psyc3010 lecture 3

following up significant effects
magnitude of effects

last week: logic and computations of factorial anova
next week: No class (Ekka holiday)
In 2 weeks: higher-order / complex anova; 3-way designs

last week → this week

- last week we went over the conceptual and computational processes involved in **between-subjects factorial anova**
- this week we look at how to follow-up significant main effects and interactions
  - Re-cap from 2nd year: Following up variables with >2 levels
  - Following up main effects: "Main effect" comparisons
  - Following up the interaction: Simple effects and simple comparisons
- we also consider the issue of **effect sizes**
  - Eta squared, Omega squared, partial eta-squared

Overview of today – take 2

- Last week we learned how with a good conceptual understanding of anova, you can do hand calculations (see this week's tutes) or generate the computer output that tells you if your main effects of Factor A and B and interaction are significant
  - Is there a difference among the marginal means of A?
  - Is there a difference among the marginal means of B?
  - Are the simple effects of A different at different levels of B (or vice versa)?
- BUT…a couple of questions still remain for today …
  - how do we follow-up our main effects?
    - Conduct main effect comparisons for sig main effects with >2 levels
  - how do we follow-up our interactions?
    - Conduct simple effect tests
    - Conduct simple comparisons for sig simple effects with >2 levels
  - how *substantial* are any of these effects?
    - Calculate an effect size for each main, interaction, & simple effect
    - Sometimes for comparisons too!

Source table from last week

Summary Table

Source	df	SS	MS	F	sig
A (cons)	2	3332.3	1666.15	20.07	.000
B (dist)	1	168.75	168.75	2.03	.161
AB	2	1978.12	989.06	11.91	.000
Error	42	3487.5	83.02		
Total	47	8966.7			

- in factorial anova get 3 tests - of ***main effects*** and ***interaction***

Main effect of A – H_0: $\mu_1 = \mu_2 = \mu_3$

reject H_0 if:

1. MS_A / MS_{error} results in a significant obtained F

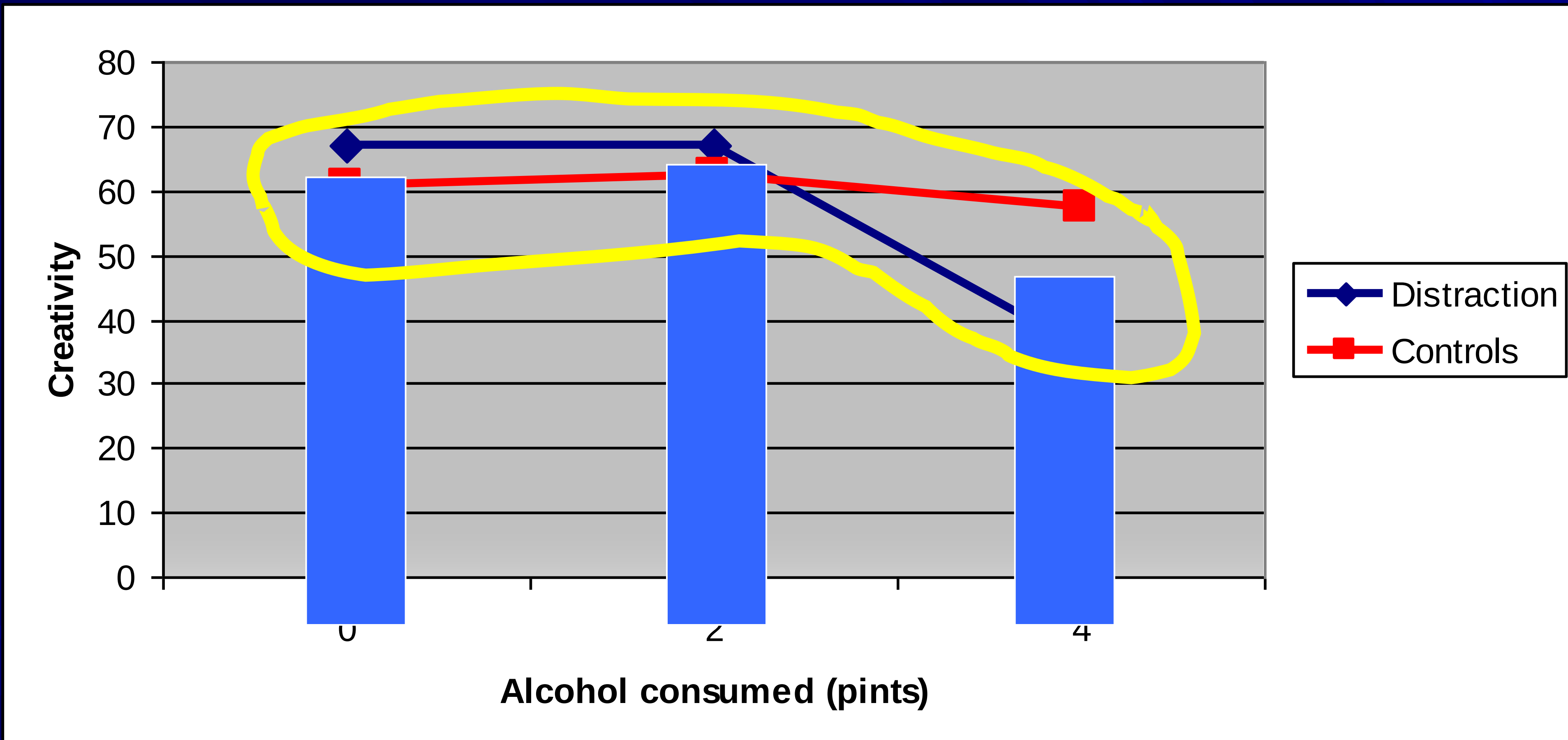
   $F(2,42) = 20.06, p<.05$

2. Indicates that the 3 levels of factor A differ (collapsed across, i.e., ignoring, factor B)

   marginal means of A differ

Participant Distraction	**Alcohol Consumption (pints) 0**	**2**	**4**	Marginal Totals (B) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (*A*)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

main effect of alcohol consumed



Main effect of B – H_0: $\mu_1 = \mu_2$

reject H_0 if:

1. MS_B / MS_{error} results in a significant obtained F

   In our example F was NOT significant:

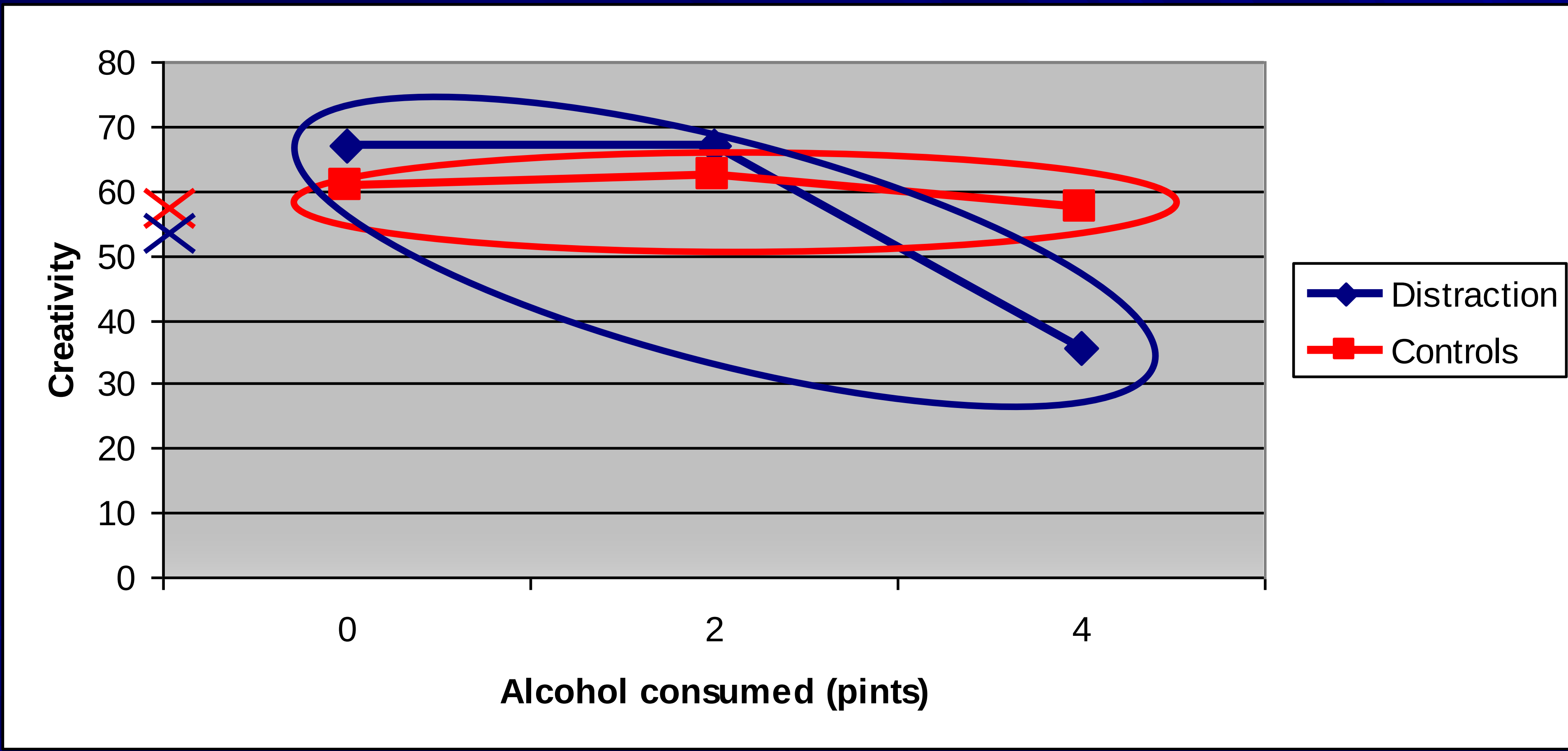
   F(1,42) = 2.03, p>.05, <u>ns</u>

2. If F *was* significant it would indicate that the 2 levels of factor B differ **(collapsed across, i.e., ignoring, factor A)**

   marginal means of B differ

Participant Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (B) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (*A*)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

no main effect of distraction



Interaction of A x B

H_0: $\mu_{11} - \mu_{21} = \mu_{12} - \mu_{22} = \mu_{13} - \mu_{23}$

reject H_0 if:

1. MS_{AB} / MS_{error} results in a significant obtained F

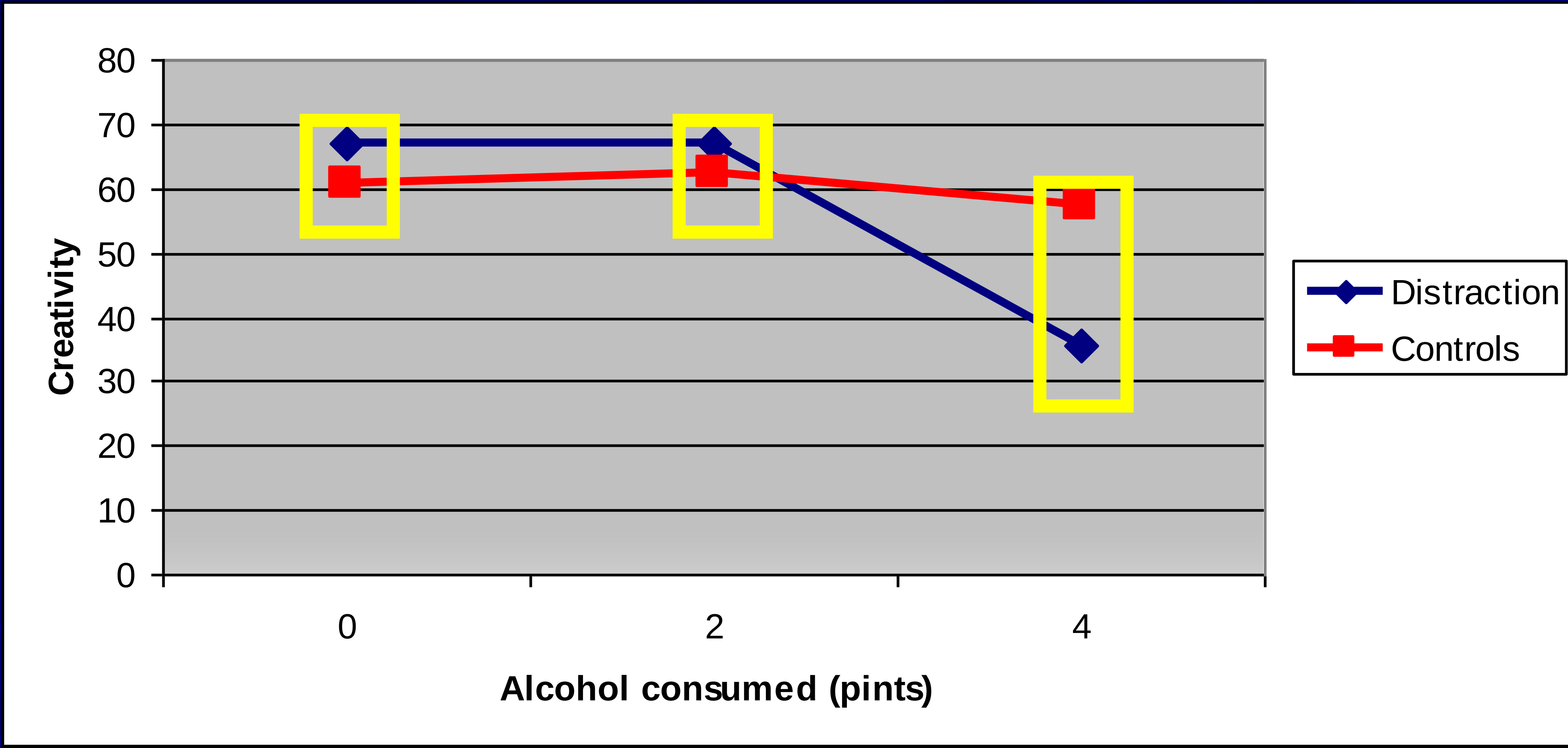
   $F(2,42) = 11.91, p<.05$

2. Indicates that the simple effect of factor B is not the same at all levels of factor A (or vice versa)

   Difference between cell means for levels of factor B changes depending on level of factor A

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (B) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (*A*)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

Interaction



omnibus versus follow-up tests

any test resulting from the preliminary partitioning of variance in anova is called an *omnibus test.*

- "Omni" means all in Latin – omnibus test looks for all possible differences among the levels of a factor
- In factorial ANOVA we have three omnibus tests
- Vs. one-way anova (from 2nd year) – one omnibus test

e.g., comparing 3 means:

H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: there is a difference (somewhere!) among the means

- Omnibus tests sometimes need follow-up tests if significant
- to fully interpret a main effect with 3 or more levels you need to conduct **follow-up tests** such as:
  - a priori t-tests
  - multiple comparisons

following-up effects in one-way designs when the variable has >2 levels (recap 2nd year)

- the ***"protected t-test"*** is used to conduct pairwise comparisons (i.e., compare 2 means),
- (*"protected"* against type-1 error rate inflation)
- just the same as a normal t-test but the error term used is MS_{error}

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{2MS_{error}}{n}}}$$

$$df_{error} = N - ab$$

following-up effects in one-way designs when the variable has >2 levels (recap 2nd year)

- as an alternative, could use **Linear Contrasts** to determine if one group ***or set of groups*** is different from another group ***or set of groups***
- a set of weights, a_j, is used to define the contrast
  e.g., $\bar{X}_1$ & $\bar{X}_2$ vs $\bar{X}_3$ **[1 1 -2]**
- (the protected t-test is a special case of this technique)

$$t = \frac{L}{\sqrt{\frac{\sum a_j^2 MS_{error}}{n}}}$$

$$L = \sum a_j \bar{X}_j$$

$$df_{error} = N - ab$$

following-up main effects in factorial ANOVA

- Look at differences among more than 2 ***marginal*** means
- As for one-way ANOVA, can use ***"protected t-test"*** to conduct pairwise comparisons (only comparing marginal means instead of 'group means')
  - *only do this if the main effect is significant* (we don't follow up ns effects)
  - e.g., to compare effect of 4 pints to 2 pints
  - Have to change the n (must be based on the number of observations in each level we're comparing so n x number of levels of the other IV)

following-up main effects:
(differences among ***marginal*** means)

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{2MS_{error}}{n \times d}}}$$

This formulae would be what you could use to follow up the MAIN EFFECT OF ALCOHOL

$$df_{error} = N - ab$$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{2MS_{error}}{n \times d}}}$$

The **d** here represents "number of levels of the **d**istraction variable" (but you could change the letter!)

$$df_{error} = N - ab$$

so, to follow up our main effect of A (alcohol consumption)…

"are creativity ratings lower after 4 pints than after 0 pints?"

