The ratio $MS_{between}$ to MS_{within}

The ratio $MS_{between}$ to MS_{within}

THE CHI-SQUARE DISTRIBUTION

Let $X_1, X_2, \ldots, X_\nu$ be ν independent normally distributed random variables with mean zero and variance 1. Consider the random variable

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_\nu^2 \tag{38}$$

where χ^2 is called *chi square*. Then we can show that for $x \geqq 0$,

$$P(\chi^2 \leqq x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2}\, du \tag{39}$$

and $P(\chi^2 \leqq x) = 0$ for $x > 0$.

The ratio $MS_{between}$ to MS_{within}

THE CHI-SQUARE DISTRIBUTION

Let $X_1, X_2, \ldots, X_\nu$ be ν independent normally distributed random variables with mean zero and variance 1. Consider the random variable

$$\chi^2 = X_1^2 + X_2^2 + \cdots + X_\nu^2 \tag{38}$$

where χ^2 is called *chi square*. Then we can show that for $x \geqq 0$,

$$P(\chi^2 \leqq x) = \frac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2}\, du \tag{39}$$

and $P(\chi^2 \leqq x) = 0$ for $x > 0$.

Suppose $y_1, y_2, ..., y_n$ is a random sample from a Normal distribution with mean μ and variance σ^2. Then

$$\frac{SS}{\sigma^2} = \frac{\sum_{i=1}^{n} (y_i - \bar{y})^2}{\sigma^2} \sim \chi^2_{n-1}$$

The ratio $MS_{between}$ to MS_{within}

THE *F* DISTRIBUTION

A random variable is said to have the *F distribution* (named after R. A. Fisher) *with* ν_1 *and* ν_2 *degrees of freedom* if its density function is given by

$$f(u) = \begin{cases} \dfrac{\Gamma\left(\dfrac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\dfrac{\nu_1}{2}\right)\Gamma\left(\dfrac{\nu_2}{2}\right)} \nu_1^{\nu_1/2}\nu_2^{\nu_2/2}u^{(\nu_1/2)-1}(\nu_2+\nu_1 u)^{-(\nu_1+\nu_2)/2} & u > 0 \\ 0 & u \leqq 0 \end{cases} \tag{45}$$

