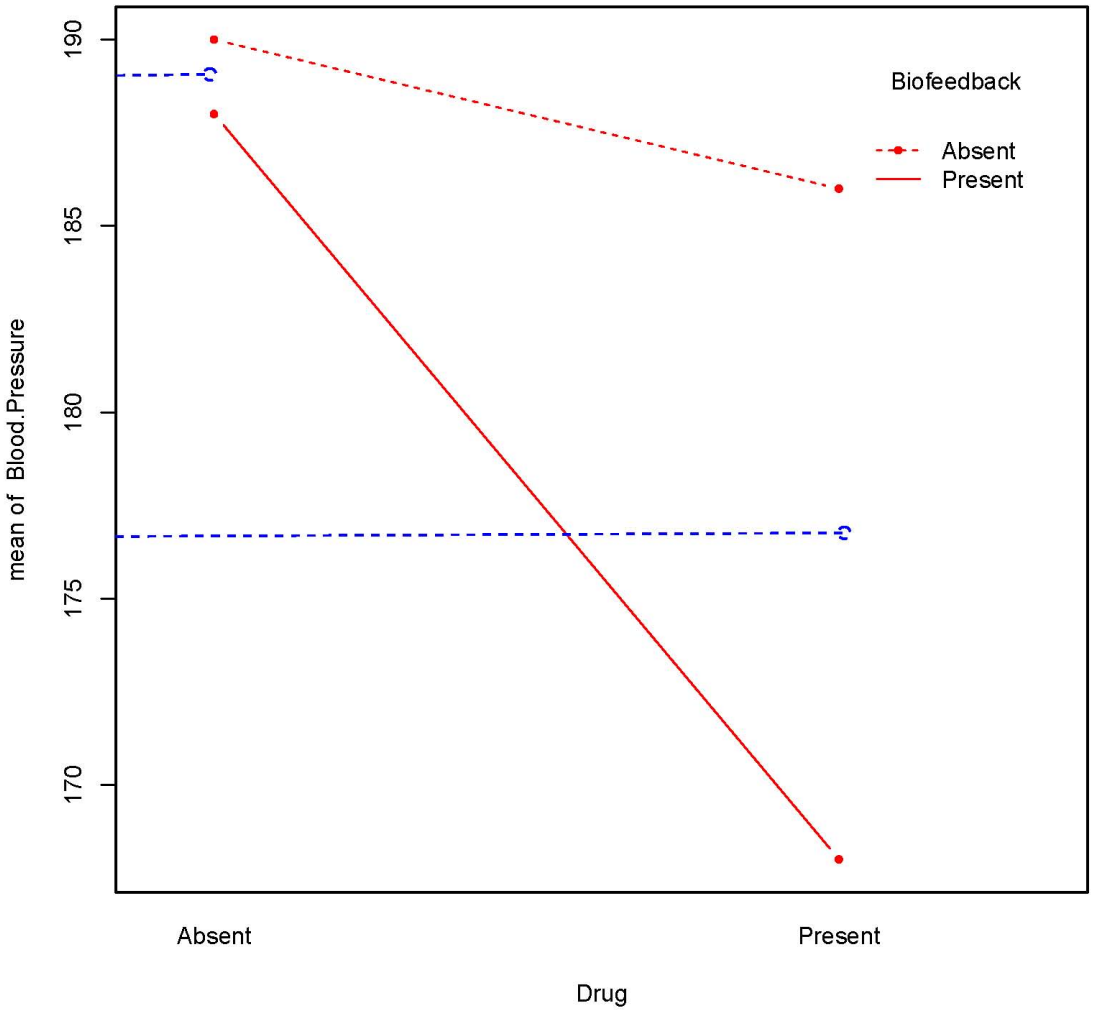
In the case of just two levels of *Biofeedback*, this amounts to plotting a point halfway between the two values. I've done that on the graph.

Next, we do the same thing for the case where the *Drug* is present.

If these two points were at the same level on the vertical axis, then there would be no main effect for *Drug*.

An Introductory Example

Main Effects



An Introductory Example

Main Effects

Next, we ask what turns out to be a somewhat more difficult question to visualize.

Is there a main effect for *Biofeedback*?

Here, we have to "average a line value" by averaging across the points on each line in the interaction plot.

