

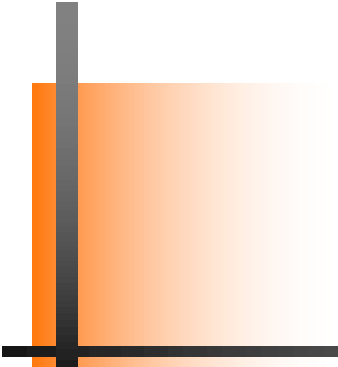


IN5060

Performance in distributed systems

User studies (cntd)





Does blur hide asynchrony?

study by Ragnhild Eg (Simula) et al., 2011

Perception of synchrony

Sensitivity for perceptual synchrony is subjective and depends on the content

Spoken sentences (Grant et al., 2003)

- Discrimination thresholds: ≈ 50 ms audio lead, ≈ 200 ms audio lag

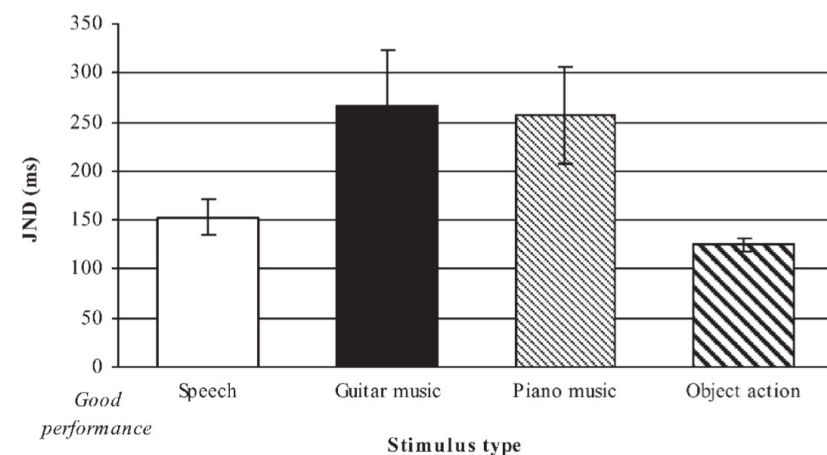
Hitting table with wand (Levitin et al., 2000)

- Synchrony thresholds set to 75 %:
41 ms Ahead to 45 ms Alag

Music, baseball, speech

(Vatakis & Spence, 2006)

- Temporal order judgements
(audio/video first)