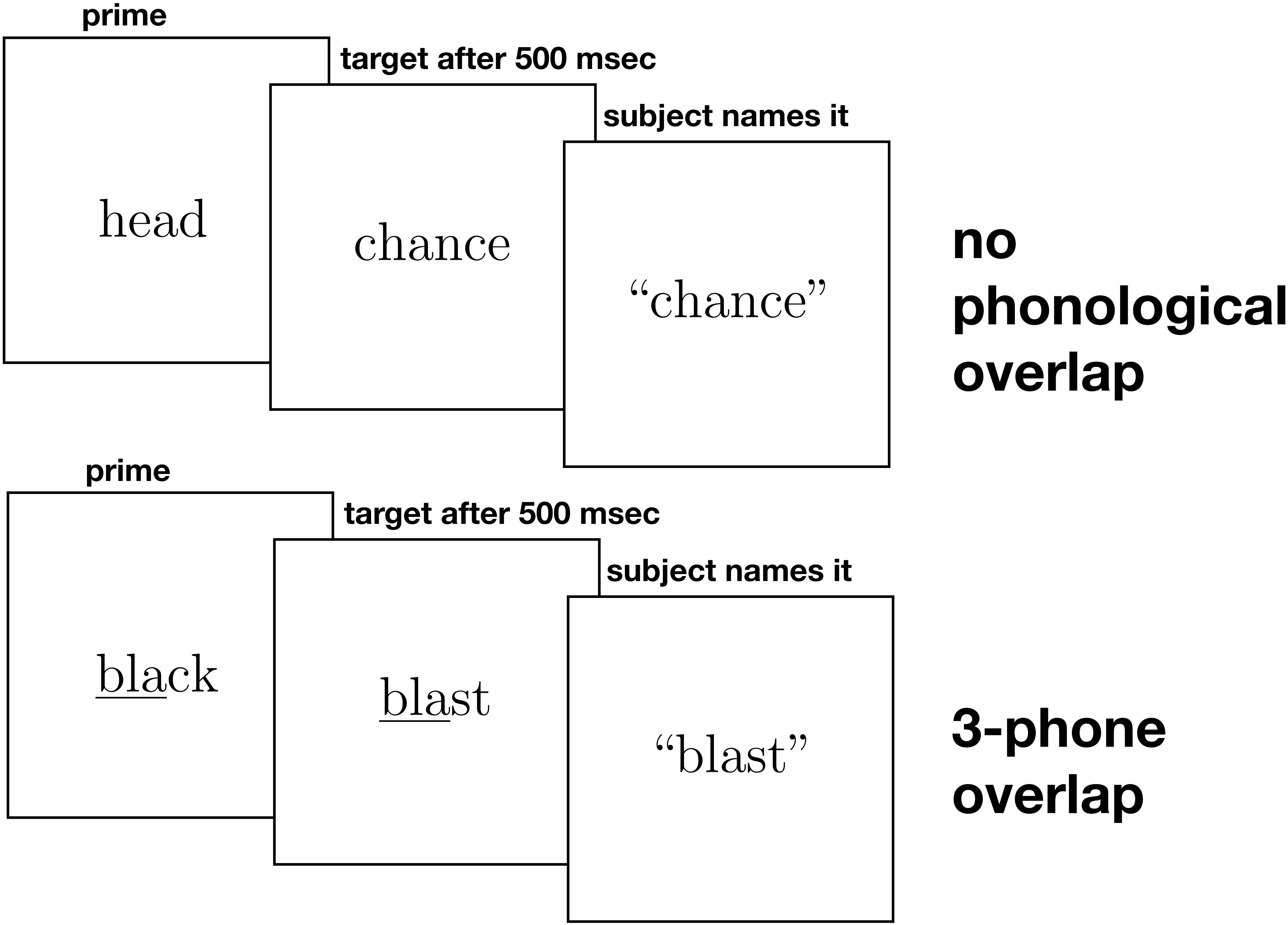
Percentile values of the F distribution for ν_1, ν_2 degrees of freedom are denoted by F_{p,ν_1,ν_2}, or briefly F_p if ν_1, ν_2 are understood.

Theorem 4-7: Let V_1 and V_2 be independent random variables that are chi square distributed with ν_1 and ν_2 degrees of freedom, respectively. Then the random variable

$$V = \frac{V_1/\nu_1}{V_2/\nu_2} \tag{48}$$

has the F distribution with ν_1 and v_2 degrees of freedom.

Pitt & Shoaf 2002


prime
head
target after 500 msec
chance
subject names it
“chance”
no phonological overlap
prime
black
target after 500 msec
blast
subject names it
“blast”
3-phone overlap

Johnson divides up the data for us

```
> head(ps1)
  subj position   rt
1 AALA      mid  845
2 AAMS     late  744
3 ADAN      mid  804
4 AING      mid  825
5 AITT     late  720
6 ALER    early 1179
```

Johnson divides up the data for us

```
> head(ps1)
  subj position   rt
1 AALA      mid  845
2 AAMS     late  744
3 ADAN      mid  804
4 AING      mid  825
5 AITT     late  720
6 ALER    early 1179
```

Get the SS by multiplying by n-1

```
> tapply(ps1$rt,ps1$position,length)
early  late   mid
   31    33    32
> attach(ps1)
> var(rt[position=="early"])*(31-1) + var(rt[position=="mid"])*(32-1) + var(rt[position=="late"])*(33-1)
[1] 3136368
> length(rt)
[1] 96
> var(rt)*(96-1)
[1] 3483523
```