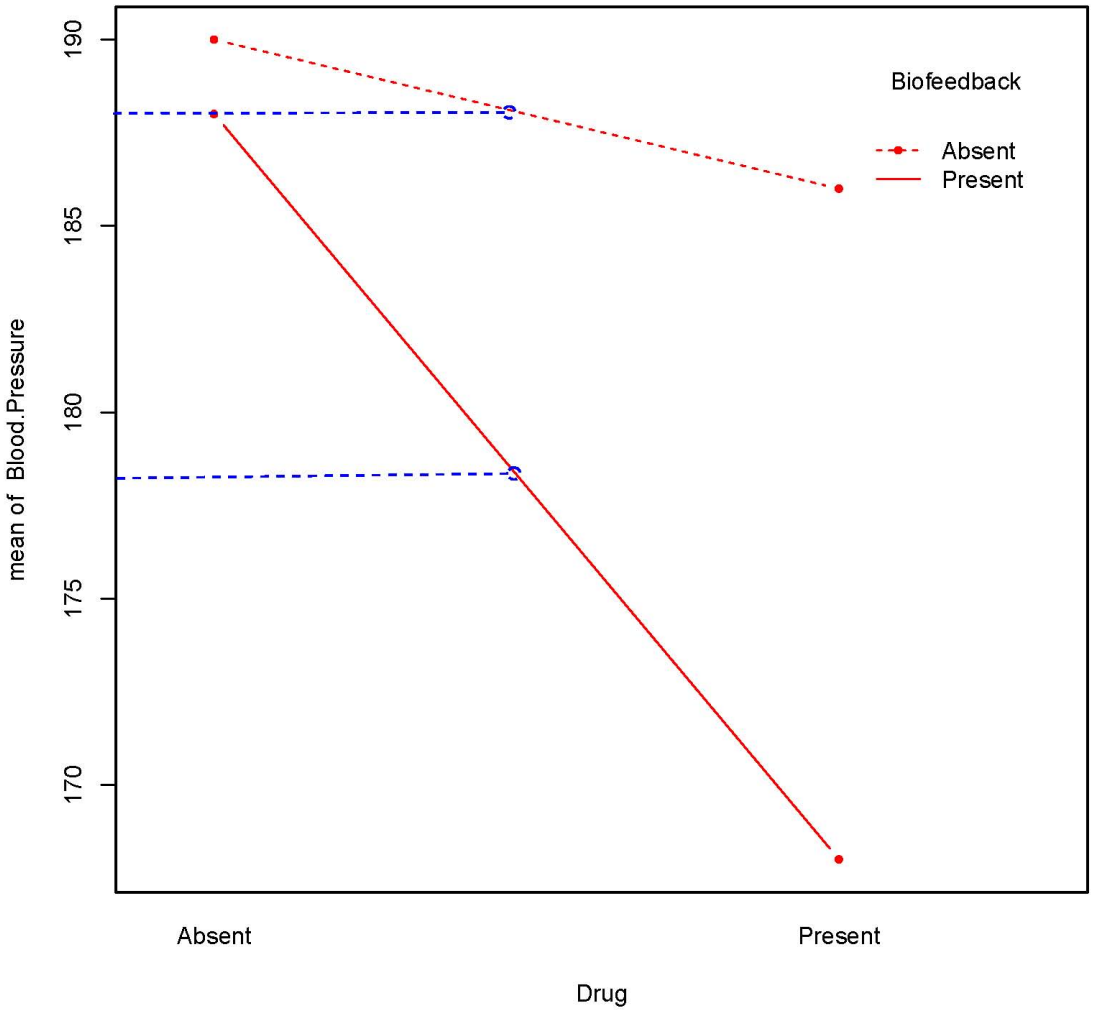
We draw a horizontal line from these points to the vertical axis.

This is (relatively) easy with two levels of a factor, but somewhat more difficult with more than two levels.

If the points on the vertical axis do not all coincide, then there is a main effect for *Biofeedback*, otherwise there is not.

An Introductory Example

Main Effects



Basic ANOVA Analysis

The Linear Model ANOVA Parameterization

We can fit the ANOVA model as a linear model with *Drug* and *Biofeedback* as *factors*.

The complete model includes main effect terms and interaction terms.

Note how the interaction term is indicated: *Drug* : *Biofeedback*.

Both main effects and the interaction are significant.

However, since the interaction is significant, main effects are not necessarily of much interest.