Stimuli

3 content types

Chess game



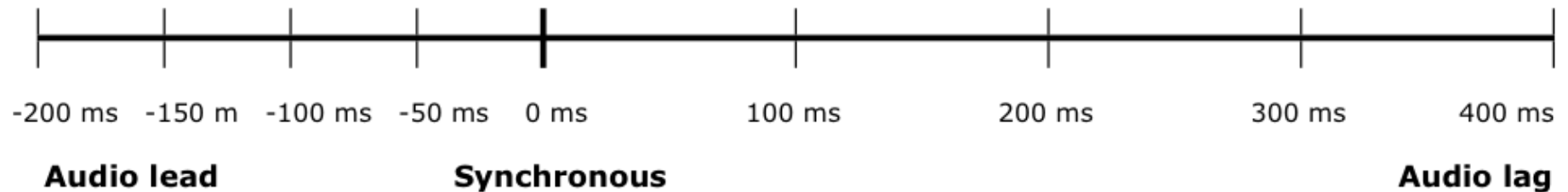
News broadcast



Drummer

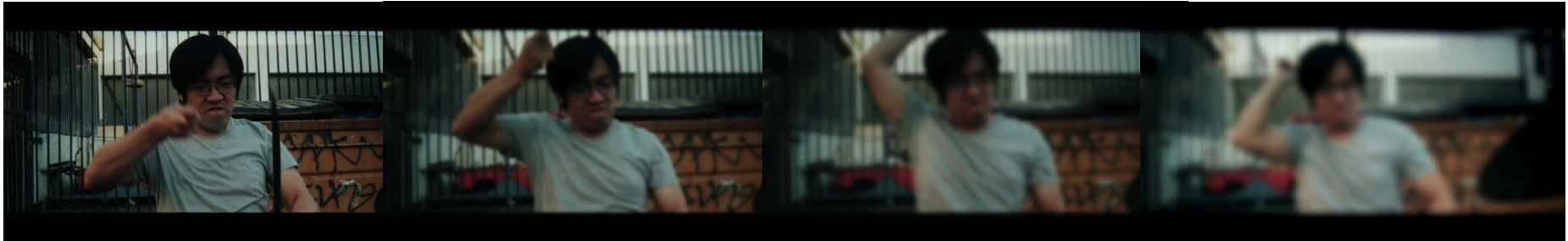


9 asynchrony levels



Stimuli

Visual distortion, 4 levels, Gaussian blur filter



Undistorted

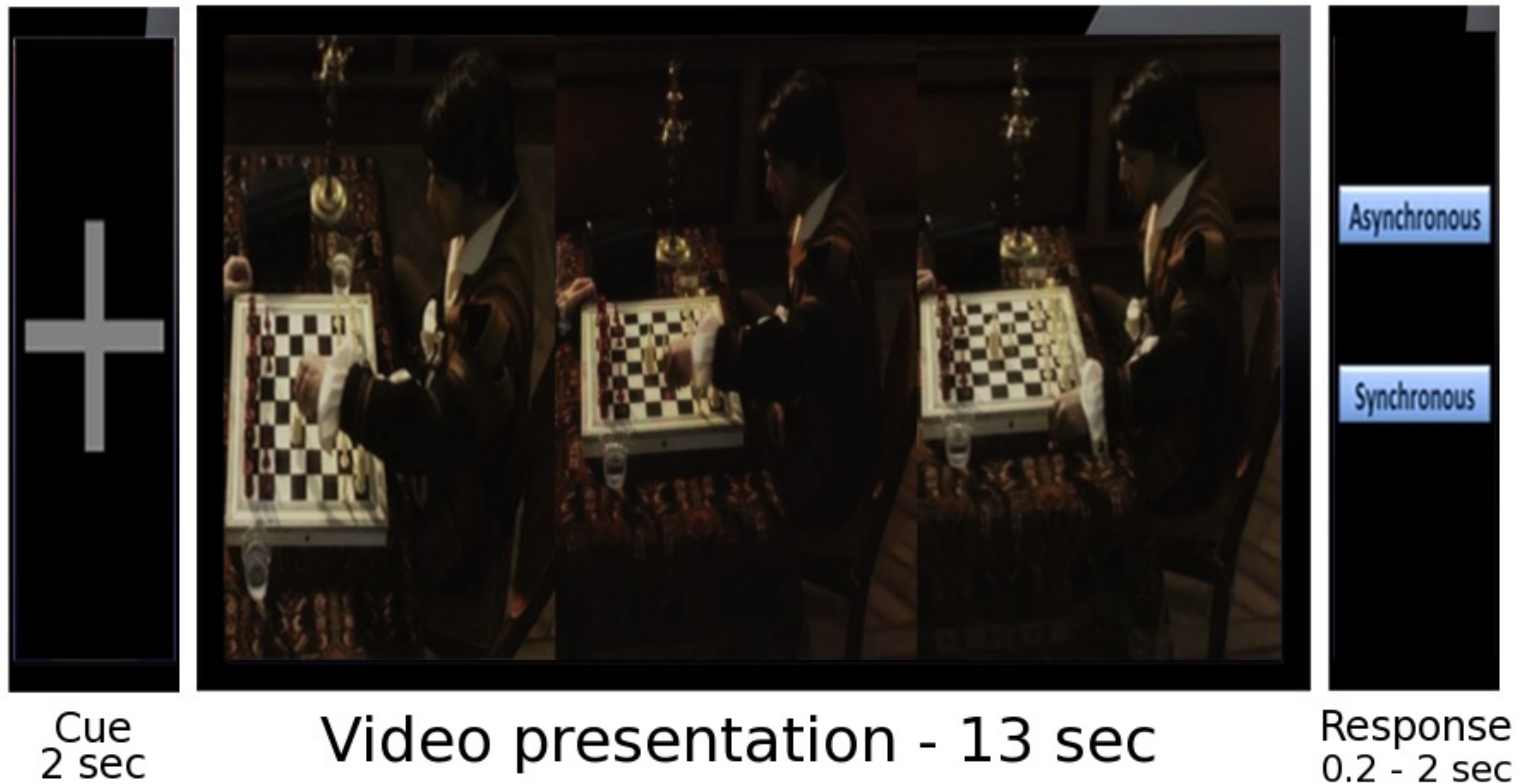
Blur 2x2 pixels

Blur 4x4 pixels

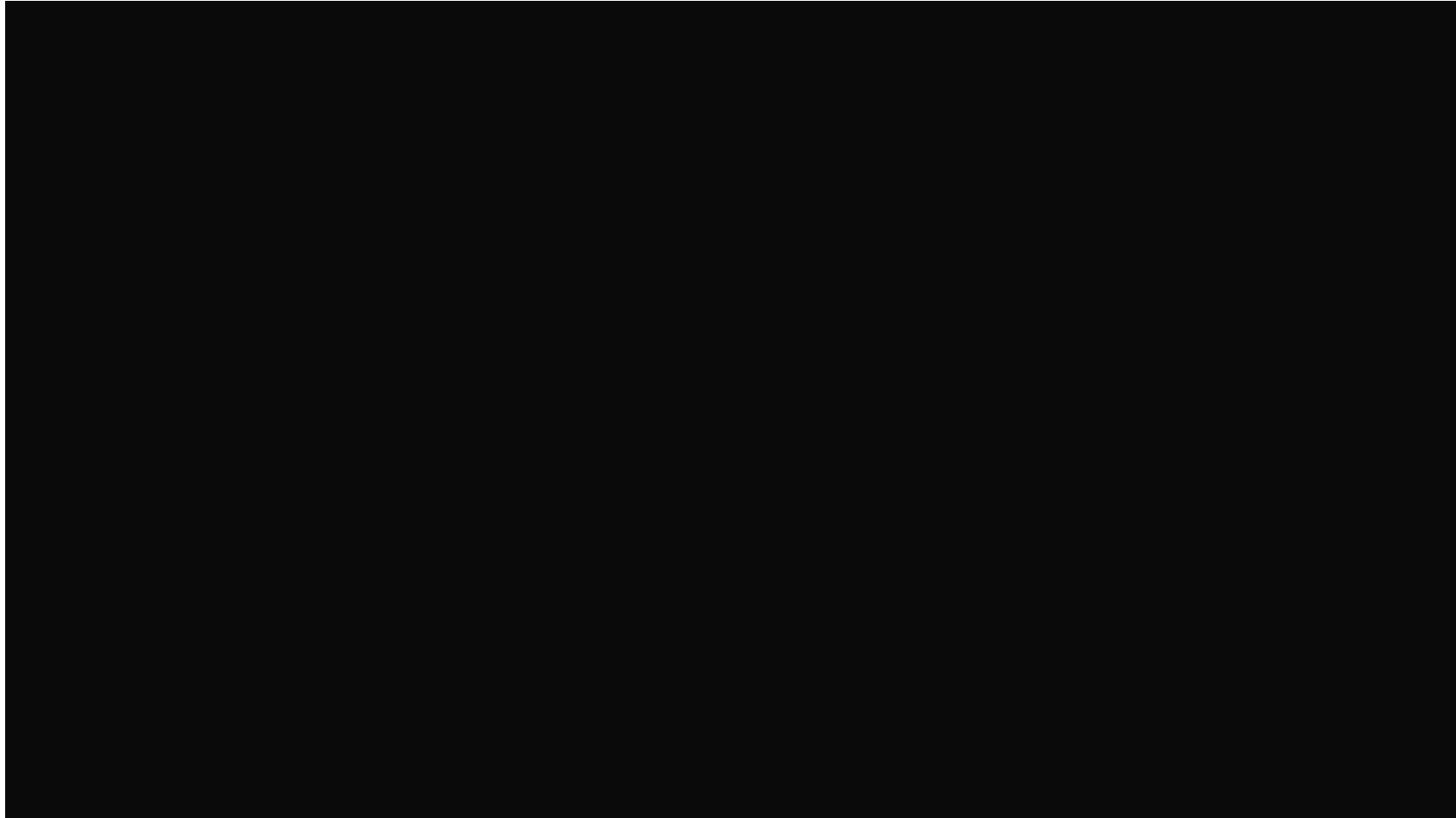
Blur 6x6 pixels

Procedure

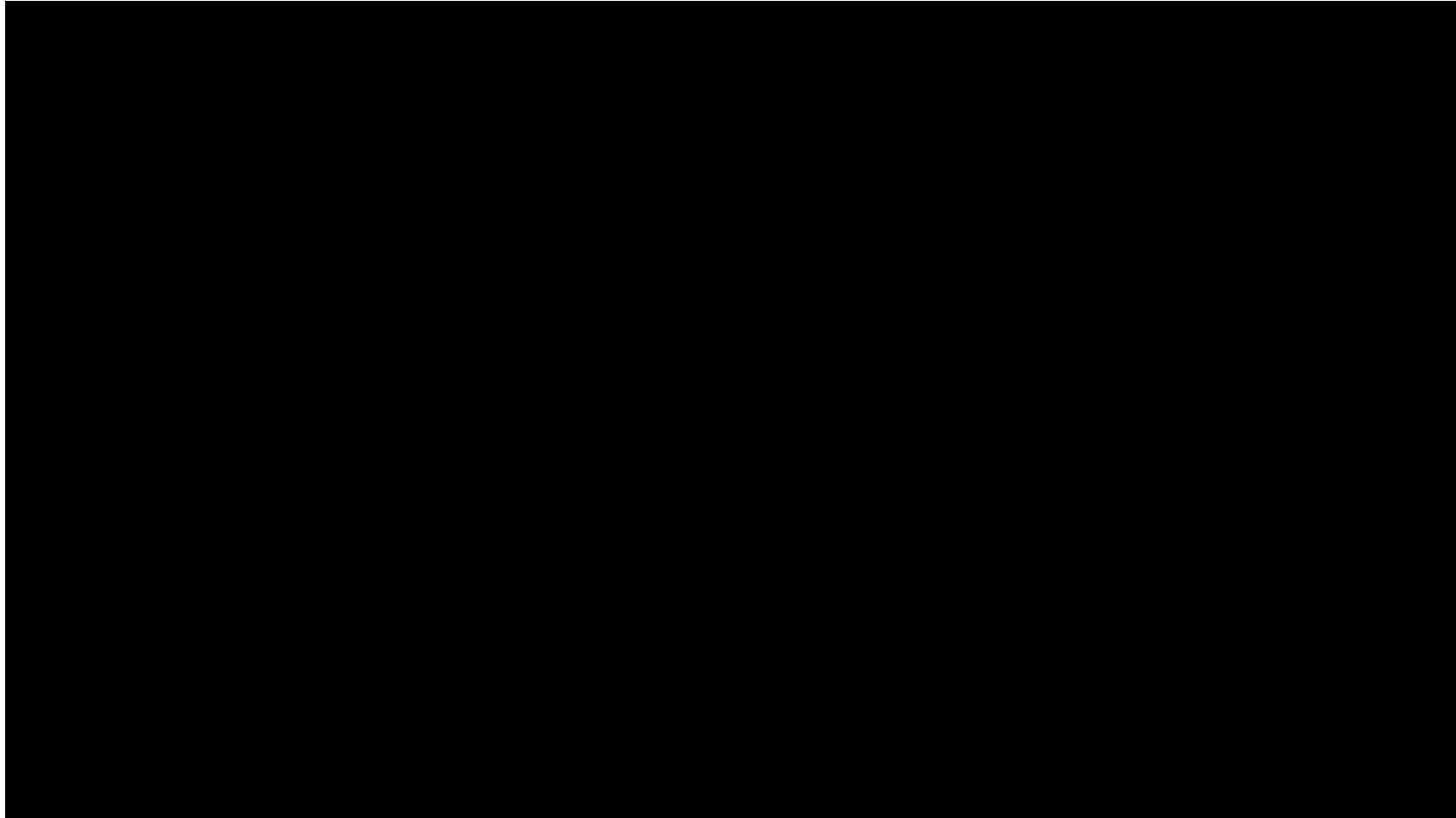
- Carried out at the Speech Lab, NTNU



