Between-Subjects ANOVA with Multiple Factors

Consider the following design

					Tricky Equal Sign Problems			
8-9 year olds					0	0	0	1
9-10 year olds					2	3	2	1

Consider the following design

	Normal Equal Sign Problems				Tricky Equal Sign Problems			
8-9 year olds	3	4	3	4	0	0	0	1
9-10 year olds	4	4	4	3	2	3	2	1

Null Hypotheses for “Main Effects”

H_{O1}: The two age groups perform similarly

H_{O2}: Children perform similarly on normal and tricky problems

H_{O1}: $\mu_8 = \mu_9$

H_{O2}: $\mu_{Normal} = \mu_{Tricky}$

	Normal Equal Sign Problems				Tricky Equal Sign Problems				
8-9 year olds	3	4	3	4	0	0	0	1	μ_8
9-10 year olds	4	4	4	3	2	3	2	1	μ_9
	μ_{Normal}				μ_{Tricky}				

Interaction between the Factors

H_{O1}: The four age groups perform similarly

H_{O2}: Children perform similarly on normal and tricky problems

H_{O3}: The differences between age groups are the same regardless of type of problem

H_{O3}: Differential performance on problem types is independent of age group

Null Hypotheses

H_{O1}: $\mu_8 = \mu_9$

H_{O2}: $\mu_{Normal} = \mu_{Tricky}$

H_{O3}: $(\mu_{N8} - \mu_{N9}) = (\mu_{T8} - \mu_{T9})$

H_{O3}: $(\mu_{N8} - \mu_{T8}) = (\mu_{N9} - \mu_{T9})$

Mean Scores

	Normal Equal Sign Problems	Tricky Equal Sign Problems
8-9 year olds	3.50	0.25
9-10 year olds	3.75	2.00
Difference between Rows	**0.25**	**1.75**

Mean Scores

	Normal Equal Sign Problems	Tricky Equal Sign Problems	Difference between Column
8-9 year olds	3.50	0.25	**3.25**
9-10 year olds	3.75	2.00	**1.75**

“Old” Models

Restricted: $Y_{ij} = \mu + e_{ij}$

Full: $Y_{ij} = Ybar_{j} + e_{ij}$

Another Perspective

Restricted: $Y_{ij} = \mu + e_{ij}$

Full: $Y_{ij} = \mu + Age_j + e_{ij}$

Full Model for Two-Way ANOVA

Full: $Y_{ijk} = \mu + Age_j + Problem_k + (AP)_{jk} + e_{ijk}$

Three Restricted Models

When testing the Age main effect…

- $Y_{ijk} = \mu + Problem_k + (AP)_{jk} + e_{ijk}$

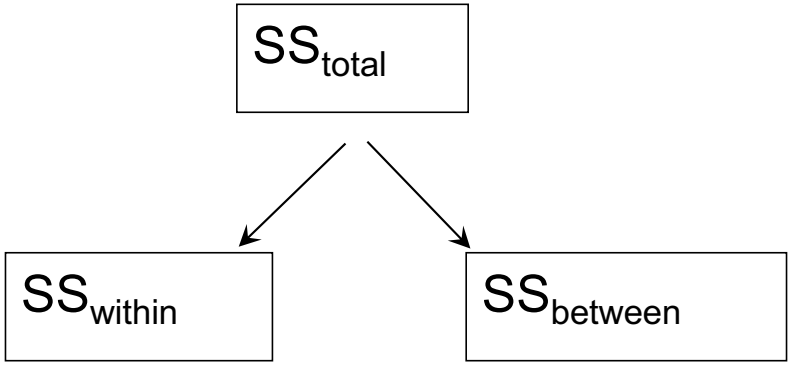
When testing the Problem main effect…

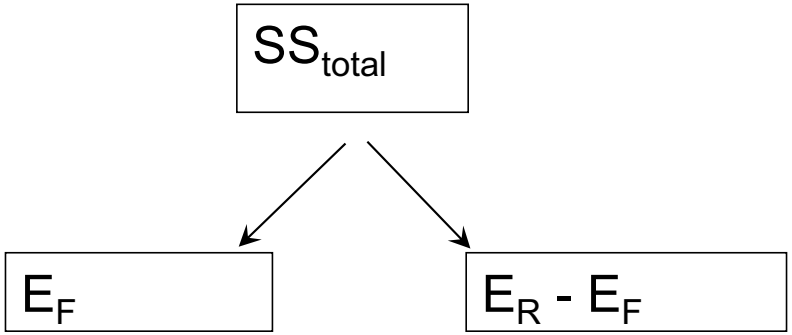
- $Y_{ijk} = \mu + Age_j + (AP)_{jk} + e_{ijk}$