Basic ANOVA Analysis

The Linear Model ANOVA Parameterization

```
> anova(lm(Blood.Pressure ~ Drug + Biofeedback + Drug:Biofeedback,
+     data = bp))

Analysis of Variance Table

Response: Blood.Pressure
                Df Sum Sq Mean Sq F value   Pr(>F)
Drug             1    720   720.0   11.52 0.003706 **
Biofeedback      1    500   500.0    8.00 0.012109 *
Drug:Biofeedback 1    320   320.0    5.12 0.037917 *
Residuals       16   1000    62.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Basic ANOVA Analysis

The Linear Model ANOVA Parameterization

An alternative way of entering a model that includes all possible interactions among its factors to enter the design in complete factorial notation as $Drug * Biofeedback$. If the model is entered this way, it is not necessary to enter the interaction term explicitly.

Basic ANOVA Analysis

The Linear Model ANOVA Parameterization

```
> anova(lm(Blood.Pressure ~ Drug * Biofeedback, data = bp))

Analysis of Variance Table

Response: Blood.Pressure
                 Df Sum Sq Mean Sq F value   Pr(>F)
Drug              1    720   720.0   11.52 0.003706 **
Biofeedback       1    500   500.0    8.00 0.012109 *
Drug:Biofeedback  1    320   320.0    5.12 0.037917 *
Residuals        16   1000    62.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Simple Main Effects Calculation

To perform the analysis without using the error term from the complete two-way analysis, we simply set up a 1-Way ANOVA on the data for which *Biofeedback* == *Present*.

```
> anova(lm(Blood.Pressure ~ Drug, data = subset(bp, Biofeedback ==
+     "Present")))

Analysis of Variance Table

Response: Blood.Pressure
          Df Sum Sq Mean Sq F value  Pr(>F)
Drug       1   1000  1000.0      16 0.00395 **
Residuals  8    500    62.5
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Simple Main Effects Calculation

To perform the analysis using the error term from the 2-Way Analysis, we calculate by hand:

```
> my.F <- 1000/62.5
> my.F

[1] 16

> p.value <- 1 - pf(my.F, 1, 16)
> p.value

[1] 0.001032025
```

In this case, the error mean square (62.5) was identical for the full and reducted data sets, so the F statistic remained the same.

This usually will not happen. However, note that the degrees of freedom for the denominator increased from 8 to 16, resulting in a lower p-value.

The General Linear Model

Cell Means Parameterization

In the linear model for the 2-way design, there are a levels of factor A, b levels of factor B, and, as a result, ab populations Ω_{jk} with means μ_{jk} representing the combination of level A_j with level B_k. The n observations in any particular cell represent a random sample from Ω_{jk}

On pages 201–203 of RDASA3, MWL work through the general linear model for the 2-way factorial design.

The *cell means parameterization* is simply

$$Y_{ijk} = \mu_{jk} + \varepsilon_{ijk} \tag{1}$$

where the ε_{ijk} have the distribution

$$\varepsilon_{ijk} \underset{i.i.d}{\sim} N(0, \sigma_e^2) \tag{2}$$

The General Linear Model

Linear Structural Model Parameterization

The *linear structural model parameterization* for *ab* populations is

$$Y_{ijk} = \mu + \alpha_j + \beta_k + (\alpha\beta)_{jk} + \varepsilon_{ijk} \tag{3}$$

where the ε_{ijk} have the distribution

$$\varepsilon_{ijk} \underset{i.i.d}{\sim} N(0, \sigma_e^2) \tag{4}$$

and

$$\begin{aligned}
\mu &= \sum_j \sum_k \mu_{jk}/ab && (5)\\
\mu_{j\bullet} &= \sum_k \mu_{jk}/b && (6)\\
\mu_{\bullet k} &= \sum_j \mu_{jk}/a && (7)\\
\alpha_j &= \mu_{j\bullet} - \mu && (8)\\
\beta_k &= \mu_{\bullet k} - \mu && (9)\\
(\alpha\beta)_{jk} &= \mu_{jk} - (\mu + \alpha_j + \beta_k) && (10)
\end{aligned}$$

Calculating Population Model Coefficients

It is instructive to imagine that we actually knew population means for each cell, and to calculate the parameters of the linear structural model from those means.

Calculating Population Model Coefficients

Table 9.2 Treatment population means and effects

(a) Original population means

	B_1	B_2	B_3	B_4	μ_j	$\alpha_j = \mu_j - \mu$
A_1	65	50	47	58	55	5
A_2	43	48	51	38	45	−5
μ_k	54	49	49	48	$\mu = 50$	
$\beta_k = \mu_k - \mu$	4	−1	−1	−2		

(b) Population means after removing the A (α_j) and B (β_k) main effects

	B_1	B_2	B_3	B_4	Mean
A_1	56	46	43	55	50
A_2	44	54	57	45	50
Mean	50	50	50	50	$\mu = 50$

(c) Interaction effects; $(\alpha\beta)_{jk} = (\mu_{jk} - \mu) - \alpha_j - \beta_k$

	B_1	B_2	B_3	B_4	Mean
A_1	6	−4	−7	5	0
A_2	−6	4	7	−5	0
Mean	0	0	0	0	

Calculating Population Model Coefficients

Table 9.3 Treatment population means with no interaction effects

	B_1	B_2	B_3	B_4	μ_j	$\alpha_j = \mu_j - \mu$
A_1	59	54	54	53	55	5
A_2	49	44	44	43	45	−5
μ_k	54	49	49	48	$\mu = 50$	
$\beta_k = \mu_k - \mu$	4	−1	−1	−2		

(*b*) Population means after removing the A (α_j) and B (β_k) main effects

	B_1	B_2	B_3	B_4	Mean
A_1	50	50	50	50	50
A_2	50	50	50	50	50
Mean	50	50	50	50	$\mu = 50$

Calculating Population Model Coefficients

Table 9.4 Simple effects of A at each level of B ($\mu_{1k} - \mu_{2k}$) for the data of Tables 9.2 and 9.3

(a) Population means from Table 9.2 with interaction effects

	B_1	B_2	B_3	B_4	μ_j
A_1	65	50	47	58	55
A_2	43	48	51	38	45
$\mu_{1k} - \mu_{2k}$	22	2	–4	20	10

(b) Population means from Table 9.3 with no interaction effects

	B_1	B_2	B_3	B_4	μ_j
A_1	59	54	53	53	55
A_2	49	44	43	43	45
$\mu_{1k} - \mu_{2k}$	10	10	10	10	10

Source Table and Expected Mean Squares

Table 9.5 The analysis of variance (ANOVA) table for the two-factor between-subjects design (*a*) and expected mean squares (*b*)

(*a*) ANOVA

Source	*df*	*SS*	*MS*	*F*
Total	$abn - 1$	$\sum_{j}^{a}\sum_{k}^{b}\sum_{i}^{n}(Y_{ijk} - \overline{Y}_{...})^2$		
Between cells	$ab - 1$	$n\sum_{j}^{a}\sum_{k}^{b}(\overline{Y}_{.jk} - \overline{Y}_{...})^2$		
A	$a - 1$	$nb\sum_{j}^{a}(\overline{Y}_{.j.} - \overline{Y}_{...})^2$	SS_A/df_A	$MS_A/MS_{S/AB}$
B	$b - 1$	$na\sum_{k}^{b}(\overline{Y}_{..k} - \overline{Y}_{...})^2$	SS_B/df_B	$MS_B/MS_{S/AB}$
AB	$(a - 1)(b - 1)$	$SS_{cells} - SS_A - SS_B$	SS_{AB}/df_{AB}	$MS_{AB}/MS_{S/AB}$
S/AB	$ab(n - 1)$	$SS_{total} - SS_{cells}$	$SS_{S/AB}/df_{S/AB}$	

(*b*) Expected mean squares

SV	*EMS*
A	$\sigma_e^2 + nb\sum_j \alpha_j^2/(a-1) = \sigma_e^2 + nb\theta_A^2$
B	$\sigma_e^2 + na\sum_k \beta_k^2/(b-1) = \sigma_e^2 + na\theta_B^2$
AB	$\sigma_e^2 + n\sum_j\sum_k (\alpha\beta)_{jk}^2/(a-1)(b-1) = \sigma_e^2 + n\theta_{AB}^2$
S/AB	σ_e^2