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{2MS_{error}}{n \times d}}} \rightarrow t = \frac{46.56 - 63.75}{\sqrt{\frac{2 \times 83.02}{8 \times 2}}}$$

$$df_{error} = N - ab$$

$t_{obt}(42) = -5.34 > t_{crit}(42) = 2.021$

"Yes, there is a significant difference"

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (B) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (*A*)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

following-up main effects:

(differences among ***marginal*** means)

- as an alternative, could use **Linear Contrasts** to determine if one group ***or set of groups*** is different from another group ***or set of groups***
- a set of weights, a_j, is used to define the contrast
  e.g., $\bar{X}_1$ & $\bar{X}_2$ vs $\bar{X}_3$ **[1 1 -2]**
- (the protected t-test is a special case of this technique)

$$t = \frac{L}{\sqrt{\dfrac{\sum a_j^{\,2} MS_{error}}{n \times \#.levels.of.other.IV}}}$$

$$\mathrm{L} = \sum \mathrm{a_j} \overline{\mathrm{X}}_\mathrm{j}$$

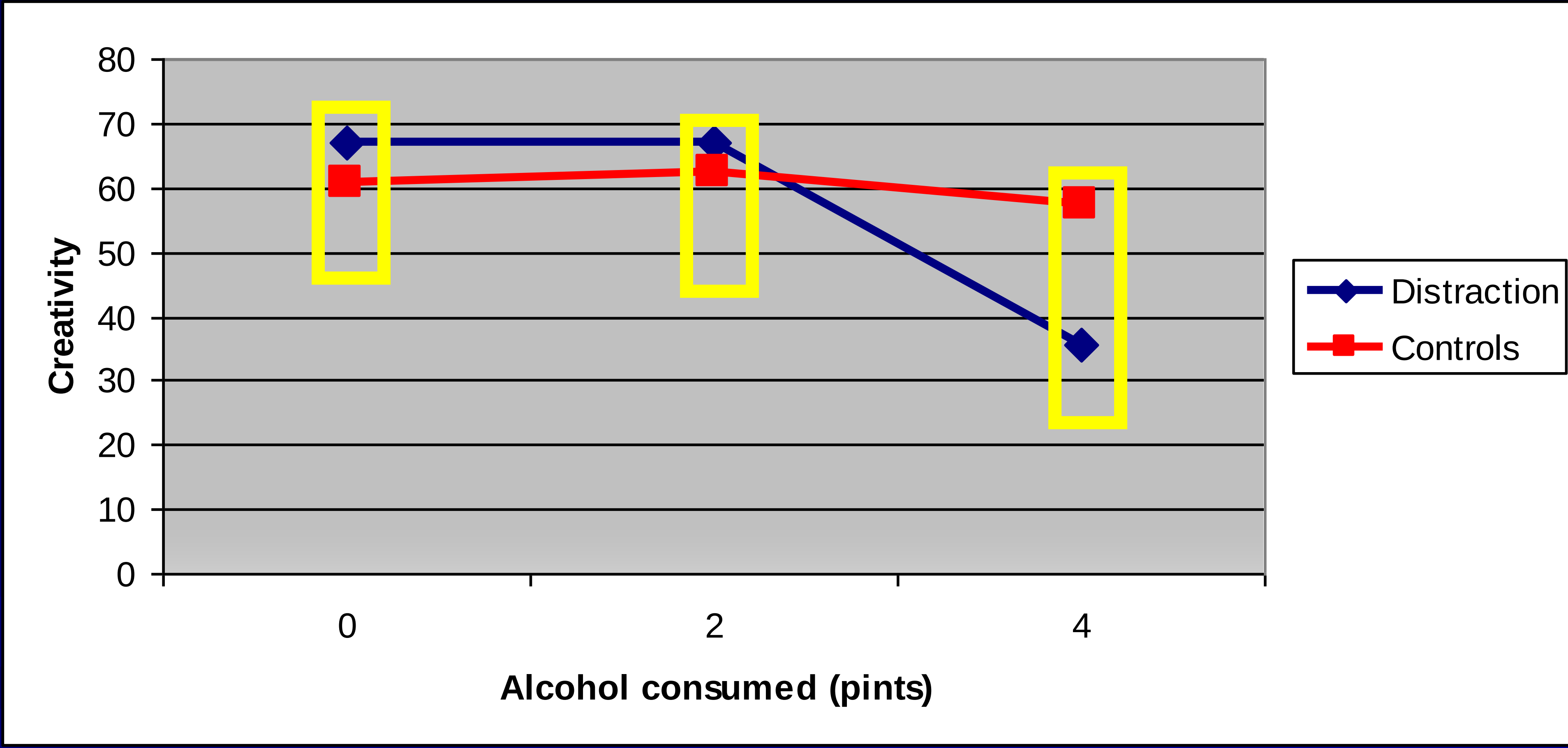
$$df_{error} = N - ab$$

NB: THIS METHOD WILL BE COVERED IN TUTORIALS

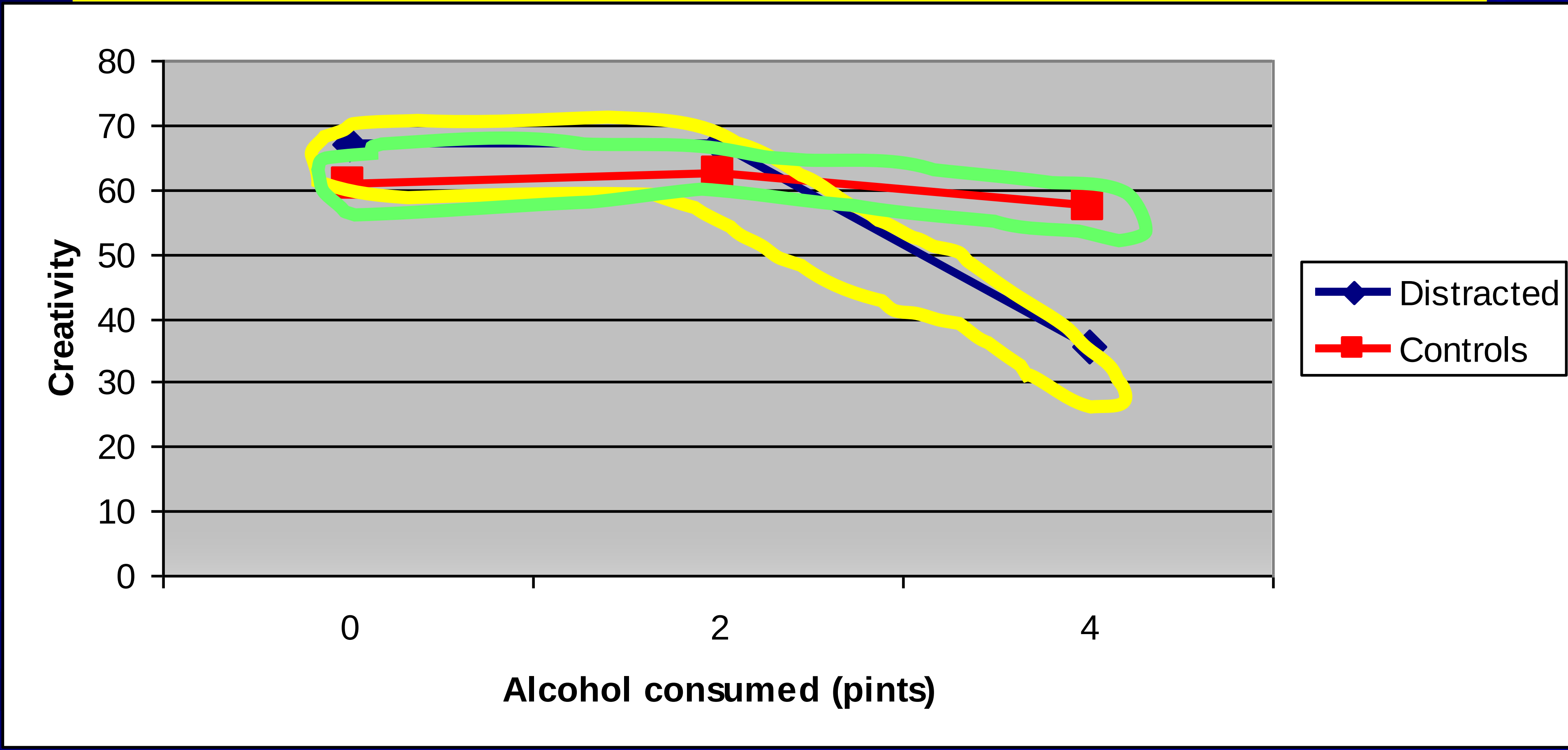
following-up interactions
(differences among ***cell*** means)