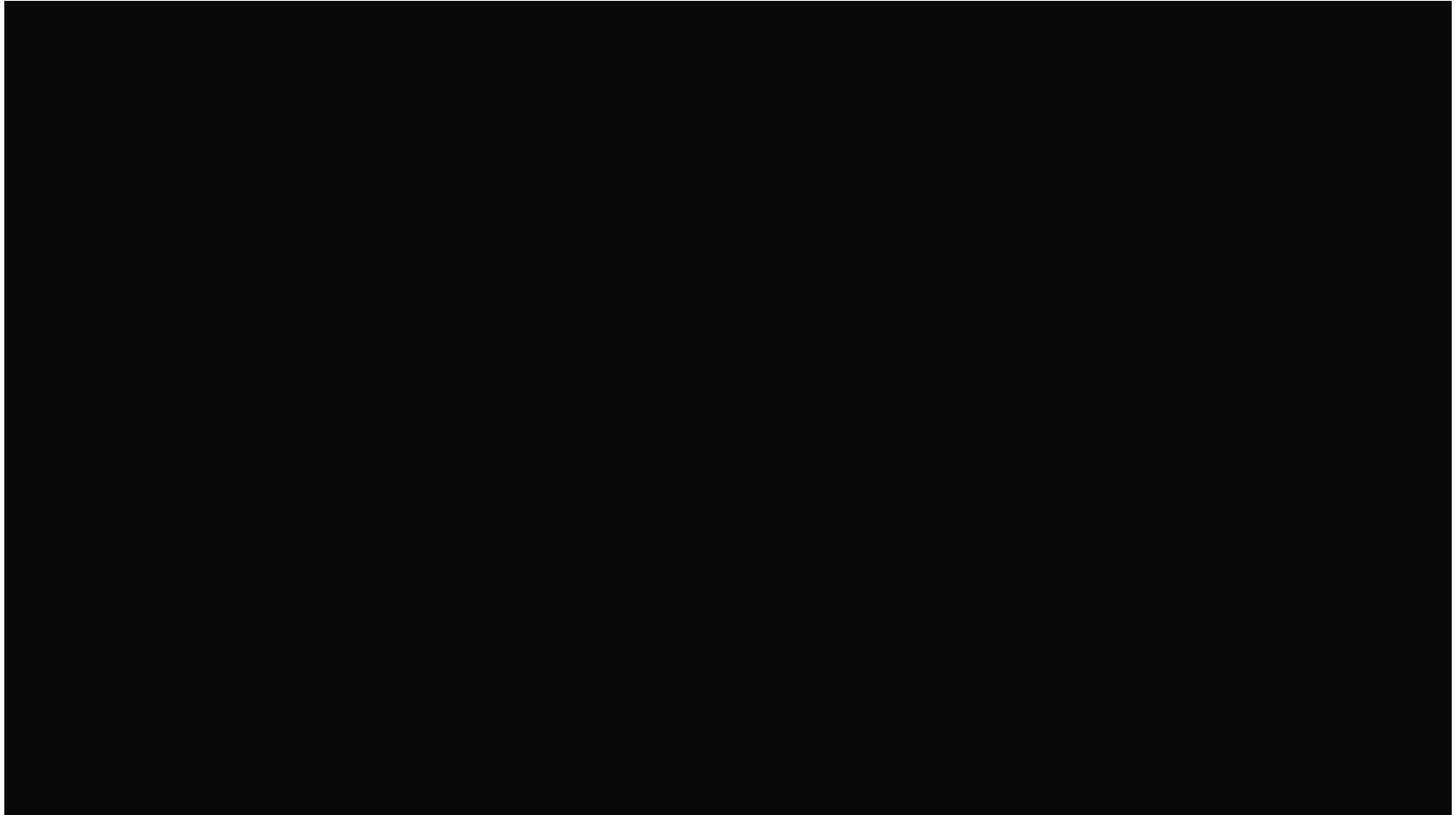
Chess content - 200 ms audio lead



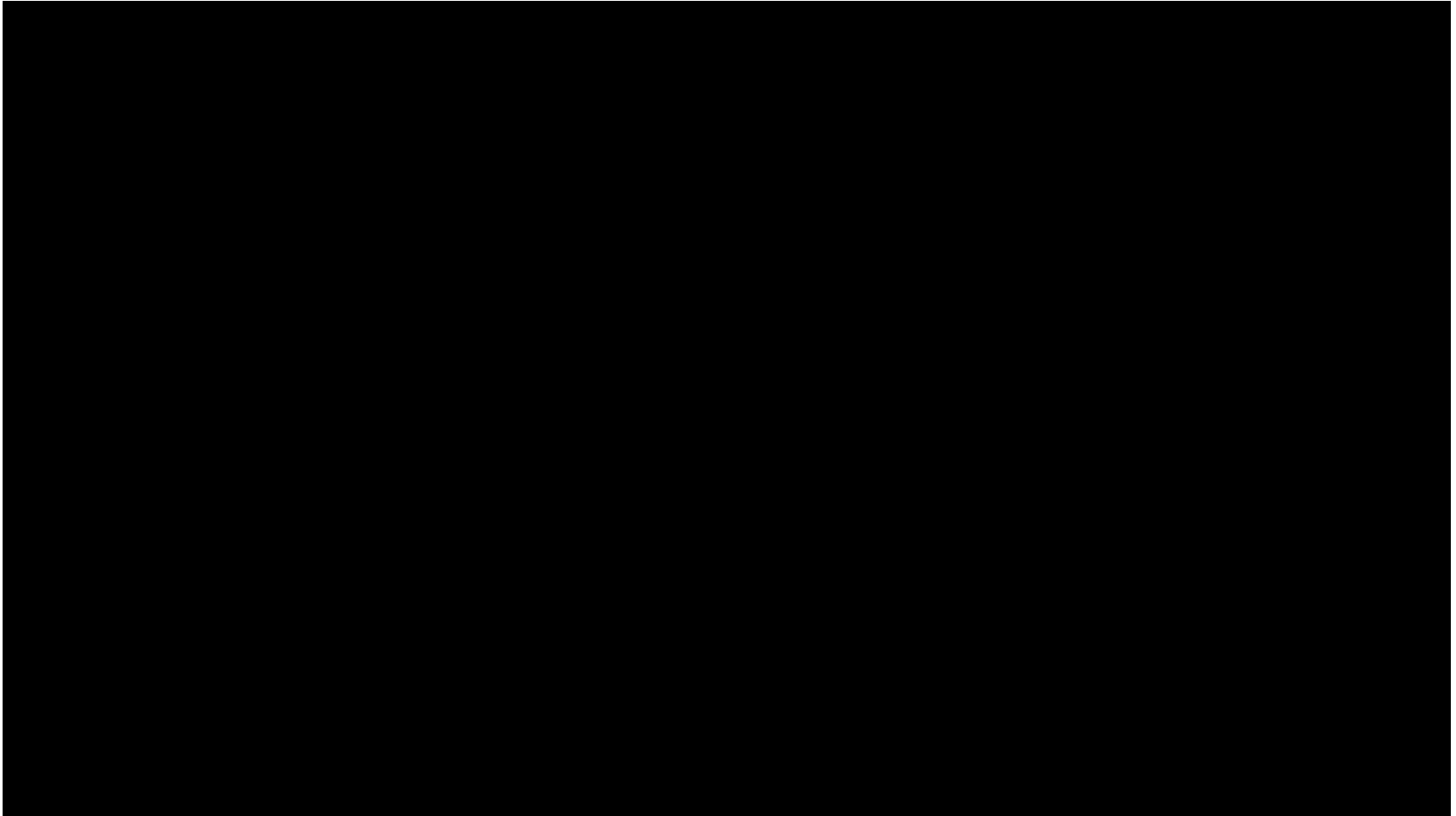
Chess content - 200 ms audio lag, blurred



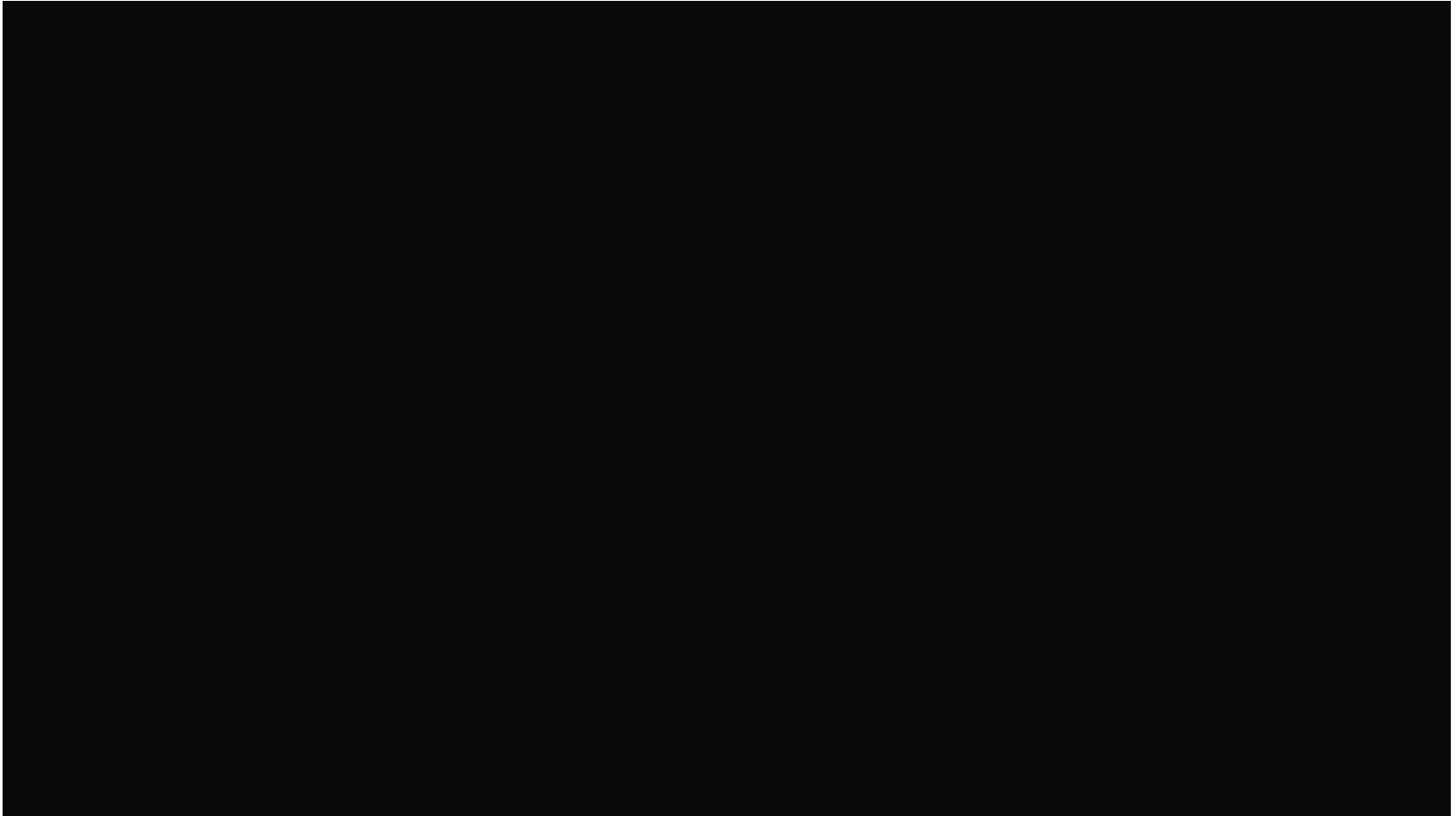
News - 300 ms audio lag, blurred



Drums - 100 ms audio lag, blurred



Drums - 150 ms audio lead, slightly blurred





Audio streaming from PPT in Zoom is really bad.

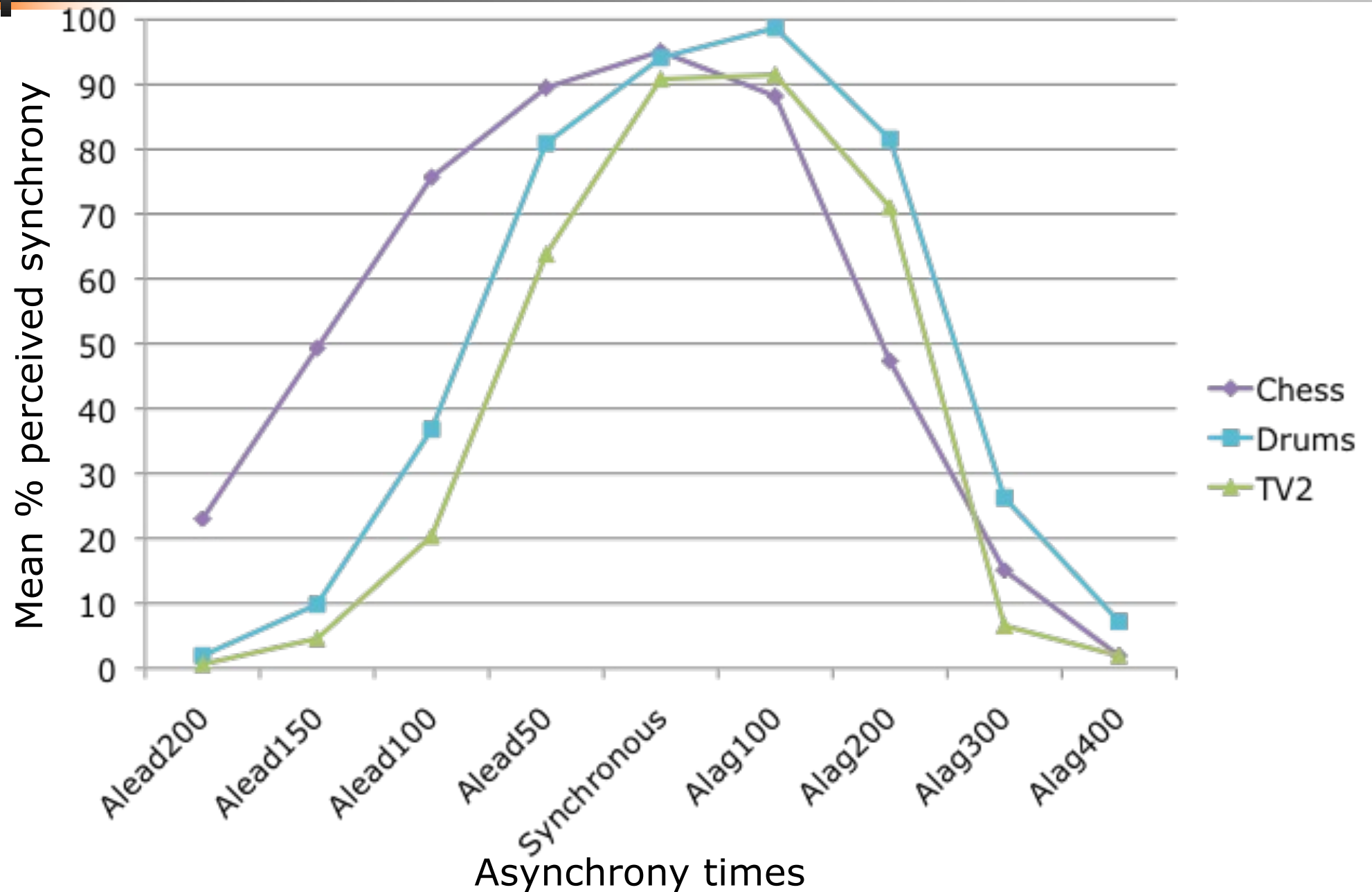
See the examples here:

<https://drive.google.com/drive/folders/1hxXFdh5xCeN1pMril2kZzNwPC3ZmuL-u?usp=sharing>

Design & Analysis

- 2 independent studies
- Full-factorial design
- 2 repetitions of each condition
- Binomial responses converted to percentages
- Repeated-measures ANOVAs
- Separate analyses for:
 - Audio lag and audio lead (different scales)
 - Content types (different response patterns)

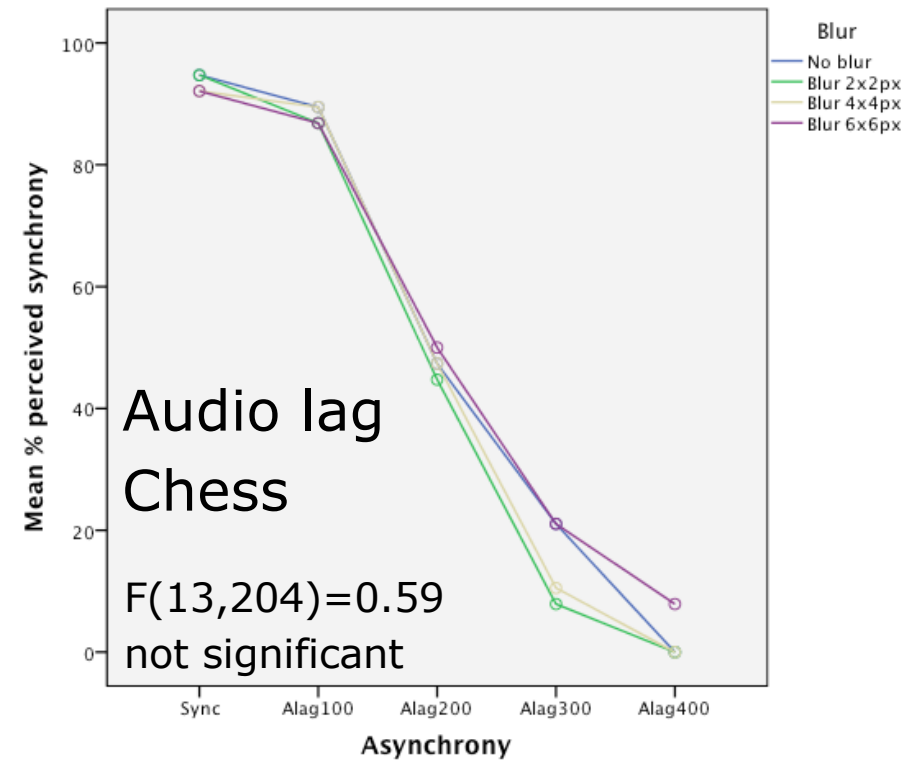
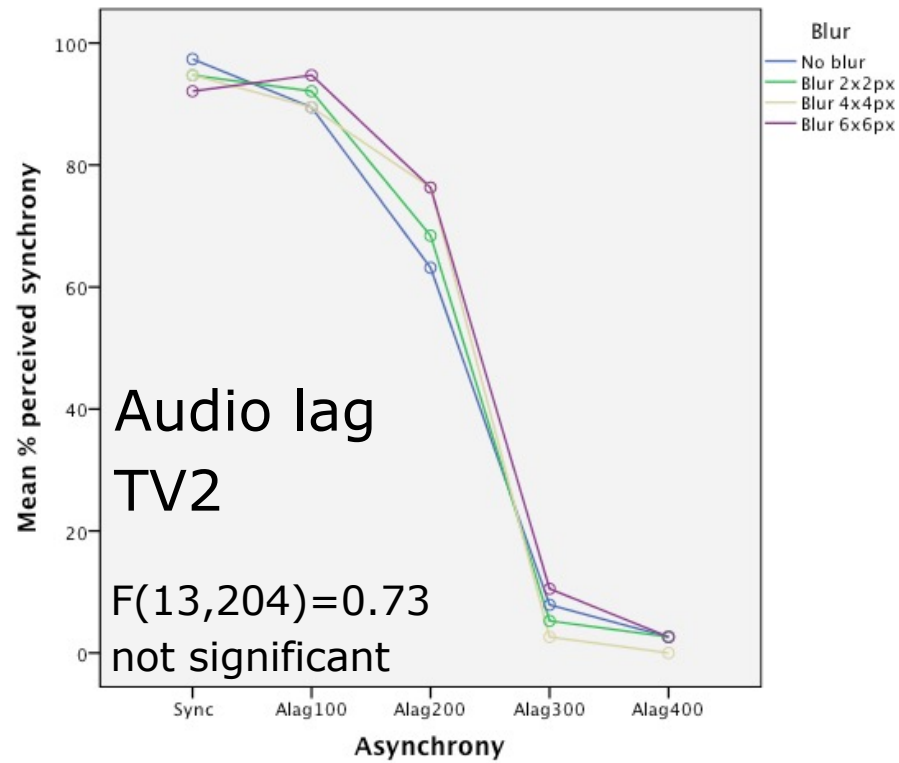
Mean perceived synchrony, averaged across **blur** levels