Johnson divides up the data for us

```
> head(ps1)
  subj position   rt
1 AALA      mid  845
2 AAMS     late  744
3 ADAN      mid  804
4 AING      mid  825
5 AITT     late  720
6 ALER    early 1179
```

Get the SS by multiplying by n-1

```
> tapply(ps1$rt,ps1$position,length)
early  late   mid
   31    33    32
> attach(ps1)
> var(rt[position=="early"])*(31-1) + var(rt[position=="mid"])*(32-1) + var(rt[position=="late"])*(33-1)
[1] 3136368
> length(rt)
[1] 96
> var(rt)*(96-1)
[1] 3483523
```

weighted average of within-level variance

total variance across the whole experiment

Johnson divides up the data for us

```
> head(ps1)
  subj position   rt
1 AALA      mid  845
2 AAMS     late  744
3 ADAN      mid  804
4 AING      mid  825
5 AITT     late  720
6 ALER    early 1179
```

Get the SS by multiplying by n-1

```
myvariance <- function(x) {
  n <- length(x);   # to figure out what the mean is,
  xbar <- sum(x)/n; # we need to divide by the number of observations

  nminusone <- length(x) -1; # the degrees of freedom are one less
  sum((x-xbar)^2)/nminusone  # divide sum of squared deviations by the degrees of freedom
}
```

```
> tapply(ps1$rt,ps1$position,length)
early  late   mid
   31    33    32
> attach(ps1)
> var(rt[position=="early"])*(31-1) + var(rt[position=="mid"])*(32-1) + var(rt[position=="late"])*(33-1)
[1] 3136368
> length(rt)
[1] 96
> var(rt)*(96-1)
[1] 3483523
```

weighted average of within-level variance

total variance across the whole experiment

The treatment effect

$$x_{ij} - \bar{x} = (\bar{x}_j - \bar{x}) + (x_{ij} - \bar{x}_j)$$

The treatment effect

$$x_{ij} - \bar{x} = (\bar{x}_j - \bar{x}) + (x_{ij} - \bar{x}_j)$$

347,155 **=** **3,483,523** **-** **3,136,368**

Acknowledge degrees of freedom

$$MS = \frac{SS}{df}$$

```
> msb <- 3136368 / (3-1)
> msw <- 347155/ ((31-1)+(32-1)+(33-1))
> 1-pf(msb/msw,2,93)
[1] 0.007585556
```