- **a significant interaction needs to be followed up with *simple effects***
  - simple effects describe differences among cell means within a row or column, or the ***effects of one factor at each level of the other factor***
  - just like a series of one-way anovas conducted at each level of a factor, except the pooled error term is used (MS_{error})

the ***simple effects of distraction*** describe the differences in creativity between distracted and controls ***at each level of alcohol consumed***



the ***simple effects of alcohol consumed*** describe the differences in creativity after 0, 2 or 4 pints consumed at ***each level of distraction***



simple effects of distraction

"what is the effect of distraction at each level of consumption?"

is there an effect of distraction for participants who have consumed….

0 pints?

2 pints?

4 pints?

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (B) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	***535***	***535***	***285***	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
Controls				
Cell Totals	[illegible]	[illegible]	[illegible]	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (*A*)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

simple effects of distraction

effect after 0 pints

$$SS_{Distraction.at.Consumption1} = \frac{\sum T_{D.at.C1}^{2}}{n} - \frac{T_{C1}^{2}}{nd}$$

$$= \frac{535^2 + 485^2}{8} - \frac{1020^2}{16}$$

$$= 156.25$$

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D)
	50			
	55			
	80			
	65			
Distraction	70			
	75			
	75			
	65			
Cell Totals	**535**			
Cell Means	***66.88***			
	65			
	70			
	60			
Controls	60			
	60			
	55			
	60			
	55			
Cell Totals	**485**			
Cell Means	***60.63***			
Marginal Totals (C)	***1020***			
Means	***63.75***			

note, to help remember which factor we are talking about, we can use labels other than A and B

– e.g., D = distraction and C = consumption

simple effects of distraction

effect after 2 pints

$$SS_{Distraction.at.Consumption\,2} = \frac{\sum T_{D.at.C_2}^{\;2}}{n} - \frac{T_{C_2}^2}{nd}$$

$$= \frac{535^2 + 500^2}{8} - \frac{1035^2}{16}$$

$$= 76.56$$

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D)
Distraction		45		
		60		
		85		
		65		
		70		
		70		
		80		
		60		
Cell Totals		**535**		
Cell Means		***66.88***		
Controls		70		
		65		
		60		
		70		
		65		
		60		
		60		
		50		
Cell Totals		**500**		
Cell Means		***62.50***		
Marginal Totals (C)		***1035***		
Means		***64.69***		

simple effects of distraction

effect after 4 pints

$$SS_{Distraction.at.Consumption3} = \frac{\sum T_{D.at.C_3}^{\;2}}{n} - \frac{T_{C_3}^2}{nd}$$

$$= \frac{285^2 + 460^2}{8} - \frac{745^2}{16}$$

$$= 1914.06$$

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D)
			30	
			30	
			30	
			55	
Distraction			35	
			20	
			45	
			40	
Cell Totals			**285**	
Cell Means			***35.63***	
			55	
			65	
			70	
Controls			55	
			55	
			60	
			50	
			50	
Cell Totals			**460**	
Cell Means			***57.50***	
Marginal Totals (C)			***745***	
Means			***46.56***	

summary table for simple effects of distraction

Source	*SS*	*df*	*MS*	*F*	*p*
D at C1	156.25	1	156.25	1.88	0.177
D at C2	76.56	1	76.56	0.92	0.342
D at C3	1914.06	1	1914.06	23.05	0.000
Error	3487.5	42	83.04		

critical *F* at alpha=.05 (1,42) = 4.08

if obtained F exceeds critical F ***reject the null hypothesis***

These are your calculated SS values

Degrees of freedom for a simple effect are just the df for the associated main effect

df = $df_{distraction}$ (2-1) = 1

Source	*SS*	*df*	*MS*	*F*	*p*
D at C1	**156.25**	**1**	**156.25**	**1.88**	0.177
D at C2	**76.56**	**1**	**76.56**	**0.92**	0.342
D at C3	**1914.06**	**1**	**1914.06**	**23.05**	0.000
Error	**3487.5**	**42**	**83.04**		

critical F at alpha=.05 (1,42) = 4.08

if obtained F exceeds critical F ***reject the null hypothesis***

SS_{error} term (and df) is taken from the main anova (calculated last week)

Mean Squares and F values calculated as SS/df and MSeffect/MSerror

simple effects of consumption

"what is the effect of consumption at each level of distraction?"

is there an effect of consumption for….

distracted?

controls?

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (C)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

simple effects of consumption

effect in distracted group

$$SS_{Consumption.at.Distraction_1} = \frac{\sum T_{C.at.D_1}^{\ \ 2}}{n} - \frac{T_{D_1}^2}{nc}$$

$$= \frac{535^2 + 535^2 + 285^2}{8} - \frac{1355^2}{24}$$

= 5208.33

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D) (means)
	50	45	30	
	55	60	30	
	80	85	30	
	65	65	55	
Distraction	70	70	35	
	75	70	20	
	75	80	45	
	65	60	40	
Cell Totals	**535**	**535**	**285**	***1355***
Cell Means	***66.88***	***66.88***	***35.63***	***56.46***

simple effects of consumption

effect in control group

$$SS_{Consumption.at.distraction_2} = \frac{\sum T_{C.at.D_2}^{\ 2}}{n} - \frac{T_{D_2}^2}{nc}$$

$$= \frac{485^2 + 500^2 + 460^2}{8} - \frac{1445^2}{24}$$

= 102.08

Distraction	Alcohol Consumption (pints) 0	2	4	Marginal Totals (D)
	65	70	55	
	70	65	65	
	60	60	70	
Controls	60	70	55	
	60	65	55	
	55	60	60	
	60	60	50	
	55	50	50	
Cell Totals	**485**	**500**	**460**	***1445***
Cell Means	***60.63***	***62.50***	***57.50***	***60.21***
Marginal Totals (C)	***1020***	***1035***	***745***	***2800***
Means	***63.75***	***64.69***	***46.56***	***58.33***

summary table for simple effects of consumption

Source	*SS*	*df*	*MS*	*F*	*p*
C at D1	5208.33	2	2604.17	31.36	0.000
C at D2	102.08	2	51.04	0.61	0.546
Error	3487.5	42	83.04		

critical F at alpha=.05 (242) = 3.23

if obtained F exceeds critical F ***reject the null hypothesis***

These are your calculated SS values

Degrees of freedom for a simple effect are just the df for the associated main effect