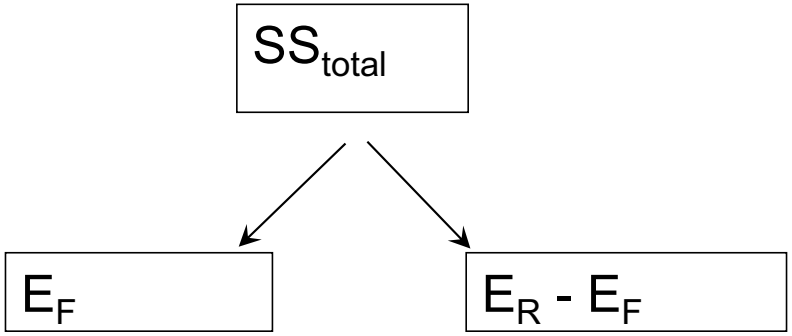
When testing the Interaction between Age and Problem…

- $Y_{ijk} = \mu + Age_j + Problem_k + e_{ijk}$

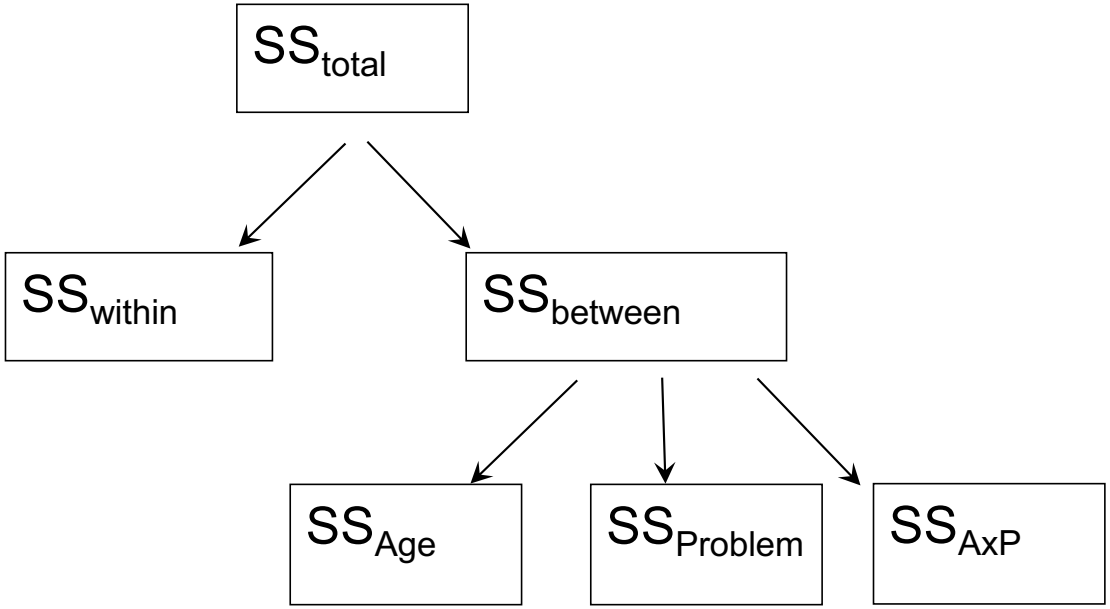
SS_{total}

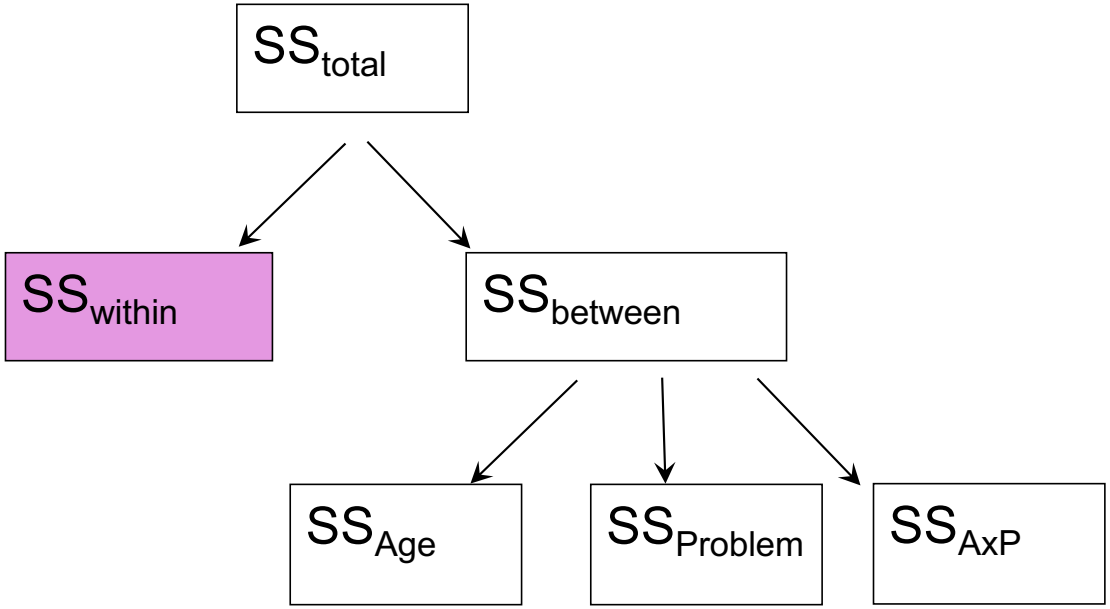

SS_{total}
SS_{within}
$SS_{between}$


SS_{total}
E_F
$E_R - E_F$



Reminder: This is basically our F test


SS_{total}
SS_{within}
$SS_{between}$
SS_{Age}
$SS_{Problem}$
SS_{AxP}



MS_{within} becomes the denominator for all F-tests

F Table for Findings

	SS	df	MS	F
Age	SS_{Age}	df_{Age}	$\frac{SS_{Age}}{df_{Age}}$	$\frac{MS_{Age}}{MS_W}$
Problem	$SS_{Problem}$	$df_{Problem}$	$\frac{SS_{Problem}}{df_{Problem}}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	df_{AxS}	$\frac{SS_{AxS}}{df_{AxS}}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	df_W	$\frac{SS_W}{df_W}$	

Putting the Puzzle Together

F Table for Findings

	SS	df	MS	F
Age	SS_{Age}	**(A – 1)**	$\frac{SS_{Age}}{df_{Age}}$	$\frac{MS_{Age}}{MS_W}$