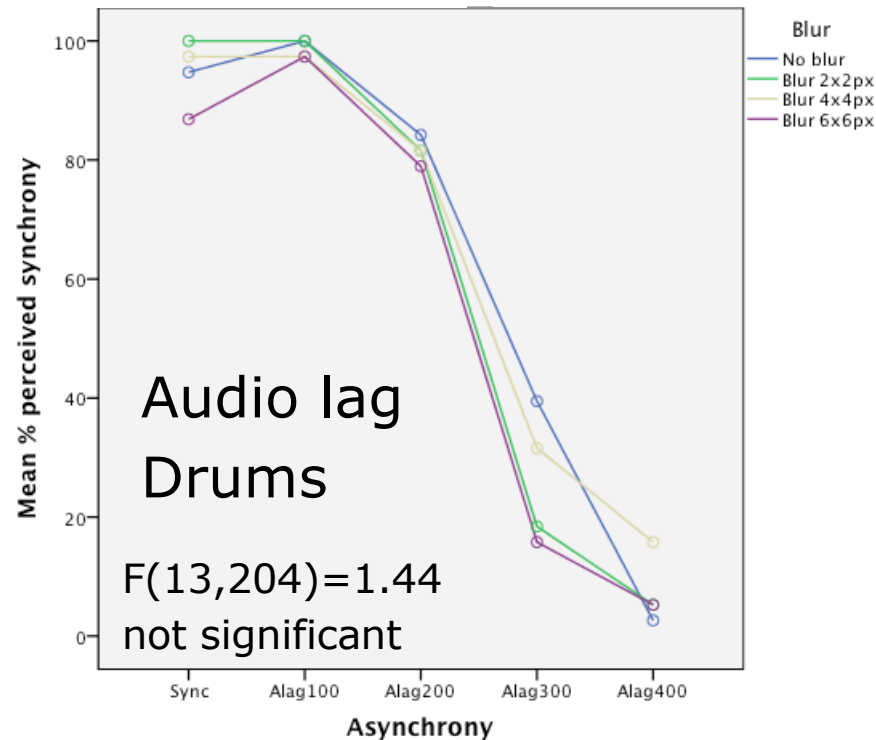
Assessment of relevance

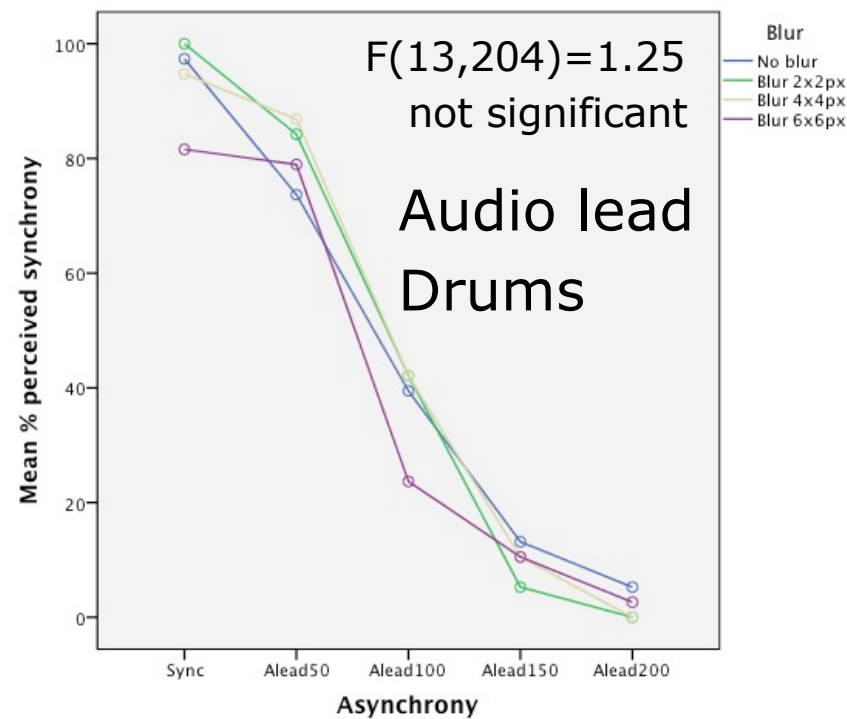
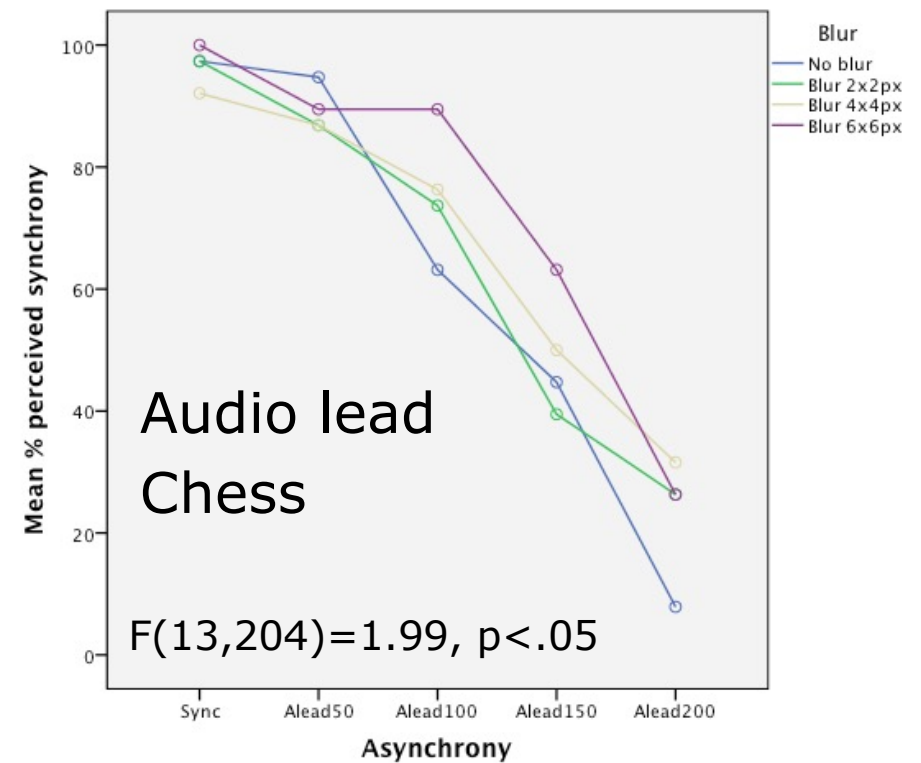
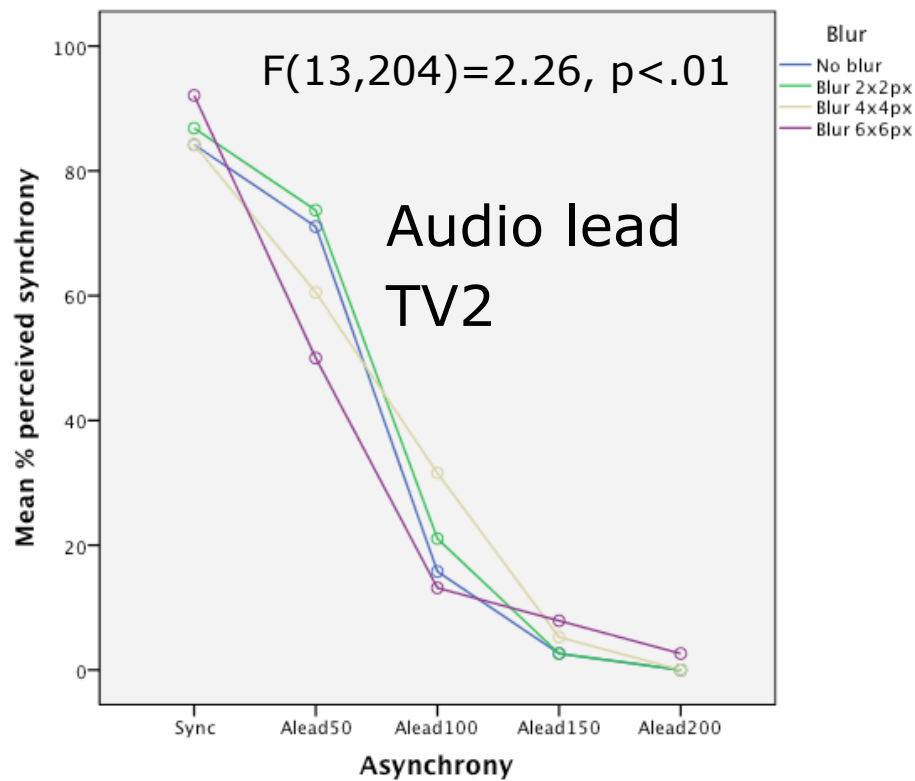
	Visual distortion	
	Content	F-statistics
Audio lag	Chess	$F(4,85)=88.79, p<.001$
	TV2	$F(4,85)=232.54, p<.001$
	Drums	$F(4,85)=197.57, p<.001$
Audio lead	Chess	$F(4,85)=71.77, p<.001$
	TV2	$F(4,85)=100.26, p<.001$
	Drums	$F(4,85)=126.31, p<.001$

5 settings
18 participants



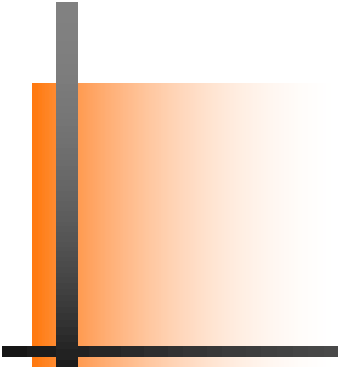
Blur distortion





Blur
distortion

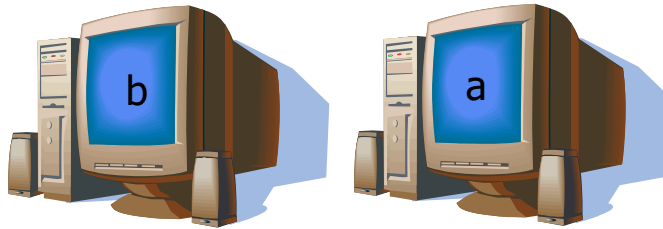




ANOVA

Analysis of Variance

A-B Comparison



Candidate $[i]$	Audio lag $[b_i]$	Audio lead $[a_i]$	Difference $[d_i = b_i - a_i]$
1	5	6	-1
2	3	8	-5
3	14	10	4
4	10	15	-5
5	8	11	-3
6	7	3	4

Mean of differences $\bar{d} = -1$, Standard deviation $\sigma_d = 4.15$

A-B Comparison

Mean of differences $\bar{d} = -1$

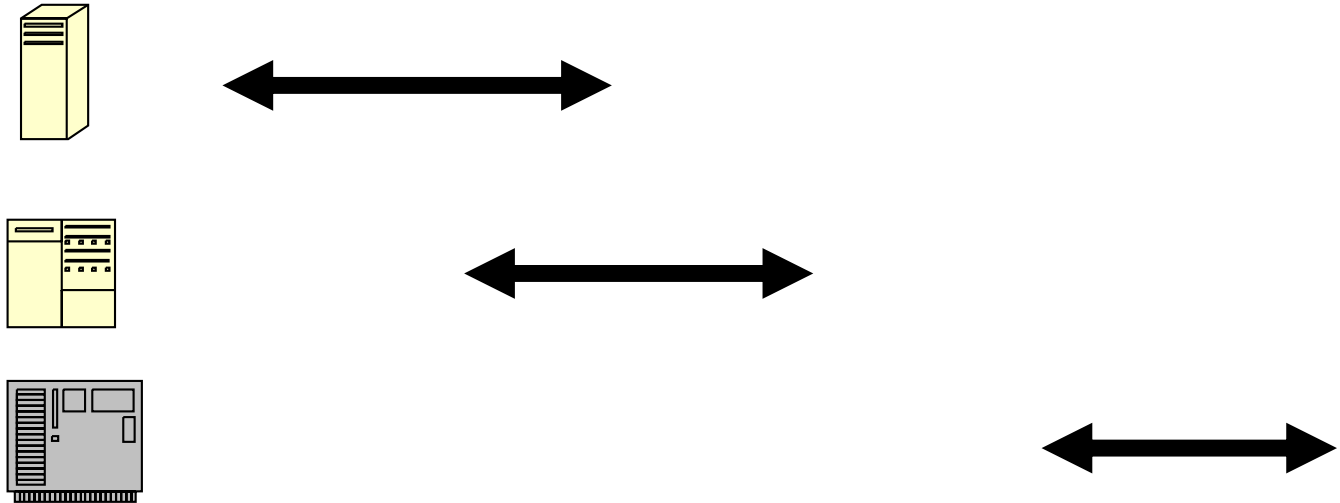
Standard deviation $\sigma_d = 4.15$

- From mean of differences, appears that audio lag reduced performance
- However, standard deviation is large
- Is the variation between the two alternatives greater than the variation (error) in the measurements?
- Confidence intervals can work, but what if there are more than two alternatives?

Comparing more than two alternatives

- Naïve approach

- Compare confidence intervals



- Need to do for all *pairs*. This grows very quickly.
- Example: 7 alternatives would require 21 pair-wise comparisons
 - possible combinations: $\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1}$
 - for our case: $\binom{7}{2} = \frac{7*6}{2*1} = \frac{42}{2} = 21$
- Would not be surprising to find 1 pair differed [at 95%]

ANOVA – Analysis of Variance

- Partitioning *variation* (not variance) into the part that can be explained and the part that cannot be explained
- Separates total variation observed in a set of measurements into:
 1. Variation within one system
due to uncontrolled measurement errors
 2. Variation between systems
due to real differences + random error
- Is variation (2) statistically greater than variation (1)?