$df = df_{consumption}$ (3-1) = 2

Source	*SS*	*df*	*MS*	*F*	*p*
C at D1	**5208.33**	**2**	**2604.17**	**31.36**	0.000
C at D2	**102.08**	**2**	**51.04**	**0.61**	0.546
Error	**3487.5**	**42**	**83.04**		

critical F at alpha=.05 (242) = 3.23

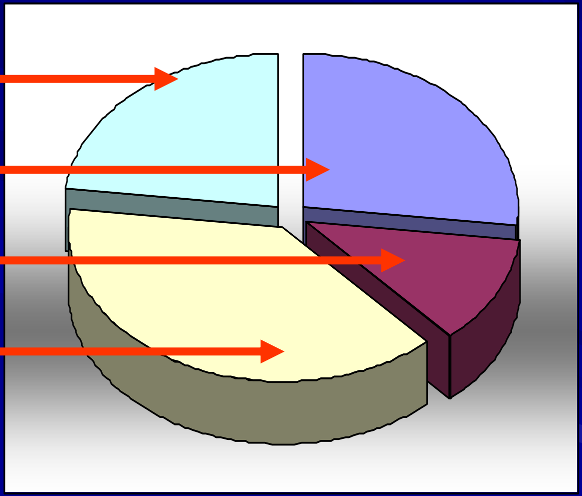
if obtained F exceeds critical F ***reject the null hypothesis***

SS_{error} term (and df) is taken from the main anova (calculated last week)

Mean Squares and F values calculated as per last week

additivity of omnibus tests

- *remember* - in anova what we are doing is **partitioning variance**
- a 2x3 between-subjects anova partitions the total variance into 4 parts:
  - effect due to first factor
  - effect due to second factor
  - effect due to interaction
  - error/residual/within group variance



$$SS_{total} = SS_C + SS_D + SS_{CD} + SS_{error}$$

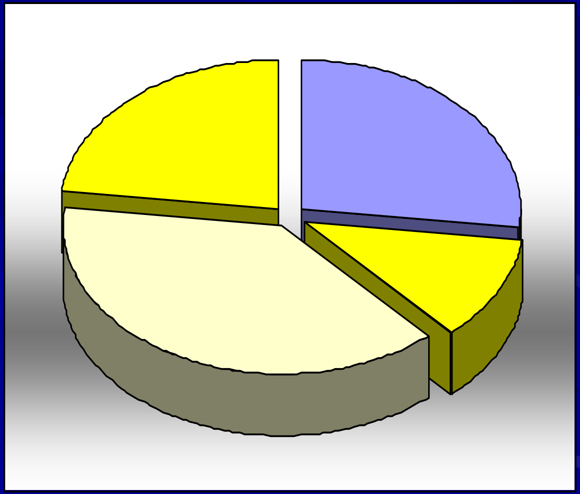
additivity of simple effects

- simple effects re-partition the main effect and interaction variance
- ∑ simple effects of factor 1 = ∑ main effect 1 + interaction

e.g., in our example

$SS_{distraction} + SS_{interaction}$
= 168.75 + 1978.12
= 2146.87

$SS_{distraction\ at\ C1} + SS_{distraction\ at\ C2} + SS_{distraction\ at\ C3}$
= 156.25 + 76.56 + 1914.06
= 2146.87



And ∑ df for simple effects of factor 1 = ∑ df main effect 1 + df interaction

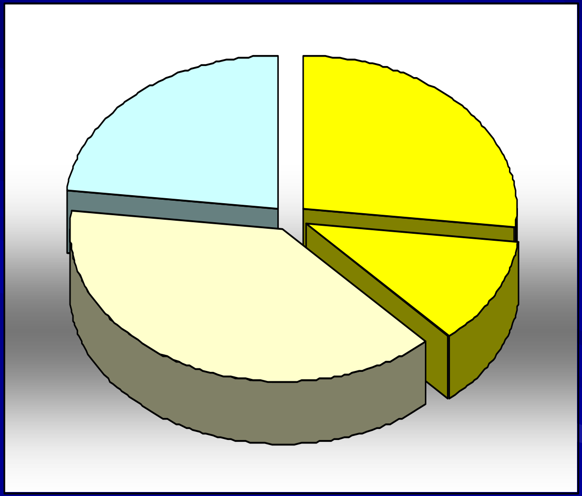
additivity of simple effects

- simple effects re-partition the main effect and interaction variance
- ∑ simple effects of factor 2 = ∑ main effect 2 + interaction

e.g., in our example

$SS_{consumption} + SS_{interaction}$
= 3332.29 + 1978.12
= 5310.41

$SS_{consumption\ at\ D1} + SS_{consumption\ at\ D2}$
= 5208.33 + 102.08
= 5310.41



And ∑ df for simple effects of factor 2 = ∑ df main effect 2 + df interaction

But immediately upon this I observed that, whilst I thus wished to think that all was false, it was absolutely necessary that I, who thus thought, should be somewhat; and as I observed that this truth, I think, therefore I am, was so certain and of such evidence that no ground of doubt, however extravagant, could be alleged by the sceptics capable of shaking it, I concluded that I might, without scruple, accept it as the first principle of the philosophy of which I was in search.

- René Descartes

following-up simple effects:

Linear Contrasts and simple comparisons

- **consider the significant simple effect of consumption for *distracted participants*:**
  - indicates that, *for distracted*, there is a difference among the means over the 3 levels of consumption (0 pints, 2 pints, 4 pints)

- **follow-up using Simple Comparisons (Linear Contrasts)**
  - the procedure is *identical* to that used for following up main effects, except comparisons are between **cell means**, not **marginal means**
  - *note:* only ***significant simple effects*** should be followed up

simple comparisons for consumption (distracted)

	Consumption		
	0 pints	2 pints	4 pints
Distracted	66.88	66.88	35.63
Contrast 1	2	-1	-1
Contrast 2	0	1	-1

these are the cell means for distracted participants from our data table earlier

a set of weights (a_j) is used to define the contrasts:

contrast 1 compares 0 vs 2 & 4

contrast 2 compares 2 vs 4

	Consumption		
	0 pints	2 pints	4 pints
Distracted	**66.88**	**66.88**	**35.63**
Contrast 1	**2**	**-1**	**-1**
Contrast 2	**0**	**1**	**-1**

contrasts are orthogonal:

- $\sum a_j = 0$
- $\sum a_j b_j = 0$
- Number of contrasts = df for effect

Calculations for contrast 1

$$t = \frac{L}{\sqrt{\frac{\sum a_j^{\,2} MS_{error}}{n}}}$$

$$L = \sum a_j \overline{X}_j$$

$$df_{error} = N - ab$$

	Consumption		
	0 pints	2 pints	4 pints
Distracted	**66.88**	**66.88**	**35.63**
Contrast 1	**2**	**-1**	**-1**
Contrast 2	0	1	-1