How likely is this F-ratio?

```
> fitted <- aov(rt~position,data=ps1)
> anova(fitted)
Analysis of Variance Table

Response: rt
          Df  Sum Sq Mean Sq F value   Pr(>F)
position   2  347155  173577  5.1469 0.007586 **
Residuals 93 3136368   33724
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Flynn & Lust 80

Relative clauses

subject-extracted

(Gayle hates) the guy who telephoned Oprah

object-extracted

(Gayle hates) the guy who Oprah telephoned

Relative clauses

“free”

(I’ll have) <u>what</u> he’s having

lexically-headed

(I’ll have) <u>the drink</u> which he’s having

Relative Types

Lexically Headed Relatives

I.	Determinate Head	Big Bird pushes the balloon which bumps Ernie
II.	Non-determinate Head	Ernie pushes the thing which touches Big Bird

Nonlexically Headed

III.	Free Relative ("what")	Cookie Monster hits what pushes Big Bird

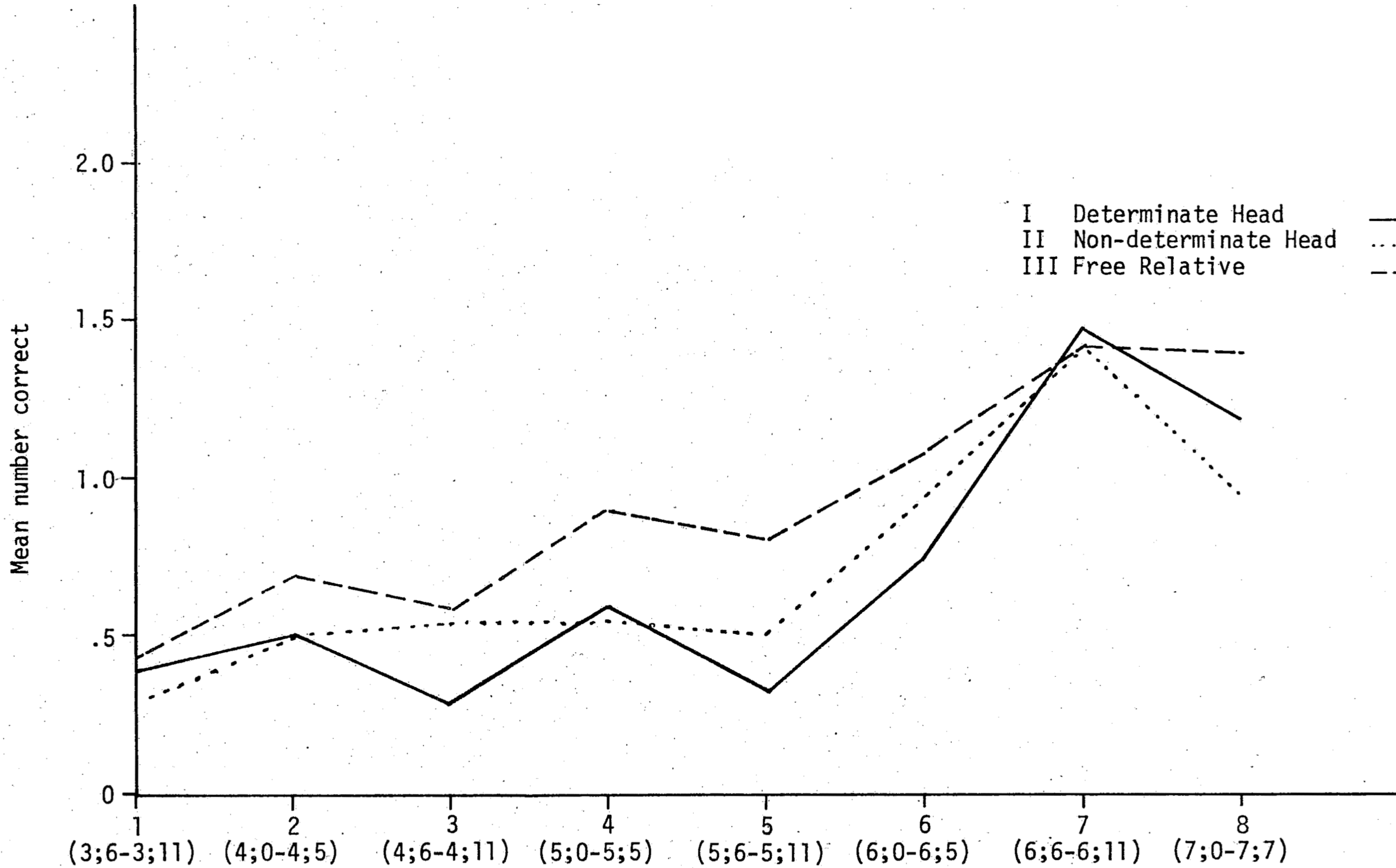
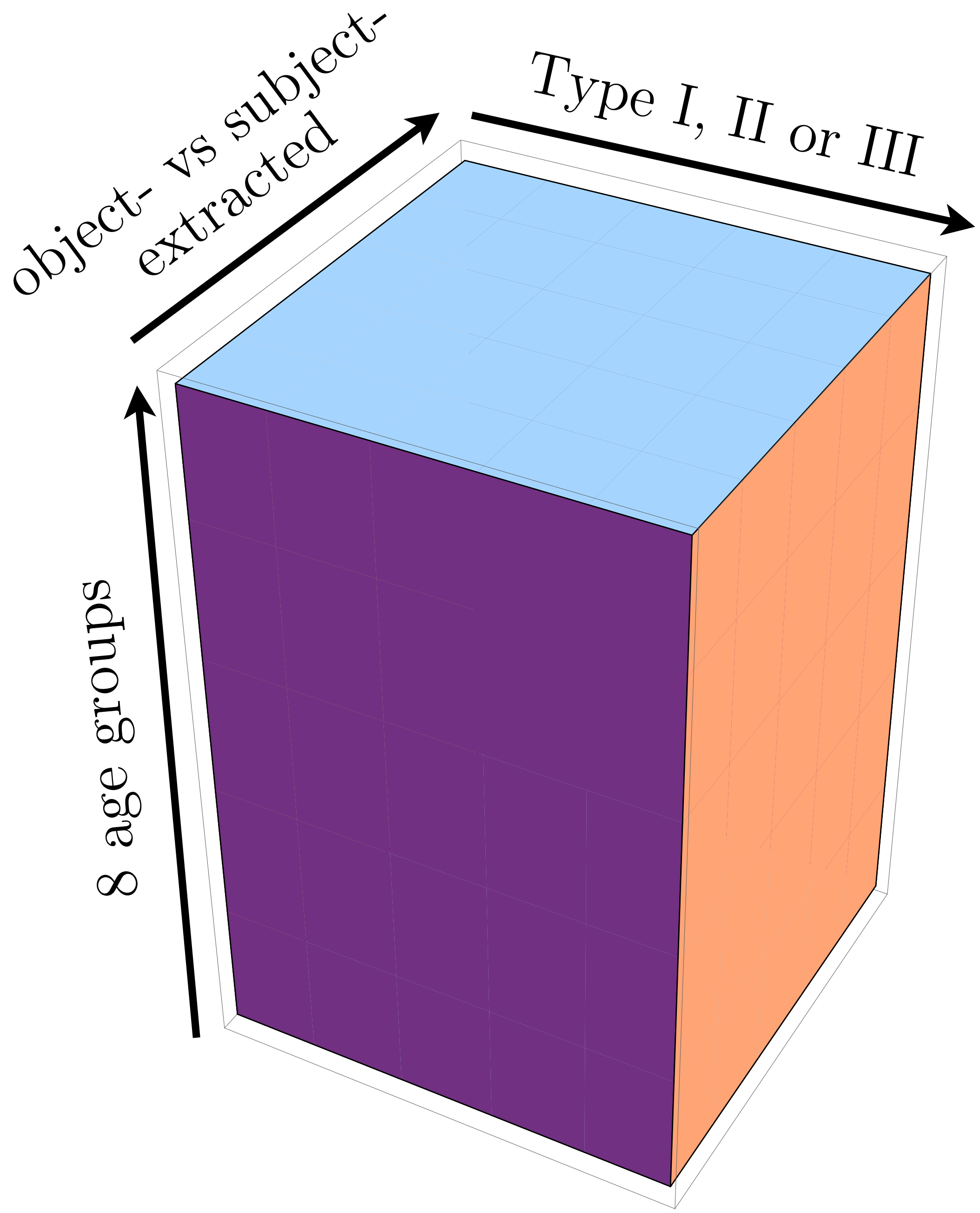


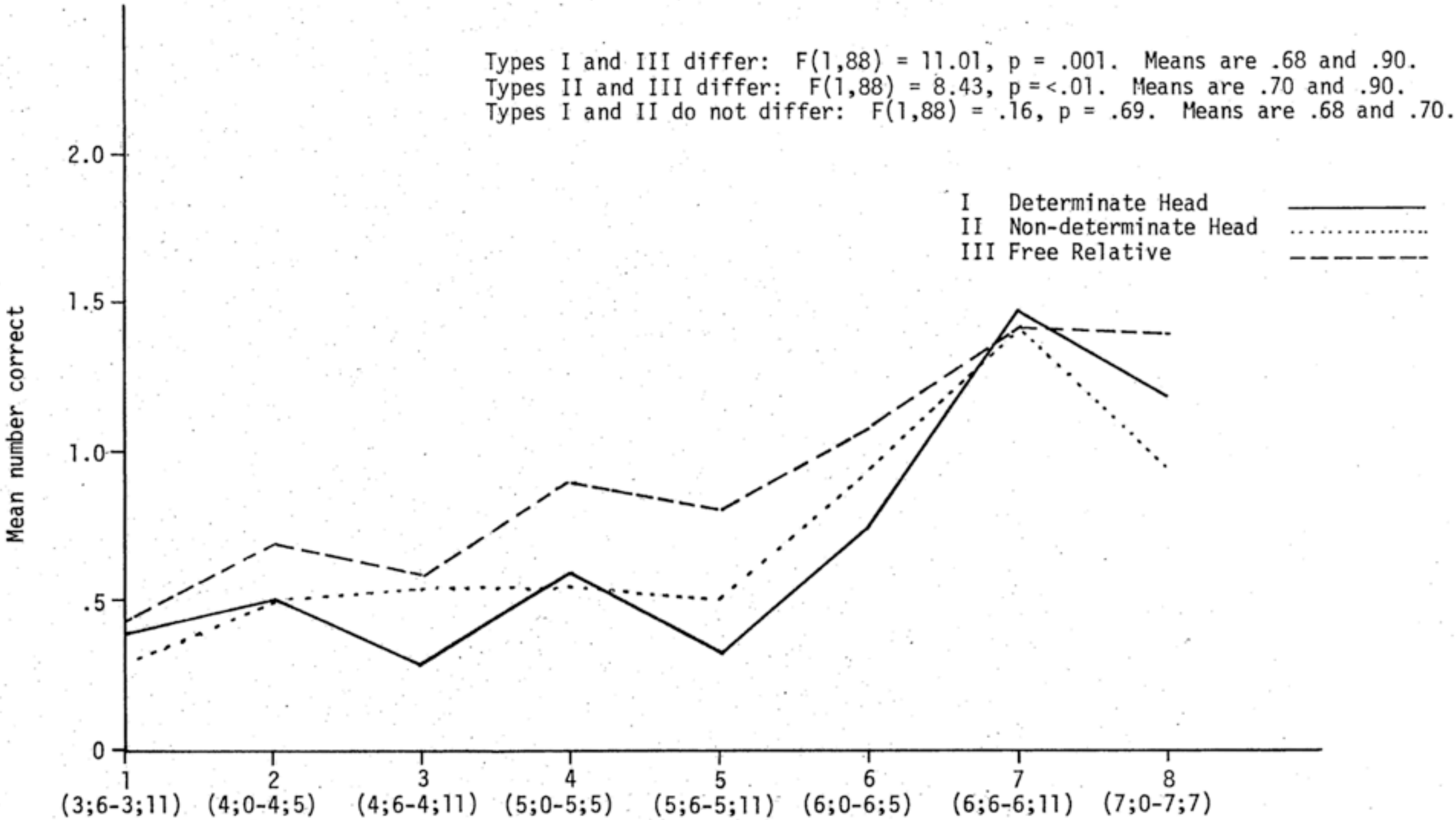
Figure 1: Development of Imitation Success on 3 Relative Types Over 8 Age Groups

3x2x8



Relative clause Type had a significant effect on the success of children's imitations
$F(2,176)=7.49$ $p < .001$

Planned comparisons



t^2 = F

t statistic is a ratio

$$\frac{Z}{\sqrt{V/\nu}}$$

t^2 is a ratio of chi-squared r.v.'s