ANOVA – Analysis of Variance

- Make n measurements of k alternatives
- y_{ij} = i -th measurement on j -th alternative
- Assumes errors are
 - independent
 - normally distributed
- In user studies, each **measurement** is the set of responses by one **participant**

All Measurements for All Alternatives

	Alternatives					
Measure- ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}

Overall Mean

Average of all measurements made of all alternatives:

$$\bar{y} = \frac{\sum_{j=1}^k \sum_{i=1}^n y_{ij}}{kn}$$

	Alternatives					
Measure- ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}

Column Means

Column means are average values of all measurements within a single alternative

$$\overline{y_{.j}} = \frac{\sum_{i=1}^n y_{ij}}{n}$$

- average performance of a single alternative

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{1k}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$y_{.1}$	$y_{.2}$	...	$y_{.j}$	...	$y_{.k}$

Effect = Deviation From Overall Mean

- α_j : effect of alternative j = deviation of column mean from overall mean: $\alpha_j = \bar{y}_{.j} - \bar{y}$

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$\bar{y}_{.1}$	$\bar{y}_{.2}$	...	$\bar{y}_{.j}$	...	$\bar{y}_{.k}$
Effect	α_1	α_2	...	α_j	...	α_k

Error = Deviation From Column Mean

- e_{ij} : error of each measurement = deviation from column mean: $e_{ij} = y_{ij} - \bar{y}_{.j}$

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$\bar{y}_{.1}$	$\bar{y}_{.2}$	...	$\bar{y}_{.j}$	...	$\bar{y}_{.k}$



Effects and Errors

- **Effect** is distance of column mean from overall mean
 - Horizontally across alternatives
- **Error** is distance of sample from column mean
 - Vertically within one alternative
 - Error across alternatives, too
- Note that neither Effect nor Error are absolute values, they can be positive or negative
- Individual measurements are then:

$$y_{ij} = \bar{y} + \alpha_j + e_{ij}$$

Sum of Squares of Differences

- SST = differences between each measurement and overall mean

$$SST = \sum_{j=1}^k \sum_{i=1}^n (y_{ij} - \bar{y})^2$$

- SSA = variation due to effects of alternatives

$$SSA = n \sum_{j=1}^k \alpha_j^2 = n \sum_{j=1}^k (\bar{y}_{.j} - \bar{y})^2$$

- SSE = variation due to errors in measurements

$$SSE = \sum_{j=1}^k \sum_{i=1}^n e_{ij}^2 = \sum_{j=1}^k \sum_{i=1}^n (y_{ij} - \bar{y}_{.j})^2$$

- $SSE = SST - SSA \Leftrightarrow SST = SSE + SSA$

ANOVA

Separates variation in measured values into:

1. variation due to effects of alternatives
 - SSA – variation across column averages
2. variation due to errors
 - SSE – variation within a single column

If differences among alternatives are due to real differences:

→ SSA statistically greater than SSE

Comparing SSE and SSA

- Simple approach

- $\frac{SSA}{SST}$ = fraction of total variation explained by differences among alternatives

- $\frac{SSE}{SST} = \frac{SST - SSA}{SST}$ = fraction of total variation due to experimental error

- But is it statistically significant?

Comparing SSE and SSA

- Is it statistically significant?
- variance = mean square values
= total variation / degrees of freedom
$$\sigma_x^2 = \frac{SSx}{df(SSx)}$$
- df(SSx):
 - degrees of freedom
 - this is the number of independent terms in sum

Degrees of Freedom for Effects

$df(SSA) = k - 1$, since k alternatives

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$y_{.1}$	$y_{.2}$	...	$y_{.j}$	...	$y_{.k}$
Effect	α_1	α_2	...	α_j	...	α_k



Degrees of Freedom for Errors

$df(SSE) = k \cdot (n - 1)$, since k alternatives, each with $(n - 1)$ degrees of freedom

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$y_{.1}$	$y_{.2}$	...	$y_{.j}$	...	$y_{.k}$
Effect	α_1	α_2	...	α_j	...	α_k

Degrees of Freedom for Total

$df(SST) = df(SSA) + df(SSE) = k \cdot n - 1$, since we consider all kn alternatives as independent experiments, thus fixing only 1 pair

	Alternatives					
Measure-ments	1	2	...	j	...	k
1	y_{11}	y_{12}	...	y_{1j}	...	y_{k1}
2	y_{21}	y_{22}	...	y_{2j}	...	y_{2k}
...	...	...	...	...	...	...
i	y_{i1}	y_{i2}	...	y_{ij}	...	y_{ik}
...	...	...	...	...	...	...
n	y_{n1}	y_{n2}	...	y_{nj}	...	y_{nk}
Column mean	$y_{.1}$	$y_{.2}$	...	$y_{.j}$	...	$y_{.k}$
Effect	α_1	α_2	...	α_j	...	α_k

Note: $k(n - 1) + k - 1 = kn - 1$

Variances from Sum of Squares (Mean Square Value)

Variation between sample means

$$\sigma_a^2 = \frac{SSA}{k - 1}$$

in user studies, this is the variance of the average scores for the different alternatives

Variation within the samples

$$\sigma_e^2 = \frac{SSE}{k(n - 1)}$$

in user studies, this is the variance of the score differences between participants

Comparing Variances

Use F-test to compare ratio of variances

- an F-test is used to test if the standard deviations of two populations are equal

- $F = \frac{\sigma_a^2}{\sigma_e^2}$

- $F_{[1-\alpha; df(num), df(denum)]} = F_{[1-\alpha, k-1, k(n-1)]} =$
tabulated critical values

- table for $p = 0.001$ at:

http://www.tutor-homework.com/statistics_tables/f-table-0.001.html

if $F_{computed} > F_{table}$ for a given α

→ we have $(1 - \alpha)100\%$ **confidence** that
variation due to **actual differences** in alternatives, SSA,
is **statistically greater than** variation due to **errors**, SSE

Comparing Variances: α

- Probability that x is in the interval $[c_1, c_2]$
 - formally written:

$$P(c_1 \leq x \leq c_2) = 1 - \alpha$$

- $[c_1, c_2]$ *confidence interval*
- α *significance level*
- $100(1 - \alpha)$ *confidence level*

typical confidence levels are 90%, 95%, 99%

significance level $\alpha = 0.1 \equiv$ confidence level 90%

significance level $\alpha = 0.05 \equiv$ confidence level 95%

significance level $\alpha = 0.01 \equiv$ confidence level 99%

Comparing Variances

Table of F-statistics P=0.001

[t-statistics](#)

F-statistics with other P-values: [P=0.05](#) | [P=0.01](#)

[Chi-square statistics](#)

df(num)