L = 2(66.88) – 1(66.88) – 1(35.63) = ***31.25***

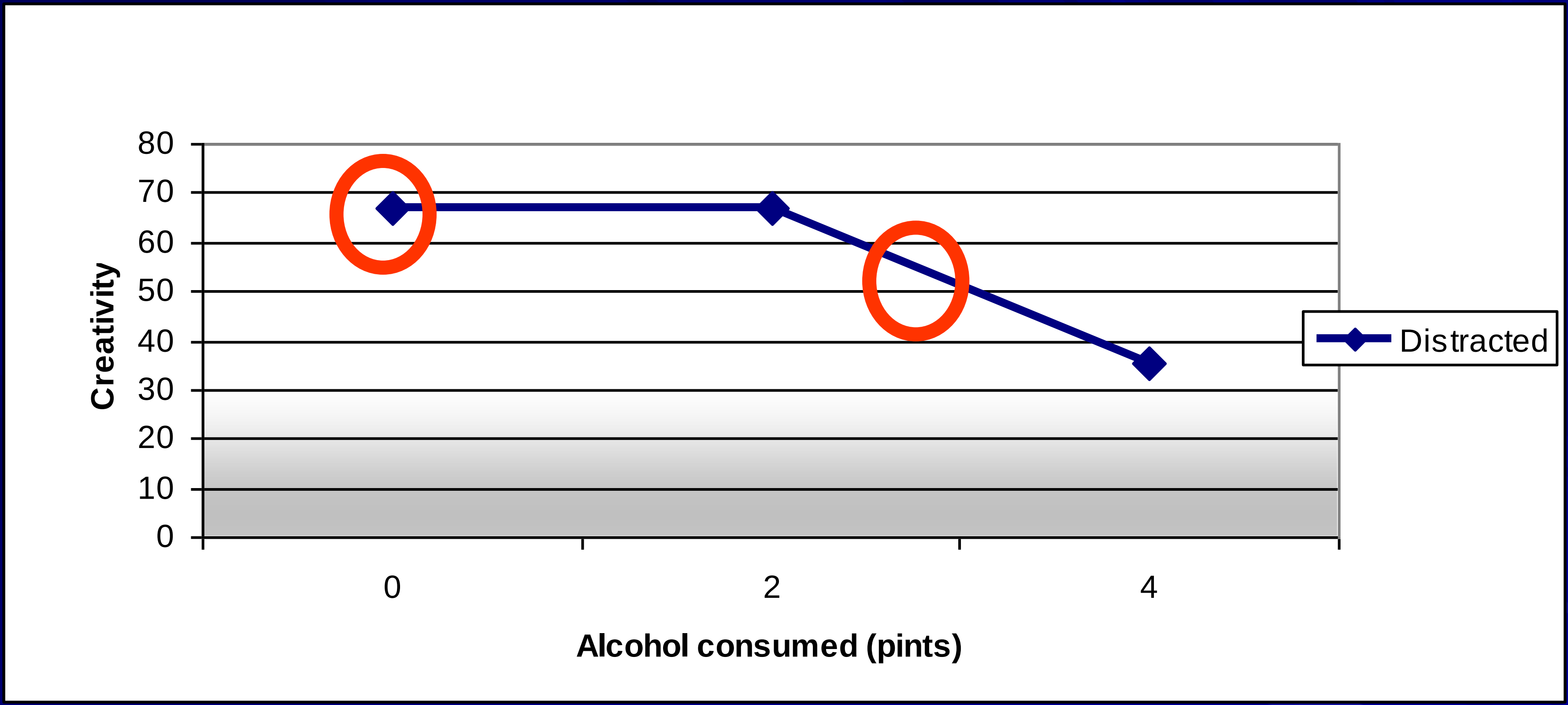
$$t = \frac{31.25}{\sqrt{\frac{(2^2 + (-1)^2 + (-1)^2)83.04}{8}}} = 3.96$$

$t_{\alpha=.05}$ (42) = 2.02 (unadjusted)

$t_{\alpha=.05}$ (42) = 2.33 (adjusted)

(Bonferroni adjustment for 2 comparisons)

contrast 1 – what does it do?



contrast 1 **compares *(for distracted participants only)* the mean creativity rating for participants who have had *0 pints* with the mean attractiveness rating for participants who have had *2 or 4 pints***

***t*(42) = 3.96, p<.05** → significant

Calculations for contrast 2

$$t = \frac{L}{\sqrt{\frac{\sum a_j^2 MS_{error}}{n}}}$$

$$L = \sum a_j \overline{X}_j$$

$$df_{error} = N - ab$$

	Consumption		
	0 pints	2 pints	4 pints
Distracted	**66.88**	**66.88**	**35.63**
Contrast 1	2	-1	-1
Contrast 2	**0**	**1**	**-1**

$L = 0(66.88) + 1(66.88) - 1(35.63) = \mathbf{31.25}$

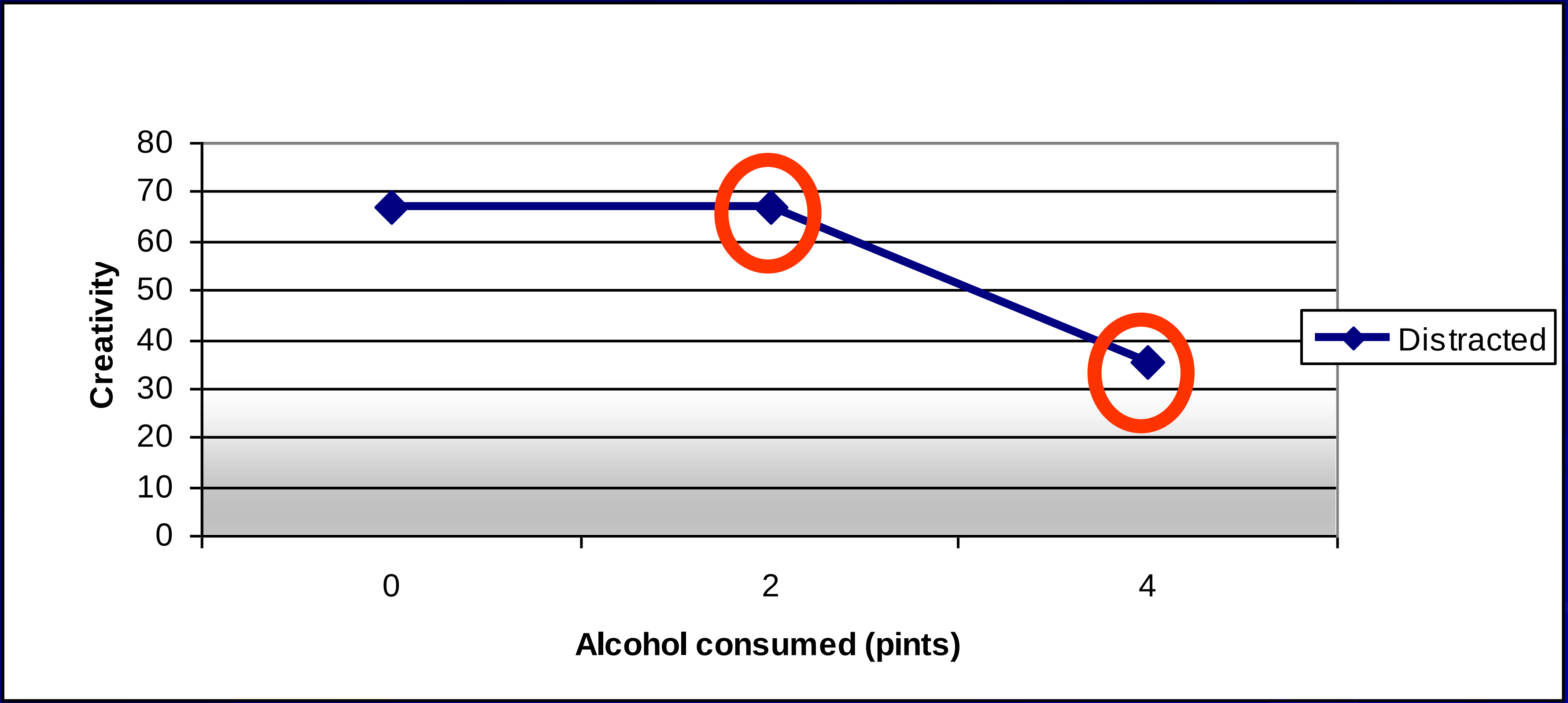
$$t = \frac{31.25}{\sqrt{\frac{(0^2 + 1^2 + (-1)^2)83.04}{8}}} = 6.86$$

$t_{\alpha=.05}(42) = 2.02$ (unadjusted)

$t_{\alpha=.05}(42) = 2.33$ (adjusted)

(Bonferroni adjustment for 2 comparisons)

contrast 2 – what does it do?