Problem	$SS_{Problem}$	**(P – 1)**	$\frac{SS_{Problem}}{df_{Problem}}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	**(A - 1) x (P - 1)**	$\frac{SS_{AxS}}{df_{AxS}}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W		$\frac{SS_W}{df_W}$	

DF for the Full Model

Always the (Number of Participants – Number of Estimated Parameters in Full Model)

	Normal Equal Sign Problems	Tricky Equal Sign Problems
8-9 year olds	3.50	0.25
9-10 year olds	3.75	2.00

F Table for Findings

	SS	df	MS	F
Age	SS_{Age}	**(A – 1)**	$\frac{SS_{Age}}{df_{Age}}$	$\frac{MS_{Age}}{MS_W}$
Problem	$SS_{Problem}$	**(P – 1)**	$\frac{SS_{Problem}}{df_{Problem}}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	**(A - 1) x (P - 1)**	$\frac{SS_{AxS}}{df_{AxS}}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	**AB(n -1)**	$\frac{SS_W}{df_W}$	

F Table for Findings

	SS	df	MS	F
Age	SS_{Age}	**1**	$\frac{SS_{Age}}{1}$	$\frac{MS_{Age}}{MS_W}$
Problem	$SS_{Problem}$	**1**	$\frac{SS_{Problem}}{1}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	**1**	$\frac{SS_{AxS}}{1}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	**12**	$\frac{SS_W}{12}$	