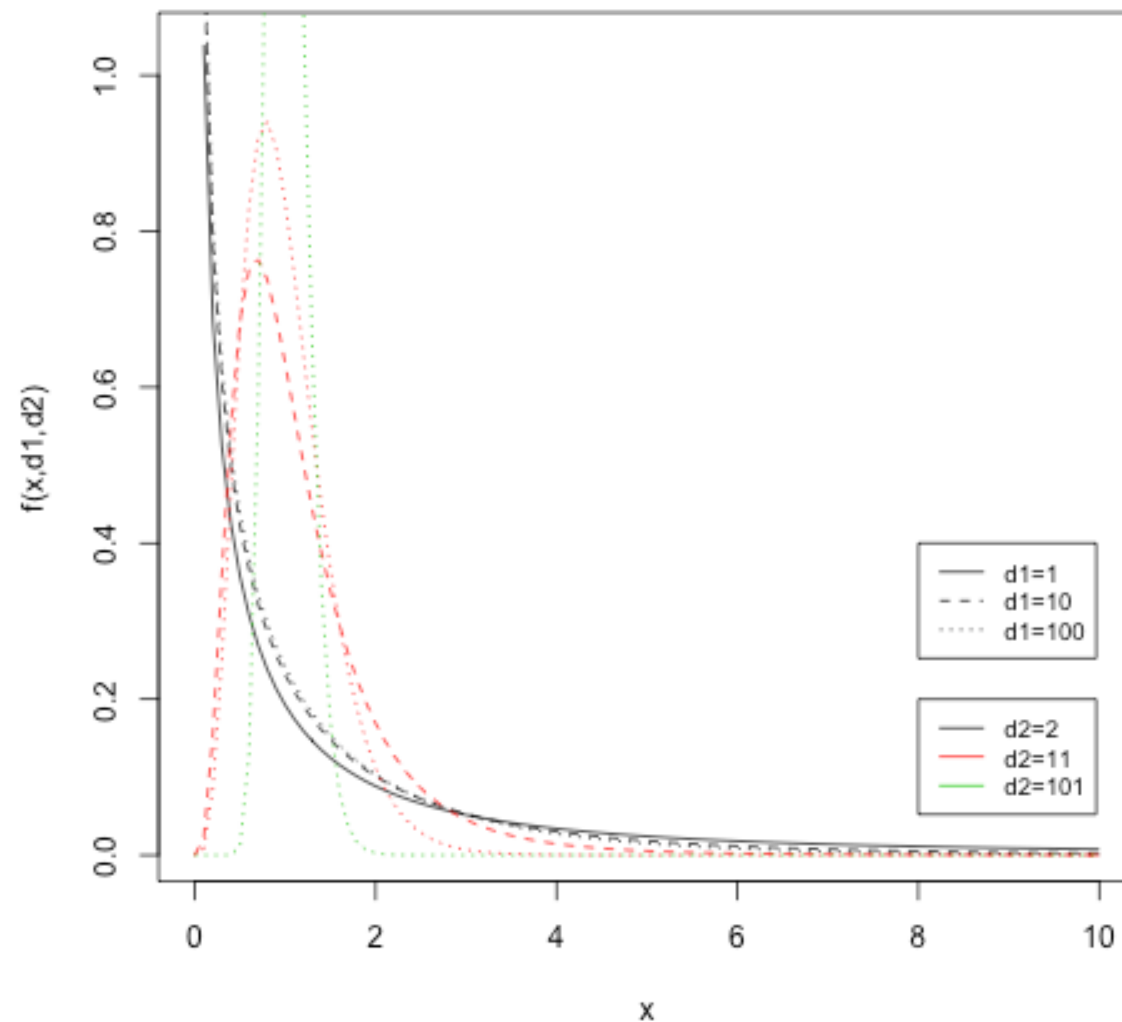
df(denum)	df2\df1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	22	24	26	28	30
3		167.03	148.50	141.11	137.10	134.58	132.85	131.59	130.62	129.86	129.25	128.74	128.32	127.96	127.65	127.38	127.14	126.93	126.74	126.57	126.42	126.16	125.94	125.75	125.59	125.45
4		74.14	61.25	56.18	53.44	51.71	50.53	49.66	49.00	48.48	48.05	47.71	47.41	47.16	46.95	46.76	46.60	46.45	46.32	46.21	46.10	45.92	45.77	45.64	45.53	45.43
5		47.18	37.12	33.20	31.09	29.75	28.84	28.16	27.65	27.25	26.92	26.65	26.42	26.22	26.06	25.91	25.78	25.67	25.57	25.48	25.40	25.25	25.13	25.03	24.95	24.87
6		35.51	27.00	23.70	21.92	20.80	20.03	19.46	19.03	18.69	18.41	18.18	17.99	17.83	17.68	17.56	17.45	17.35	17.27	17.19	17.12	17.00	16.90	16.81	16.74	16.67
7		29.25	21.69	18.77	17.20	16.21	15.52	15.02	14.63	14.33	14.08	13.88	13.71	13.56	13.43	13.32	13.23	13.14	13.06	12.99	12.93	12.82	12.73	12.66	12.59	12.53
8		25.42	18.49	15.83	14.39	13.49	12.86	12.40	12.05	11.77	11.54	11.35	11.20	11.06	10.94	10.84	10.75	10.67	10.60	10.54	10.48	10.38	10.30	10.22	10.16	10.11
9		22.86	16.39	13.90	12.56	11.71	11.13	10.70	10.37	10.11	9.89	9.72	9.57	9.44	9.33	9.24	9.15	9.08	9.01	8.95	8.90	8.80	8.72	8.66	8.60	8.55
10		21.04	14.91	12.55	11.28	10.48	9.93	9.52	9.20	8.96	8.75	8.59	8.45	8.33	8.22	8.13	8.05	7.98	7.91	7.86	7.80	7.71	7.64	7.57	7.52	7.47
11		19.69	13.81	11.56	10.35	9.58	9.05	8.66	8.36	8.12	7.92	7.76	7.63	7.51	7.41	7.32	7.24	7.18	7.11	7.06	7.01	6.92	6.85	6.79	6.73	6.68
12		18.64	12.97	10.80	9.63	8.89	8.38	8.00	7.71	7.48	7.29	7.14	7.01	6.89	6.79	6.71	6.63	6.57	6.51	6.45	6.41	6.32	6.25	6.19	6.14	6.09
13		17.82	12.31	10.21	9.07	8.35	7.86	7.49	7.21	6.98	6.80	6.65	6.52	6.41	6.31	6.23	6.16	6.09	6.03	5.98	5.93	5.85	5.78	5.72	5.67	5.63
14		17.14	11.78	9.73	8.62	7.92	7.44	7.08	6.80	6.58	6.40	6.26	6.13	6.02	5.93	5.85	5.78	5.71	5.66	5.60	5.56	5.48	5.41	5.35	5.30	5.25
15		16.59	11.34	9.34	8.25	7.57	7.09	6.74	6.47	6.26	6.08	5.94	5.81	5.71	5.62	5.54	5.46	5.40	5.35	5.29	5.25	5.17	5.10	5.04	4.99	4.95
16		16.12	10.97	9.01	7.94	7.27	6.81	6.46	6.20	5.98	5.81	5.67	5.55	5.44	5.35	5.27	5.21	5.14	5.09	5.04	4.99	4.91	4.85	4.79	4.74	4.70
17		15.72	10.66	8.73	7.68	7.02	6.56	6.22	5.96	5.75	5.58	5.44	5.32	5.22	5.13	5.05	4.99	4.92	4.87	4.82	4.78	4.70	4.63	4.58	4.53	4.48
18		15.38	10.39	8.49	7.46	6.81	6.36	6.02	5.76	5.56	5.39	5.25	5.13	5.03	4.94	4.87	4.80	4.74	4.68	4.63	4.59	4.51	4.45	4.39	4.34	4.30
19		15.08	10.16	8.28	7.27	6.62	6.18	5.85	5.59	5.39	5.22	5.08	4.97	4.87	4.78	4.70	4.64	4.58	4.52	4.47	4.43	4.35	4.29	4.23	4.19	4.14
20		14.82	9.95	8.10	7.10	6.46	6.02	5.69	5.44	5.24	5.08	4.94	4.82	4.72	4.64	4.56	4.50	4.44	4.38	4.33	4.29	4.21	4.15	4.09	4.05	4.01
22		14.38	9.61	7.80	6.81	6.19	5.76	5.44	5.19	4.99	4.83	4.70	4.58	4.49	4.40	4.33	4.26	4.20	4.15	4.10	4.06	3.98	3.92	3.86	3.82	3.78
24		14.03	9.34	7.55	6.59	5.98	5.55	5.24	4.99	4.80	4.64	4.51	4.39	4.30	4.21	4.14	4.07	4.02	3.96	3.92	3.87	3.80	3.74	3.68	3.63	3.59
26		13.74	9.12	7.36	6.41	5.80	5.38	5.07	4.83	4.64	4.48	4.35	4.24	4.14	4.06	3.99	3.92	3.86	3.81	3.77	3.72	3.65	3.59	3.53	3.49	3.45
28		13.50	8.93	7.19	6.25	5.66	5.24	4.93	4.70	4.51	4.35	4.22	4.11	4.01	3.93	3.86	3.80	3.74	3.69	3.64	3.60	3.52	3.46	3.41	3.36	3.32
30		13.29	8.77	7.05	6.13	5.53	5.12	4.82	4.58	4.39	4.24	4.11	4.00	3.91	3.83	3.75	3.69	3.63	3.58	3.54	3.49	3.42	3.36	3.30	3.26	3.22

The F-distribution

- The F-distribution is the *null distribution* of test statistics, in particular used in ANOVA
 - if you have two sets of data, and results are not outside the parameters of expectations described by the F-distribution,
 - then the null hypothesis is said to be true
 - the null hypothesis says that the two sets have no statistically significant difference
- probability density function

$$f(x, d_1, d_2) = \frac{\sqrt{\frac{(d_1 x)^{d_1} \cdot d_2^{d_2}}{(d_1 x + d_2)^{d_1 + d_2}}}}{xB(\frac{d_1}{2}, \frac{d_2}{2})}; B(x, y) = \int_0^1 t^{x-1} (t-1)^{y-1} dt$$

The F-distribution



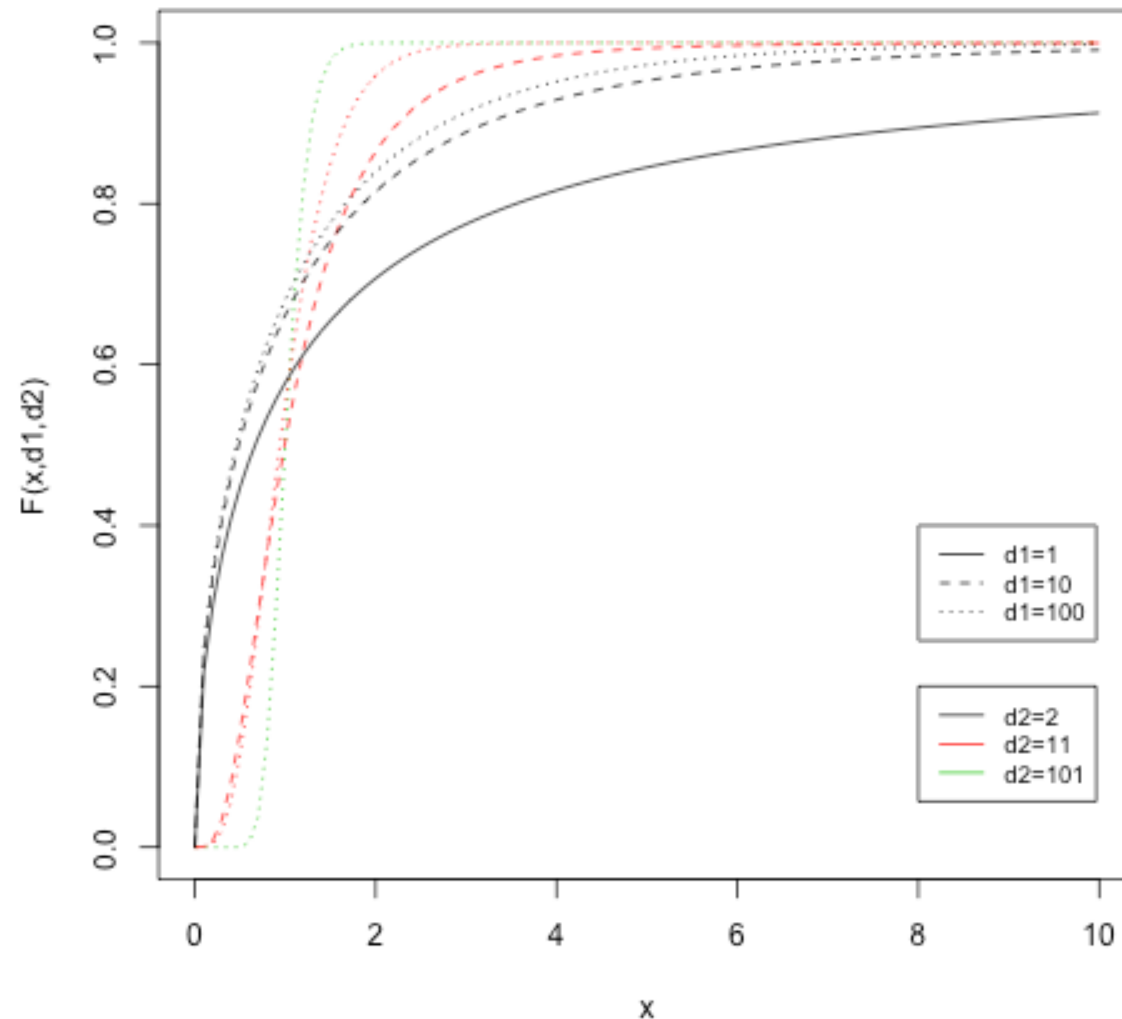
Example PDF functions plotted with R

The F-distribution

- The F-distribution is the *null distribution* of test statistics, in particular using in ANOVA
 - if you have two sets of data, and results are not outside the parameters of expectations described by the F-distribution,
 - then the null hypothesis is said to be true
 - the null hypothesis says that the two sets have no statistically significant difference
- cumulative distribution function

$$F(x, d_1, d_2) = I_{\frac{d_1 x}{d_1 x + d_2}}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)$$
$$I_a(x, y) = \frac{\int_0^a t^{x-1} (t-1)^{y-1} dt}{\int_0^1 t^{x-1} (t-1)^{y-1} dt}$$