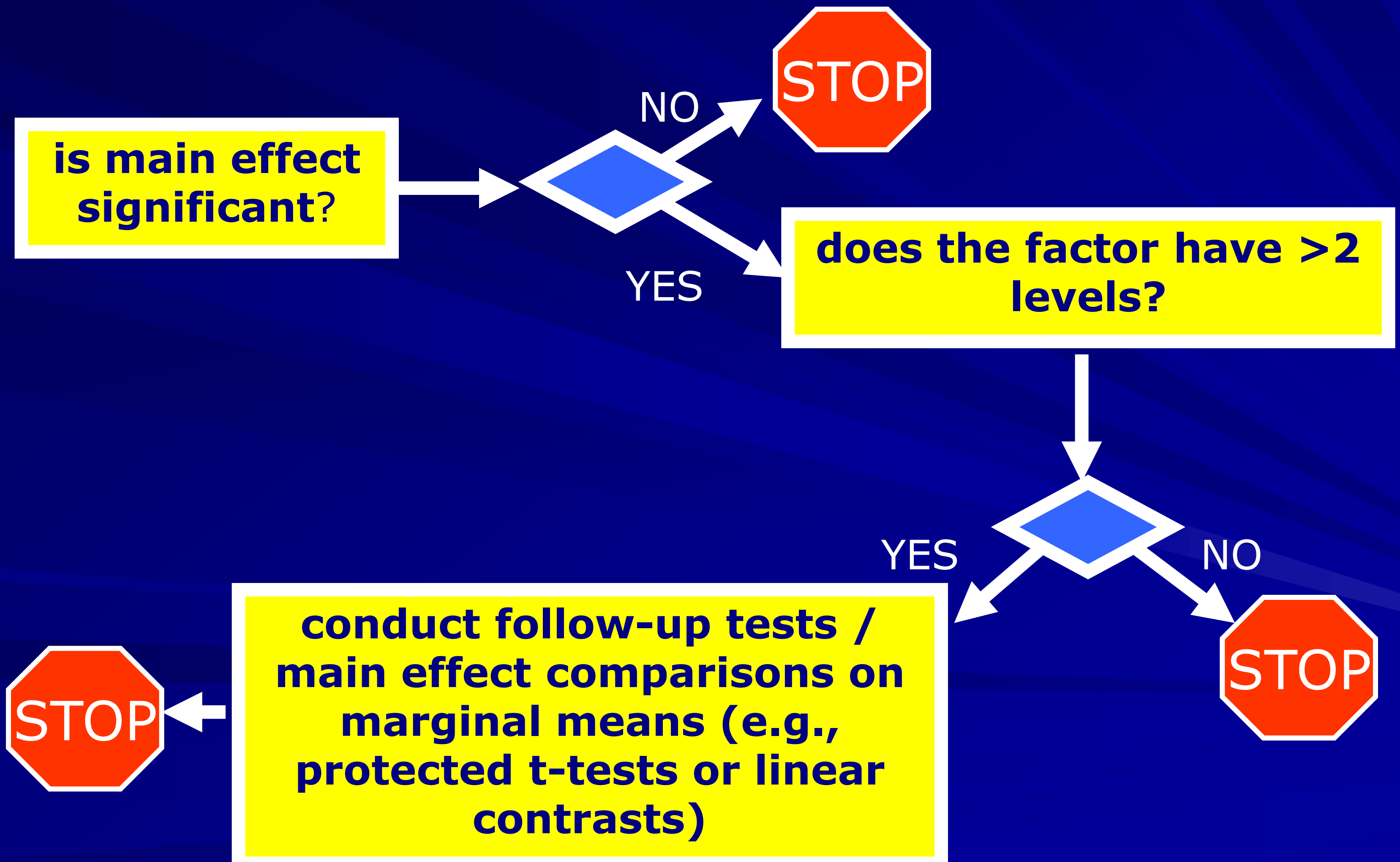
contrast 2 **compares *(for distracted participants only)* the mean creativity rating for participants who have had *2 pints* with the mean attractiveness rating for participants who have had *4 pints***

***t*(42) = 6.86, p<.05** → significant

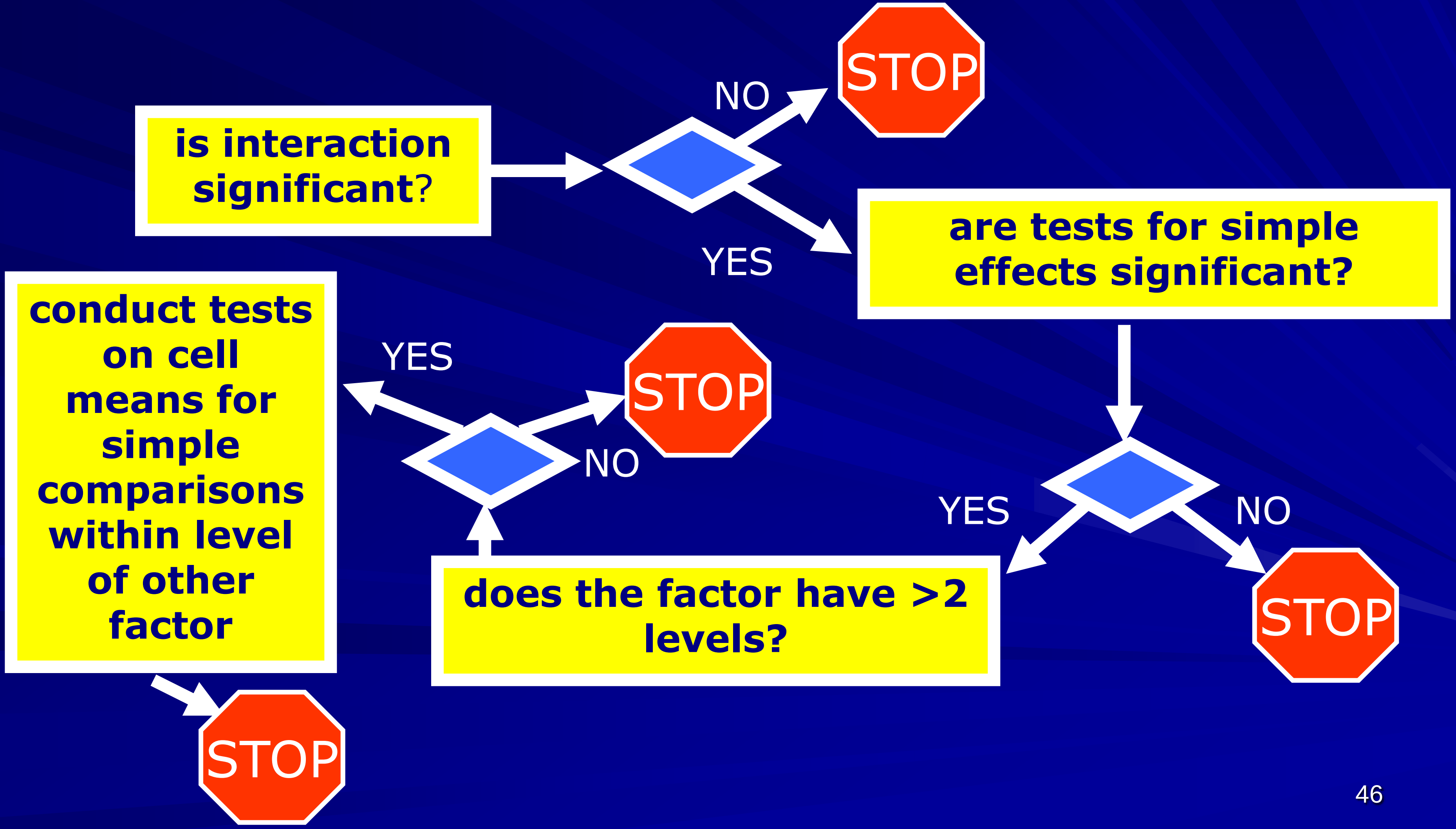
issues with follow-up comparisons

- **redundancy – explaining the same mean difference more than once**
  - solution – ***orthogonal*** (independent) linear contrasts (see Howell p 382-384)
- **increases in familywise error rate**
  - type-1 error rate is α for each test, this leads to higher probability of committing a type-1 error over all tests
  - solution 1 – use ***Bonferroni Adjustment*** (from Bonferroni *t'*-tables) for critical *t*
  - solution 2 – conduct contrasts defined *a priori*, rather than exhaustive orthogonal set (i.e., do fewer contrasts).

steps for following-up main effects



steps for following-up a 2-way interaction



significance tests:
do they tell the whole story?