How to calculate SS_{Age}

$SS_{Age} = E_R - E_F$

Restricted: $Y_{ijk} = \mu + Problem_k + (AP)_{jk} + e_{ijk}$

Full: $Y_{ijk} = \mu + Age_j + Problem_k + (AP)_{jk} + e_{ijk}$

Calculating E_R for Age

Data	Prediction	Error	Squared Error	
3	2.375	-0.625	0.390625	
4	2.375	-1.625	2.640625	
3	2.375	-0.625	0.390625	
4	2.375	-1.625	2.640625	
0	2.375	2.375	5.640625	
0	2.375	2.375	5.640625	
0	2.375	2.375	5.640625	**E_R = 35.75**
1	2.375	1.375	1.890625	
4	2.375	-1.625	2.640625	
4	2.375	-1.625	2.640625	
4	2.375	-1.625	2.640625	
3	2.375	-0.625	0.390625	
2	2.375	0.375	0.140625	
3	2.375	-0.625	0.390625	
2	2.375	0.375	0.140625	
1	2.375	1.375	1.890625	

Calculating E_F for Age

Data	Prediction	Error	Squared Error	
3	1.875	-1.125	1.265625	
4	1.875	-2.125	4.515625	
3	1.875	-1.125	1.265625	
4	1.875	-2.125	4.515625	
0	1.875	1.875	3.515625	
0	1.875	1.875	3.515625	
0	1.875	1.875	3.515625	$E_R = 31.75$
1	1.875	0.875	0.765625	
4	2.875	-1.125	1.265625	
4	2.875	-1.125	1.265625	
4	2.875	-1.125	1.265625	
3	2.875	-0.125	0.015625	
2	2.875	0.875	0.765625	
3	2.875	-0.125	0.015625	
2	2.875	0.875	0.765625	
1	2.875	1.875	3.515625	

F Table for Findings

	SS	df	MS	F
Age	**Same $E_R - E_F$ as if there was one factor**	1	$\frac{SS_{Age}}{1}$	$\frac{MS_{Age}}{MS_W}$
Problem	**Same $E_R - E_F$ as if there was one factor**	1	$\frac{SS_{Problem}}{1}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	1	$\frac{SS_{AxS}}{1}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	12	$\frac{SS_W}{12}$	

F Table for Findings

	SS	df	MS	F
Age	**4.00**	1	$\frac{SS_{Age}}{1}$	$\frac{MS_{Age}}{MS_W}$
Problem	**25.00**	1	$\frac{SS_{Problem}}{1}$	$\frac{MS_{Problem}}{MS_W}$
Age x Problem	SS_{AxP}	1	$\frac{SS_{AxS}}{1}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	12	$\frac{SS_W}{12}$	

F Table for Findings

	SS	df	MS	F
Age	**4.00**	1	**4.00**	$\frac{4.00}{MS_W}$
Problem	**25.00**	1	**25.00**	$\frac{25.00}{MS_W}$
Age x Problem	SS_{AxP}	1	$\frac{SS_{AxS}}{1}$	$\frac{MS_{AxP}}{MS_W}$
Error / Residuals	SS_W	12	$\frac{SS_W}{12}$	

Figuring Out SS_W (E_F)

One-Way Full Model: $Y_{ijk} = \mu + Age_j + e_{ijk}$

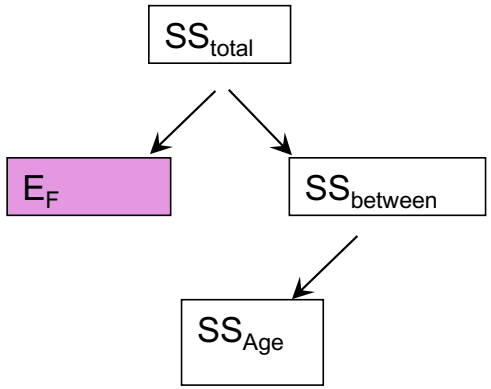
Two-Way Full Model: $Y_{ijk} = \mu + Age_j + Problem_k + (AP)_{jk} + e_{ijk}$