The F-distribution



Example CDF functions plotted with R

ANOVA Summary

Variation	Alternatives	Error	Total
Sum of squares	SSA	SSE	SST
Deg freedom	$k-1$	$k(n-1)$	$kn-1$
Mean square	$\sigma_a^2 = \frac{SSA}{k-1}$	$\sigma_e^2 = \frac{SSE}{k(n-1)}$	
Computed F	$\frac{\sigma_a^2}{\sigma_e^2}$		
Tabulated F	$F_{[1-\alpha; k-1, k(n-1)]}$		

ANOVA Example

	Alternatives			
Measurements	1	2	3	Overall mean
1	0.0972	0.1382	0.7966	
2	0.0971	0.1432	0.5300	
3	0.0969	0.1382	0.5152	
4	0.1954	0.1730	0.6675	
5	0.0974	0.1383	0.5298	
Column mean				
Effects				

ANOVA Example

	Alternatives			
Measurements	1	2	3	Overall mean
1	0.0972	0.1382	0.7966	
2	0.0971	0.1432	0.5300	
3	0.0969	0.1382	0.5152	
4	0.1954	0.1730	0.6675	
5	0.0974	0.1383	0.5298	
Column mean	$\bar{y}_{.1} = 0.1168$	$\bar{y}_{.2} = 0.1462$	$\bar{y}_{.3} = 0.6078$	$\bar{y} = 0.2903$
Effects				
Column sum	0.5840	0.7309	3.0391	

ANOVA Example

	Alternatives			
Measurements	1	2	3	Overall mean
1	0.0972	0.1382	0.7966	
2	0.0971	0.1432	0.5300	
3	0.0969	0.1382	0.5152	
4	0.1954	0.1730	0.6675	
5	0.0974	0.1383	0.5298	
Column mean	$\bar{y}_{\cdot 1} = 0.1168$	$\bar{y}_{\cdot 2} = 0.1462$	$\bar{y}_{\cdot 3} = 0.6078$	$\bar{y} = 0.2903$
Effects $\alpha_j = \bar{y}_{\cdot j} - \bar{y}$				

ANOVA Example

	Alternatives			
Measurements	1	2	3	Overall mean
1	0.0972	0.1382	0.7966	
2	0.0971	0.1432	0.5300	
3	0.0969	0.1382	0.5152	
4	0.1954	0.1730	0.6675	
5	0.0974	0.1383	0.5298	
Column mean	$\bar{y}_{\cdot 1} = 0.1168$	$\bar{y}_{\cdot 2} = 0.1462$	$\bar{y}_{\cdot 3} = 0.6078$	$\bar{y} = 0.2903$
Effects	$\bar{\alpha}_1 = -0.1735$	$\bar{\alpha}_2 = -0.1441$	$\bar{\alpha}_3 = 0.3175$	
SSA	$5 \cdot \sum_{j=1}^3 \alpha_j^2 =$			0.7585

ANOVA Example

	Errors			
Measurements	$e_{i1} = y_{i1} - \bar{y}_{.1}$	$e_{i2} = y_{i2} - \bar{y}_{.2}$	$e_{i3} = y_{i3} - \bar{y}_{.3}$	SSE
1	-0.0196	-0.0080	0.1888	$\sum_{j=1}^3 \sum_{i=1}^5 e_{ij}^2$
2	-0.0197	-0.0030	-0.0778	
3	-0.0199	-0.0080	-0.0926	
4	0.0786	0.0268	0.0597	
5	-0.0194	-0.0079	-0.0780	
Column mean	$\bar{y}_{.1} = 0.1168$	$\bar{y}_{.2} = 0.1462$	$\bar{y}_{.3} = 0.6078$	0.0685
SST=SSA+SSE	0.8270			

ANOVA Example

Variation	Alternatives	Error	Total
Sum of squares	$SSA = 0.7585$	$SSE = 0.0685$	$SST = 0.8270$
Deg freedom	$k - 1 = 2$	$k(N - 1) = 12$	$kn - 1 = 14$
Mean square	$\sigma_a^2 = 0.3793$	$\sigma_3^2 = 0.0057$	
Computed F	$\frac{0.3793}{0.0057} = 66.4$		
Tabulated F	$F_{[0.95;2,12]} = 3.89$ $F_{[0.99;2,12]} = 6.93$ $F_{[0.999;2,12]} = 12.97$		

$$SSA/SST = 0.7585/0.8270 = 0.917$$

→ **91.7%** of total variation in measurements is **due to differences** among alternatives

$$SSE/SST = 0.0685/0.8270 = 0.083$$

→ **8.3%** of total variation in measurements is **due to noise** in measurements

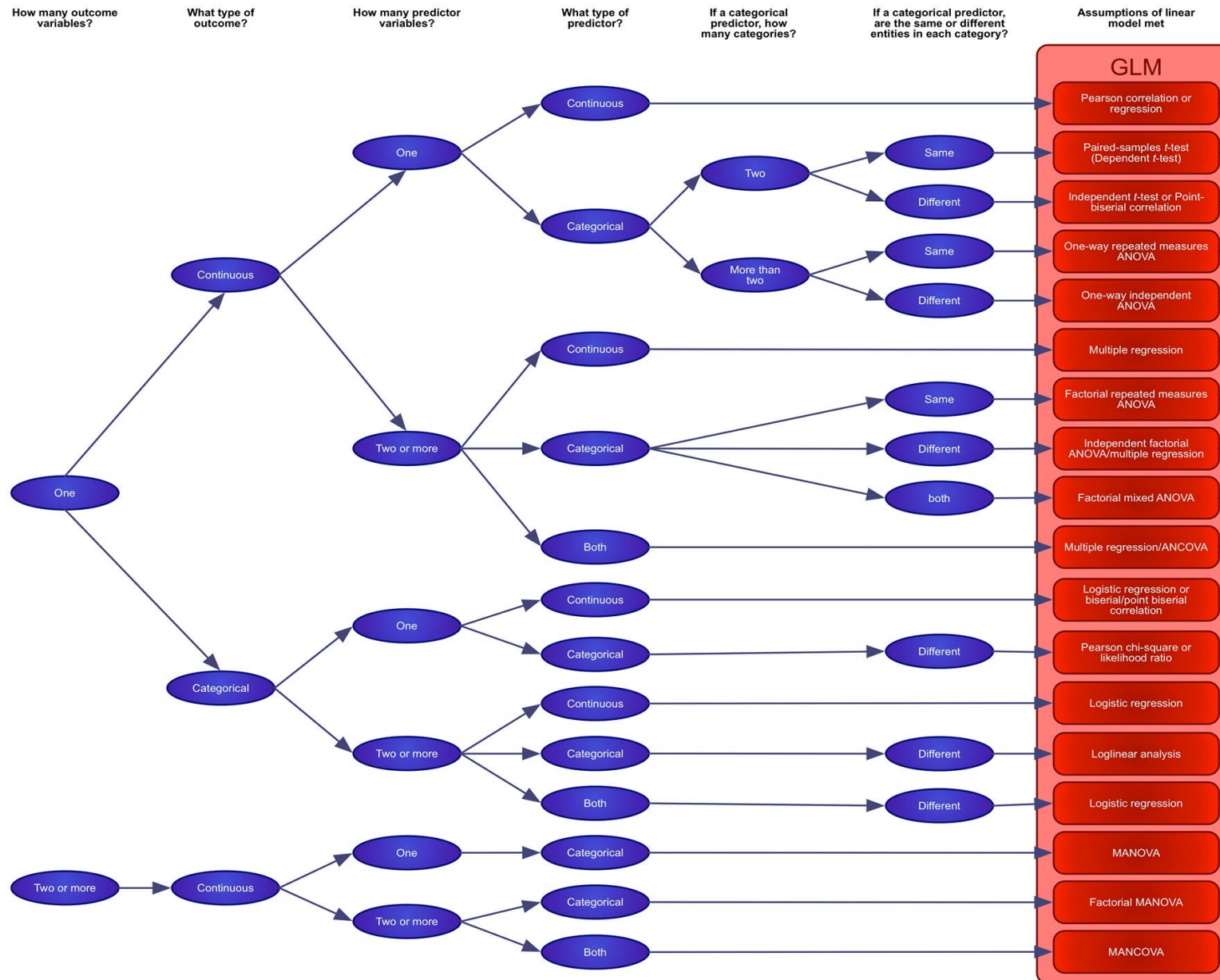
ANOVA Example

Variation	Alternatives	Error	Total
Sum of squares	$SSA = 0.7585$	$SSE = 0.0685$	$SST = 0.8270$
Deg freedom	$k - 1 = 2$	$k(N - 1) = 12$	$kn - 1 = 14$
Mean square	$\sigma_a^2 = 0.3793$	$\sigma_3^2 = 0.0057$	
Computed F	$\frac{0.3793}{0.0057} = 66.4$		
Tabulated F	$F_{[0.95;2,12]} = 3.89$ $F_{[0.99;2,12]} = 6.93$ $F_{[0.999;2,12]} = 12.97$		

Computed F statistic $>$ tabulated F statistic

→ **99.9% confidence** that differences among alternatives are statistically significant.

SPSS Chart of methods



From Field, A. P. (2013). *Discovering statistics using IBM SPSS Statistics: And sex and drugs and rock 'n' roll* (4th ed.). London: Sage.



ANOVA Summary

- Useful for partitioning total variation into components
 - Experimental error
 - Including differences between human participants – which can be due to actual differences but also due to lack of attention
 - Variation among alternatives
 - Including alternatives comprising a numeric score, a rank or a count
- Compare more than two alternatives
- Note, does not tell you *where* differences may lie
 - Use confidence intervals for pairs
 - Or use contrasts