- **significance testing was uncritically accepted as a way of determining the importance of findings until recently…**
  - the use of an arbitrary acceptance criterion (α) results in a binary outcome – significant or non-significant
  - many researchers interpret significance values (p) improperly
  - a large p-value (non-significant) will eventually slip under the acceptance criterion with increases in sample size
  - ***the magnitude of experimental effect***, or ***effect size***, has been proposed as an accompaniment (if not an outright replacement) to significance testing

magnitude of experimental effects

- **two basic approaches to estimating effect size in anova**

eta-squared η^2

- $\eta^2 = \dfrac{SS_{effect}}{SS_{total}}$
- describes the ***proportion of variance*** in the DV ***in the sample*** that is accounted for by the effect
- considered a biased estimate of the true magnitude of the effect in the population
- still the most common, widely reported effect size measure because easily interpretable (like R^2)

omega-squared ϖ^2

- $\varpi^2 = \dfrac{SS_{effect} - (df_{effect})MS_{error}}{SS_{total} + MS_{error}}$
- describes the ***proportion of variance*** in DV scores ***in the population*** that is accounted for by the effect
- a less biased (more conservative) estimate of the effect size

difference between the two estimates depends on ***sample size*** and ***error variance.***

why does effect size matter?

- a statistically ***significant*** result may be trivial or of little practical significance, e.g., age differences in height
- a statistically ***non-significant*** result may be important, e.g., suicidal behaviour, or it could account for a large proportion of variance in DV scores (particularly with small samples)
- the effect size gives you another way of assessing the reliability of the result ***in terms of variance*** (the very underpinning of ANOVA)
- Because of dependence on sample size and the fact that an acceptance criterion is used in hypothesis testing, interpreting the size of the *F*-ratio or *p*-value is spurious – need a real measure of variance accounted for (eta sq or omega sq)
- there are difficulties with defining a “large enough” effect size – Cohen (1973) suggests that .2 is small, .5 is medium and .8 is large (in social sciences, .02, .05, .08?)

Summary Table – from last week

Source	df	SS	MS	F	sig
C (cons)	2	3332.3	1666.15	20.07	.000
D (distr)	1	168.75	168.75	2.03	.161
C x D	2	1978.12	989.06	11.91	.000
Error	42	3487.5	83.02		
Total	47	8966.7			

eta-squared:

$$\eta 2 = \frac{SS_{effect}}{SS_{total}}$$

Consumption
= 3332.3 / 8966.7
= .37 (37% var)

Distraction
= 168.75 / 8966.7
= .02 (2% var)

C x D
= 1978.12 / 8966.7
= .22 (22% var)

Number of pints consumed explains 37% of the variance in DV scores (creativity ratings)

Distraction explains 2% of variance in DV scores

Interaction between distraction and number of pints consumed explains 22% in DV scores

Summary Table – from last week

Source	df	SS	MS	F	sig
C (cons)	2	3332.3	1666.15	20.07	.000
D (distr)	1	168.75	168.75	2.03	.161
C x D	2	1978.12	989.06	11.91	.000
Error	42	3487.5	83.02		
Total	47	8966.7			

omega-squared:

$$\varpi 2 = \frac{SS_{effect} - (df_{effect})MS_{error}}{SS_{total} + MS_{error}}$$

produces very similar but smaller ***(more conservative)*** *estimates*

Consumption

= [3332.3 - 2(83.02)] / (8966.7 + 83.02)

= .34 (34% var)

Distraction

= [168.75 - 1(83.02)] / (8966.7 + 83.02)

= .01 (1% var)

C x D

= [1978.12 – 2(83.02)] / (8966.7 +83.02)

= .20 (20% var)

what's a partial eta squared?

- "compare means" command in SPSS:
  asking for an ANOVA with effect size measure gives you **eta squared** – proportion of **total variance** accounted for by the effect

$$\eta^2 = \frac{SS_{\text{effect}}}{SS_{\text{total}}}$$

- UNIANOVA, MANOVA or GLM:
  asking for effect size gives you **partial eta squared** – proportion of **residual variance** accounted for by the effect

$$\text{partial}\,\eta^2 (\eta^2{}_{\text{p}}) = \frac{SS_{\text{effect}}}{SS_{\text{effect}} + SS_{\text{error}}}$$

limitations of partial η^2

$$\eta^2 = \frac{SS_{effect}}{SS_{total}}$$

$$\text{partial}\,\eta^2(\eta^2_p) = \frac{SS_{effect}}{SS_{effect} + SS_{error}}$$

1. in factorial ANOVA, [error + effect] is less than [total], so partial η^2 is more liberal or inflated

$SS_{effect1} = 4$
$SS_{effect2} = 4$
$SS_{intx} = 4$
$SS_{error} = 4$

$SS_{total} = 16$

effect size for factor 1:

$\eta^2 = 4 / 16 = .25$

partial $\eta^2 = 4 / (4 + 4)$
$= 4 / 8 = .5$

limitations of partial η^2

2. in factorial ANOVA, η^2 adds up to a maximum of 100%, but partial η^2 can add to > 100%

$SS_{effect1} = 4$
$SS_{effect2} = 4$
$SS_{intx} = 4$
$SS_{error} = 4$

$SS_{total} = 16$

(calculations as in previous slide)

	η^2	partial η^2
effect 1	.25	.5
effect 2	.25	.5
interaction	.25	.5
error	.25	.5
sum	**1.00**	**2.00 (!)**

→ ***hard to make meaningful comparisons with partial η^2***

A random illustrative example of SPSS output:

Tests of Between-Subjects Effects

Dependent Variable: I am dissatisfied with my performance recently

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	54.963[a]	14	3.926	1.822	.066	.367
Intercept	250.920	1	250.920	116.420	.000	.726
gendernum	.070	1	.070	.032	.858	.001
gpaestimate	44.031	9	4.892	2.270	.035	.317
gendernum * gpaestimate	9.326	4	2.331	1.082	.377	.090
Error	94.834	44	2.155			
Total	821.000	59				
Corrected Total	149.797	58				

a. R Squared = .367 (Adjusted R Squared = .165)

- The partial eta squared says the interaction between my two IVs accounts for 9% of the residual variance
- But calculating the eta squared (9.326 / 149.797) shows the interaction only accounts for 6% of the total variance
- Omega squared < Eta squared < partial Eta squared

summary

- **main effects and interactions are omnibus tests**
  - main effects may have to be followed up with main effect comparisons – usually use **protected t-tests** or **linear contrasts**
  - interactions may have to be followed up with simple effects, which may in turn need to be followed up with simple comparisons – usually use linear contrasts focusing on theoretically relevant differences.

- **statistical significance is not the be-all-and-end-all**
  - useful to estimate the size of effects – give more information than statistical significance
  - two approaches – eta-squared (biased, but common) and omega-squared (less biased, but uncommon)
  - partial eta-squared is an even more commonly reported effect size measure (output by SPSS) which is the portion of residual variance the effects accounts for

This week's readings:

factorial ANOVA and follow-up tests…

- Field – Chapter 10 (sections 10.1, 10.2)
- Howell – Chapter 12 (sections 12.2, 12.3)
  & Chapter 13 (sections 13.4, 13.5, 13.6)

effect sizes…

- Field – Chapter 10 (section 10.6)
- Howell – Chapter 11 (section 11.12)
  & Chapter 13 (section 13.9)

higher-order designs (next week's topic)…

- Howell – Chapter 13 (section 13.13)

Tutes this week focus on hand calculations for 2-way designs

The Ekka break and after:

- I will not be holding consult hours next Wednesday (Ekka); my office hours resume the following Wednesday the 20th (1-3pm).
- When we return, you should aim to have mastered readings, exercises, understanding for two-way designs
- We will then focus on higher-order designs and 3-way ANOVA
- Assignment 1 will be distributed in class on Wednesday the 20th and posted on the web by Monday the 18th.