Note: $\alpha_j = \mu_{j.} - \mu$, $\beta_k = \mu_{.k} - \mu$, and $(\alpha\beta)_{jk} = (\mu_{jk} - \mu) - \alpha_j - \beta_k = (\mu_{jk} - \mu_{j.} - \mu_{.k} + \mu)$. The θ^2 notation serves as a reminder that $\Sigma\alpha_j^2/(a-1)$ is not a variance; the variance of the treatment population means has *a* as the denominator.

The E(MS) Test Construction Principle

Traditional ANOVA texts place considerable emphasis on the derivation of expressions for the expected values of Mean Squares in various designs.

There are a couple of reasons why this material is important.

A primary reason is that F tests in many ANOVA designs are the ratio of two mean squares.

Looking at the expected mean squares can tell you *which* mean squares to put in the numerator and denominator to test the null hypothesis that an effect is zero.

The E(MS) Test Construction Principle

For example, suppose you wish to test the A main effect in a 2-way ANOVA.

The expected mean square for A is

$$(MS_A) = \sigma_e^2 + nb\theta_A^2 \tag{11}$$

This mean square will go in the numerator, but which value goes in the denominator?

Ask, "What will (MS_A) be if the null hypothesis is true? If the null hypothesis is true, $\theta_A^2 = 0$, and so $(MS_A) = \sigma_e^2$.

Can you find another mean square that has that expected value regardless of whether the null hypothesis is true?

If so, that mean square will go in the denominator.

In this case, consulting the table of expected mean squares, we see that $MS_{S|AB}$ satisfies that requirements, so the F statistic for testing the A main effect is $F = MS_A/MS_{S|AB}$.

Calculations: The Wiley-Voss Example

Introduction

Wiley and Voss (1999) in a *Journal of Educational Psychology* article, examined the effect of (a) various learning strategies *and* (b) the method of presentation of subject content on the learning of historical subject matter.

In this example, we use a subset of data, in which the dependent variable is performance on an *Inference Verification Task (IVT)* as a function of instructions given the subjects about how to study.

In the inference verification task, students were asked whether a statement about subject matter was true on the basis of the information they read.

Calculations: The Wiley-Voss Example

Instruction

There were 4 types of instructions. All students were presented with a writing task. The task stated: Historians work from sources including newspaper articles, autobiographies and government documents like census reports to create histories. Your task is to take the role of historian and develop a(n) _____ about what produced the significant changes in Ireland's population between 1846 and 1850. Depending on condition, the blank space was replaced with one of the following 4 words.

1. *Narrative (N).*
2. *Summary (S).*
3. *Explanation (E).*
4. *Argument (A).*

Calculations: The Wiley-Voss Example

Format of Presentation

All subjects received historical information about Ireland from 1800 to 1850, including a map, biographical accounts of King George III and Daniel O'Connell, brief descriptions of the Act of Union, the Act of Emancipation, and the Great Famine, census population data, and economic statistics between 1800 and 1850.

Calculations: The Wiley-Voss Example

Format of Presentation

There were 2 types of content presentations:

1. *Web*. This was an internet browser-like environment with eight separate source documents. The environment used in this experiment was *Sourcer's Apprentice*, a program developed by Perfetti and colleagues (described in Rouet, Britt, Mason, & Perfetti, 1996) as an aid for history classrooms. Each document is represented as a book, with a title along the spine. The books are placed on a bookshelf on the main page of the program. Readers could open up to two source documents or books at the same time. There were no hypertext links between documents in this study, and readers could return to the documents whenever they wished during the writing task. The average length of each document was around 220 words.
2. *Text*. presented as a textbook-like chapter (1,571 words). The information presented was identical in the two formats with the exception that the textbook format contained an introductory sentence and some (noncausal) transitional clauses at the beginning of paragraphs.

Calculations: The Wiley-Voss Example

Main Effects and Interactions

To analyze the data, we begin by loading in the file. To make the data more readable, we add "value labels" to the file.

This is always a good idea with qualitative independent variables, because the procedure demonstrated automatically results in the variables being "typed" as factors.