One-Way ANOVA

8-9 year olds	3	4	3	4	0	0	0	1
9-10 year olds	4	4	4	3	2	3	2	1

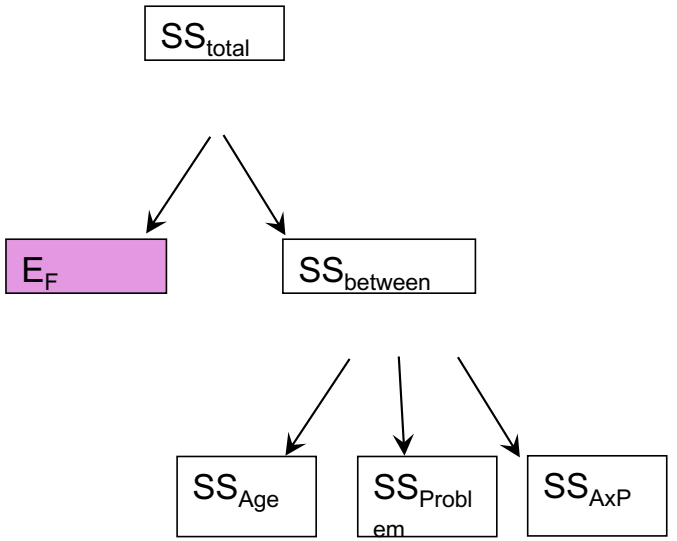
$F = \frac{4.00 / 1}{31.75 / 14}$

$F = 1.764, p = .205$



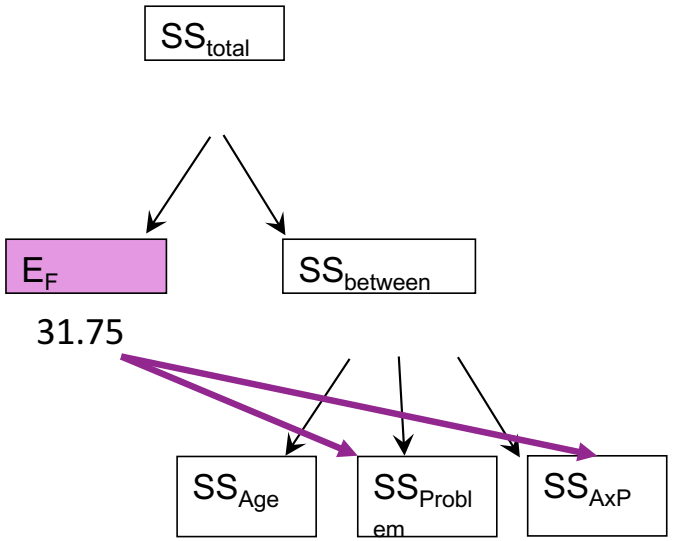
Two-Way ANOVA

	Normal Equal Sign Problems				Tricky Equal Sign Problems			
8-9 year olds	3	4	3	4	0	0	0	1
9-10 year olds	4	4	4	3	2	3	2	1



Two-Way ANOVA

	Normal Equal Sign Problems				Tricky Equal Sign Problems			
8-9 year olds	3	4	3	4	0	0	0	1
9-10 year olds	4	4	4	3	2	3	2	1



F_{Age} for One-Way and Two-Way

One-way: $F(1, 14) = 1.764$, $p = .205$

Two-way: $F(1, 12) = 10.667$, $p = .007$

Taking into account the Problem and AxP factors reduces the denominator

It also reduces the denominator degrees of freedom

Calculating the Interaction

SS_{AxP}

H_{O3}: The differences between age groups are the same regardless of type of problem

H_{O3}: Differential performance on problem types is independent of age group

Restricted model for testing the Interaction between Age and Problem…

- $Y_{ijk} = \mu + Age_j + Problem_k + e_{ijk}$

What does the restricted model say?

$Y_{ijk} = \mu + Age_j + Problem_k + e_{ijk}$

Your Score = Grand Mean + Some Age Effect + Some Problem Effect + Error

Predicted Score =

Grand Mean +

(Difference between Age Group and Grand Mean) +

(Difference between Problem Group and Grand Mean)

What this works out to…

Predicted Score = Grand Mean + ($Mean_{Row}$ – Grand Mean) + ($Mean_{Column}$ – Grand Mean)

Predicted Score = Grand Mean + $Mean_{Row}$ – Grand Mean + $Mean_{Column}$ – Grand Mean

Predicted Score = $Mean_{Row}$ + $Mean_{Column}$ – Grand Mean

Difference between Restricted and Full Models = Squared deviations for each cell from this predicted score

	Normal Equal Sign Problems				Tricky Equal Sign Problems				
8-9 year olds	3	4	3	4	0	0	0	1	μ_8
9-10 year olds	4	4	4	3	2	3	2	1	μ_9
	μ_{Normal}				μ_{Tricky}				

	Normal Equal Sign Problems	Tricky Equal Sign Problems	
8-9 year olds	3.50	0.25	**1.875**
9-10 year olds	3.75	2.00	**2.875**
	3.625	**1.125**	

	Normal Equal Sign Problems	Tricky Equal Sign Problems	
8-9 year olds	3.50	0.25	**1.875**
9-10 year olds	3.75	2.00	**2.875**
	3.625	**1.125**	**2.375**

Cell Mean	Subtract Row Mean	Subtract Column Mean	Add Grand Mean	Total
3.50	1.875	3.625	2.375	0.375
0.25	1.875	1.125	2.375	-0.375
3.75	2.875	1.125	2.375	-0.375
2.00	2.875	3.625	2.375	0.375

Sum of Deviations equal 0!!!

Cell Mean	Subtract Row Mean	Subtract Column Mean	Add Grand Mean	Total	Squared
3.50	1.875	3.625	2.375	0.375	.140625
0.25	1.875	1.125	2.375	-0.375	.140625
3.75	2.875	1.125	2.375	-0.375	.140625
2.00	2.875	3.625	2.375	0.375	.140625

Cell Mean	Subtract Row Mean	Subtract Column Mean	Add Grand Mean	Total	Squared
3.50	1.875	3.625	2.375	0.375	.140625
0.25	1.875	1.125	2.375	-0.375	.140625
3.75	2.875	1.125	2.375	-0.375	.140625
2.00	2.875	3.625	2.375	0.375	.140625

Sum of Squared Deviations = 0.5625

Multiply by n = 2.25

What we just did

Found the means for each cell in the 2 x 2 ANOVA

Took away the effect for the two factors

Added back the grand mean

The extent to which each cell differs from its respective column and row is the difference between a model without an interaction (Restricted) and a model with one (Full)

F Table for Findings

	SS	df	MS	F
Age	4.00	1	4.00	$\frac{4.00}{MS_W}$
Problem	25.00	1	25.00	$\frac{25.00}{MS_W}$
Age x Problem	**2.25**	1	**2.25**	$\frac{2.25}{MS_W}$
Error / Residuals	SS_W	12	$\frac{SS_W}{12}$	

SS_{Within} or $E_{Full\ Model}$

Full Model

The full model allows for the Age effect, the Problem effect, and the Interaction

So from the Full Model perspective, the best guess for any observation is…

Calculating SS_W or E_F

Score	Predicted Value	Deviation	Squared Deviation	
3	3.5	0.5	0.25	
4	3.5	-0.5	0.25	
3	3.5	0.5	0.25	
4	3.5	-0.5	0.25	
0	0.25	0.25	0.0625	
0	0.25	0.25	0.0625	$E_F = 4.5$
0	0.25	0.25	0.0625	
1	0.25	-0.75	0.5625	
4	3.75	-0.25	0.0625	
4	3.75	-0.25	0.0625	
4	3.75	-0.25	0.0625	
3	3.75	0.75	0.5625	
2	2	0	0	
3	2	-1	1	
2	2	0	0	
1	2	1	1	

F Table for Findings

	SS	df	MS	F
Age	4.00	1	4.00	$\frac{4.00}{MS_W}$
Problem	25.00	1	25.00	$\frac{25.00}{MS_W}$
Age x Problem	2.25	1	2.25	$\frac{2.25}{MS_W}$
Error / Residuals	4.50	12	$\frac{4.50}{12} = 0.375$	

Final Results

F Table for Findings

	SS	df	MS	F
Age	4.00	1	4.00	$\frac{4.00}{0.375}$
Problem	25.00	1	25.00	$\frac{25.00}{0.375}$
Age x Problem	2.25	1	2.25	$\frac{2.25}{0.375}$
Error / Residuals	4.50	12	0.375	

F Table for Findings

	SS	df	MS	F
Age	4.00	1	4.00	10.667
Problem	25.00	1	25.00	66.667
Age x Problem	2.25	1	2.25	6.00
Error / Residuals	4.50	12	0.375	

What do we conclude?