```
> ## Read in data
>
> wiley.voss <- read.csv("wiley0906.csv",header=T,sep=",")
>
> ## Establish Value Labels for Instruction
>
> wiley.voss$Instruction <- factor(wiley.voss$Instruction,
+ levels = c(1,2,3,4),
+ labels = c("Narrative", "Summary", "Explanation","Argument"))
>
> ## Establish Value Labels for Format
>
> wiley.voss$Format <- factor(wiley.voss$Format,
+ levels = c(1,2),
+ labels = c("Text","Web"))
```

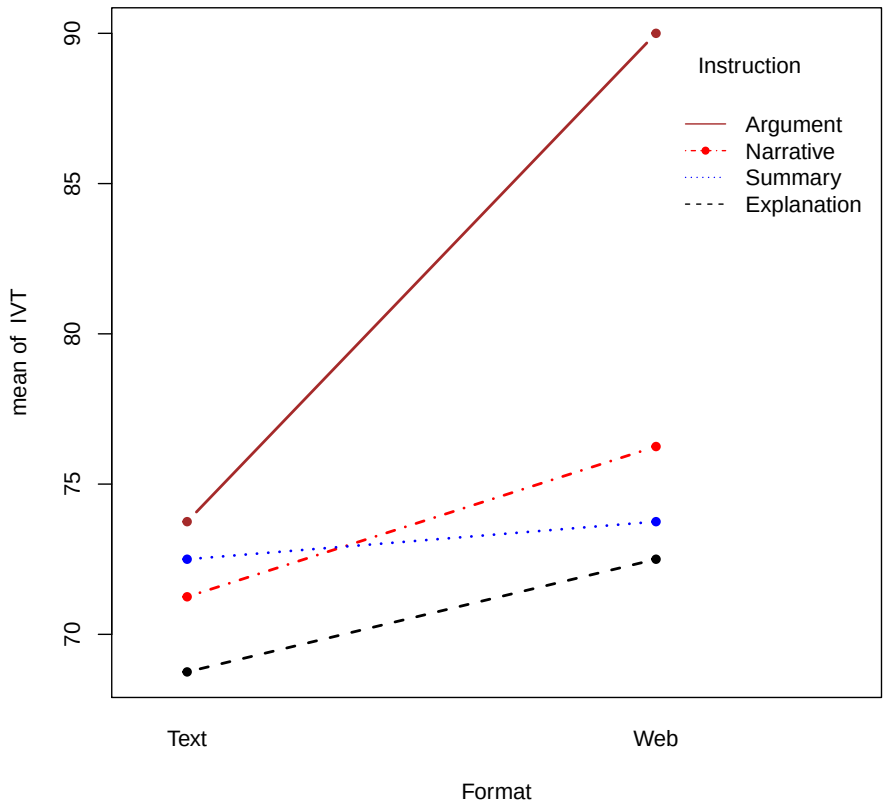
Calculations: The Wiley-Voss Example

Main Effects and Interactions

Here is an interaction plot.

Taking into account sampling error, how do you interpret this plot? (C.P.)

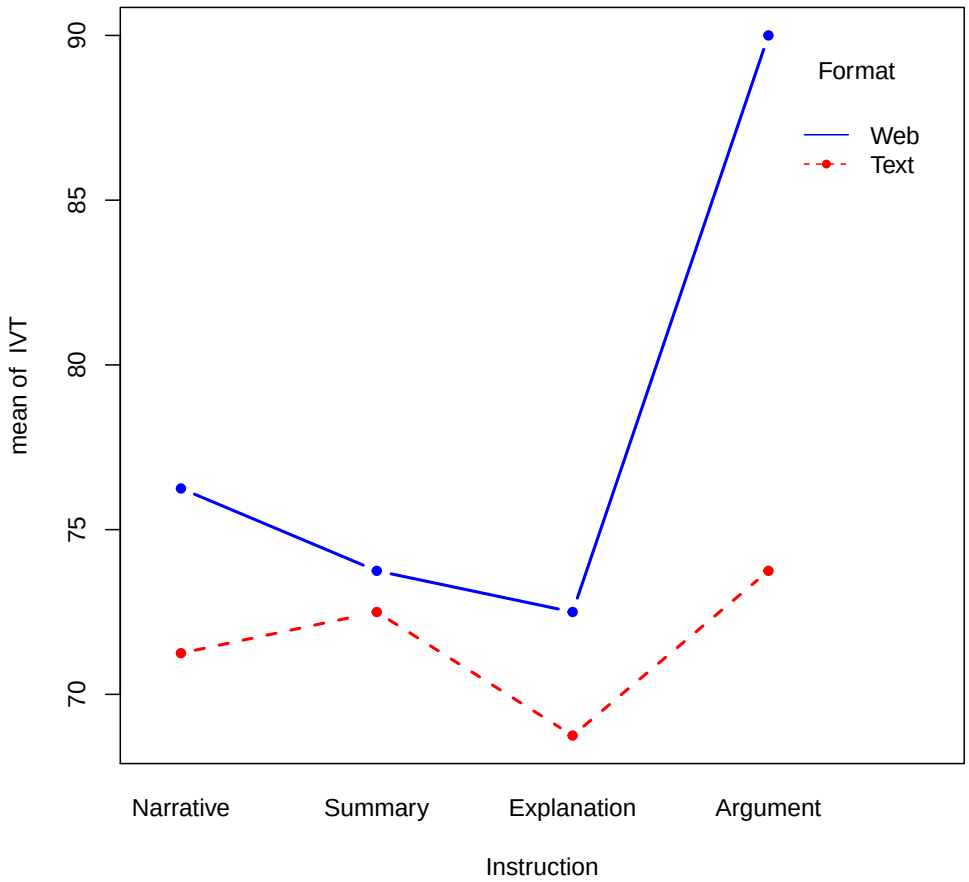
```
> with(wiley.voss, interaction.plot(Format, Instruction, IVT,
+     type = "b", col = c("red", "blue", "black", "brown"),
+     pch = 20, lwd = 2))
```



Calculations: The Wiley-Voss Example

Main Effects and Interactions