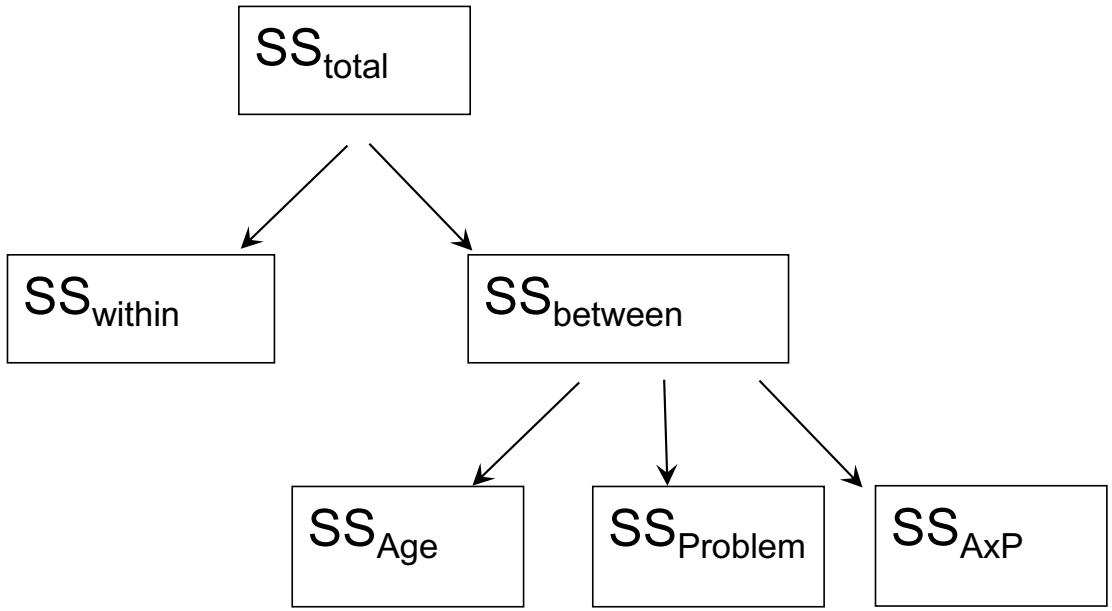
Three rejected null hypotheses

The age groups are not equal

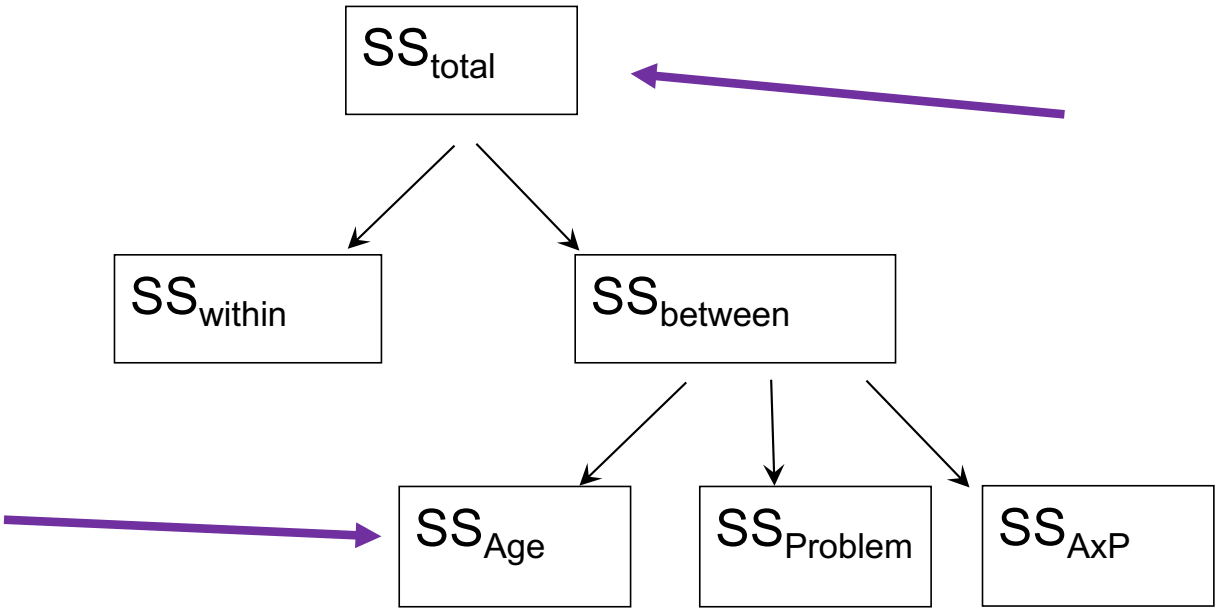
The problem groups are not equal

And the difference between age groups differs by problem groups

Effect Sizes



η^2 for Age



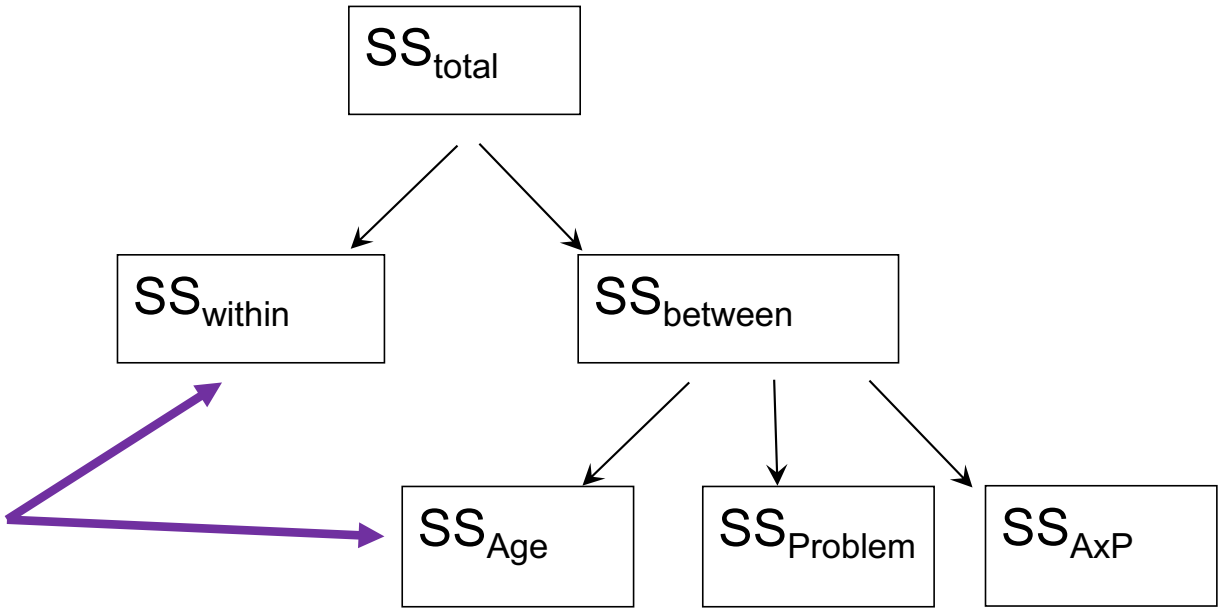
η^2 for Age

	SS
Age	4.00
Problem	25.00
Age x Problem	2.25
Error / Residuals	4.50

η^2 for Age

	SS
Age	4.00
Problem	25.00
Age x Problem	2.25
Error / Residuals	4.50

Partial η^2 for Age



Partial η^2 for Age

	SS
Age	4.00
Problem	25.00
Age x Problem	2.25
Error / Residuals	4.50

Partial η^2 for Age

	SS
Age	4.00
Problem	25.00
Age x Problem	2.25
Error / Residuals	4.50

The calculation will differ slightly when…

Sample sizes are unequal for the cells

There are more than two factors (and thus multiple interactions)