```
> with(wiley.voss, interaction.plot(Instruction, Format, IVT,
+     type = "b", col = c("red", "blue", "black", "brown"),
+     pch = 20, lwd = 2))
```



Calculations: The Wiley-Voss Example

Main Effects and Interactions

Here is a bar plot.

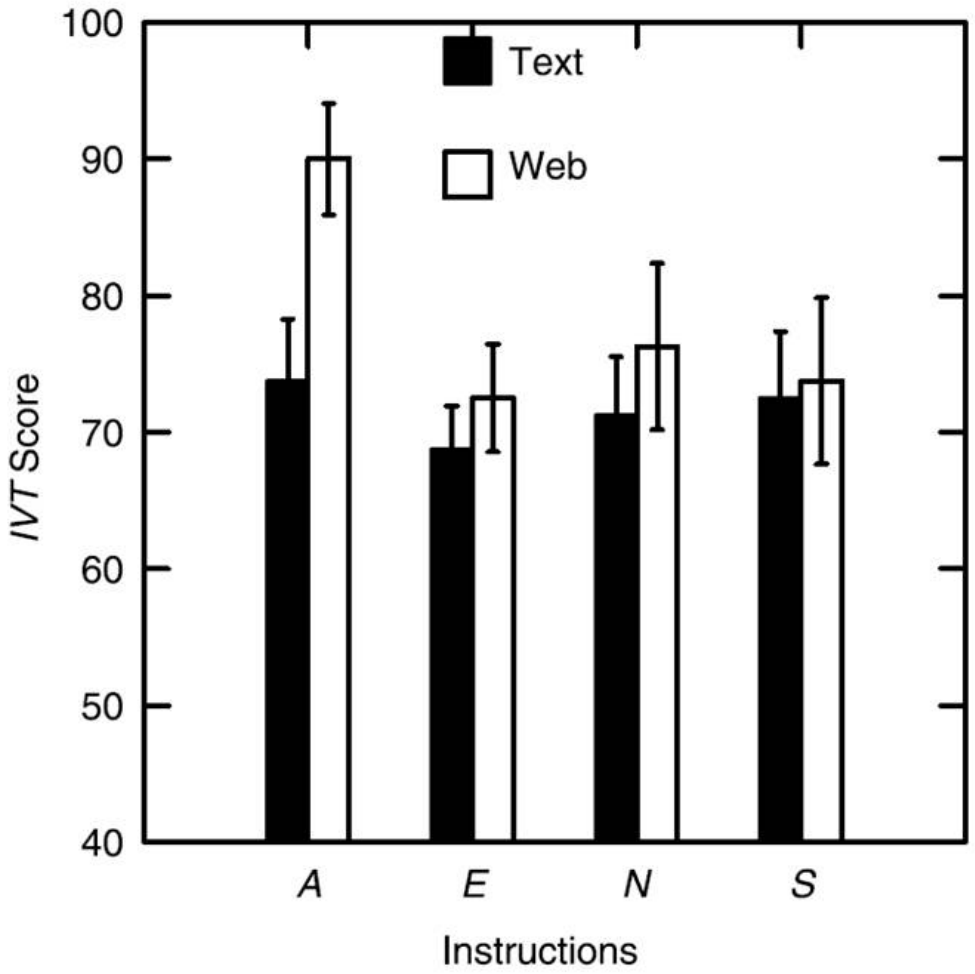


Fig. 9.2 Bar graph of the Wiley–Voss (1999) *IVT* data.

Calculations: The Wiley-Voss Example

Main Effects and Interactions

```
> anova(lm(IVT ~ Instruction * Format, data = wiley.voss))

Analysis of Variance Table

Response: IVT
                   Df Sum Sq Mean Sq F value  Pr(>F)
Instruction         3 1142.2  380.73  2.4684 0.07139 .
Format              1  689.1  689.06  4.4674 0.03901 *
Instruction:Format  3  529.7  176.56  1.1447 0.33906
Residuals          56 8637.5  154.24
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Effect Size Estimation

The RMSSE

Steiger (2004) discusses several measures of effect size and their relationship to each other and the noncentrality parameter λ of the F statistic used for testing a particular efffect θ. The Root Mean Square Standardized Effect (RMSSE) is

$$\Psi_\theta = \sqrt{\frac{\sum(\theta/\sigma)^2}{df_\theta}} \tag{12}$$

$$= \sqrt{\frac{\lambda_\theta}{n_\theta df_\theta}} \tag{13}$$

where is n_θ equal to n (the number of observations in each cell of the design) multiplied by the product of the numbers of levels in all the factors not represented in the effect currently under consideration; df_θ is the numerator degrees of freedom parameter for the effect under consideration.

Effect Size Estimation

Partial ω^2

Partial ω^2 for an effect θ is defined as

$$\omega_\theta^2 = \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_e^2} \tag{14}$$

where σ_θ^2 is the sum of squared effects divided by the number of effects, i.e., the variance of the effects if they were considered to be a finite statistical population.

Effect Size Estimation

Cohen's f

Cohen's f is defined as

$$f_\theta = \frac{\sigma_\theta}{\sigma_e} \tag{15}$$

(16)

Key Relationships between Effect Size Measures

Key Relationships between Effect Size Measures

We can switch back and forth between λ, f^2 and ω^2 with the formulas

$$f^2 = \frac{\omega^2}{1-\omega^2} = \frac{\lambda}{N_{tot}} \tag{17}$$

and

$$\omega^2 = \frac{f^2}{1+f^2} = \frac{\lambda}{\lambda + N_{tot}} \tag{18}$$

Effect Size Estimation

Effect Size Confidence Intervals

We can exploit the relationships described above to construct confidence intervals on various effect size measures.

Consider the effect of *Instruction*. The F statistic is 2.4684 with 3 and 56 degrees of freedom.

Because the significance test is one-sided, this confidence interval will include zero if and only if the F test for no effect rejects at the 0.05 level of significance.

Effect Size Estimation

Effect Size Confidence Intervals

```
> library(MBESS)
> out <- conf.limits.ncf(2.4684, 0.9, 3, 56)
> out$Lower.Limit

[1] NA

> out$Upper.Limit

[1] 16.8924
```

Using MBESS, we quickly obtain a 90% confidence interval for λ that extends from 0 to 16.8924.

Next, we compute a confidence interval on ω^2, using Equation 18. Since the lower limit on $\lambda_{Instruction}$ is zero, so is the lower limit on $\omega^2_{Instruction}$. The upper limit is

$$\omega^2_{Instruction,Upper} = \frac{16.8924}{16.8924 + 64} = 0.2088 \qquad (19)$$

So the confidence interval ranges from 0 to 0.2088. The effect is at best small.

Effect Size Estimation

Effect Size Confidence Intervals

In a similar vein, we can calculate ω^2 confidence intervals for the *Format* main effect and *Instruction* : *Format* interaction as ranging from 0.0018 to 0.1826 and from 0 to 0.1299.

In spite of the appearance of the interaction plot, it seems there is no interaction.

Power Calculations

However, a cautionary note is in order. It is often the case that power to detect interaction effects is significantly lower than the power to detect main effects in the same design.

To calculate power in this design, we make use of the results in Steiger (2004). Equation 19 in that article states the relationship between λ, the noncentrality parameter of the noncentral F distribution, and Cohen's f and the RMSSE Ψ. Rearranging that equation gives, for any effect θ,

$$\lambda_\theta = \frac{N_{tot}\Psi_\theta^2 df_\theta}{Cells_\theta} = N_{tot} f_\theta^2 \tag{20}$$

Power Calculations

For any main effect, $Cells_\theta$ is the number of levels of the effect, and for any interaction effect, $Cells_\theta$ is the number of cells involved in the interaction.

Power analysis reveals that there is indeed much lower power to detect an interaction effect than there is to detect a main effect in this design.

Specifically, if the RMSSE is 0.50 for all three effects, the power for the three effects is 0.81 for *Instruction*, 0.79 for *Format*, and only 0.49 for the interaction effects.

To calculate power, we simply use Equation 20 to compute λ, and then effortlessly compute power in one line.

For example, for the main effect of *Instruction*, there are 4 levels, so $Cells_\theta = 4$, $df_\theta = 3, N_{tot} = 64$.

Hence,

$$\lambda = \frac{\Psi^2 N_{tot} df}{Cells} = \frac{0.25 \times 64 \times 3}{4} = 12 \qquad (21)$$

Power is then the area to the right of the rejection point in a noncentral F with noncentrality parameter 12, which is

```
> 1 - pf(qf(0.95, 3, 56), 3, 56, 12)

[1] 0.8112157
```