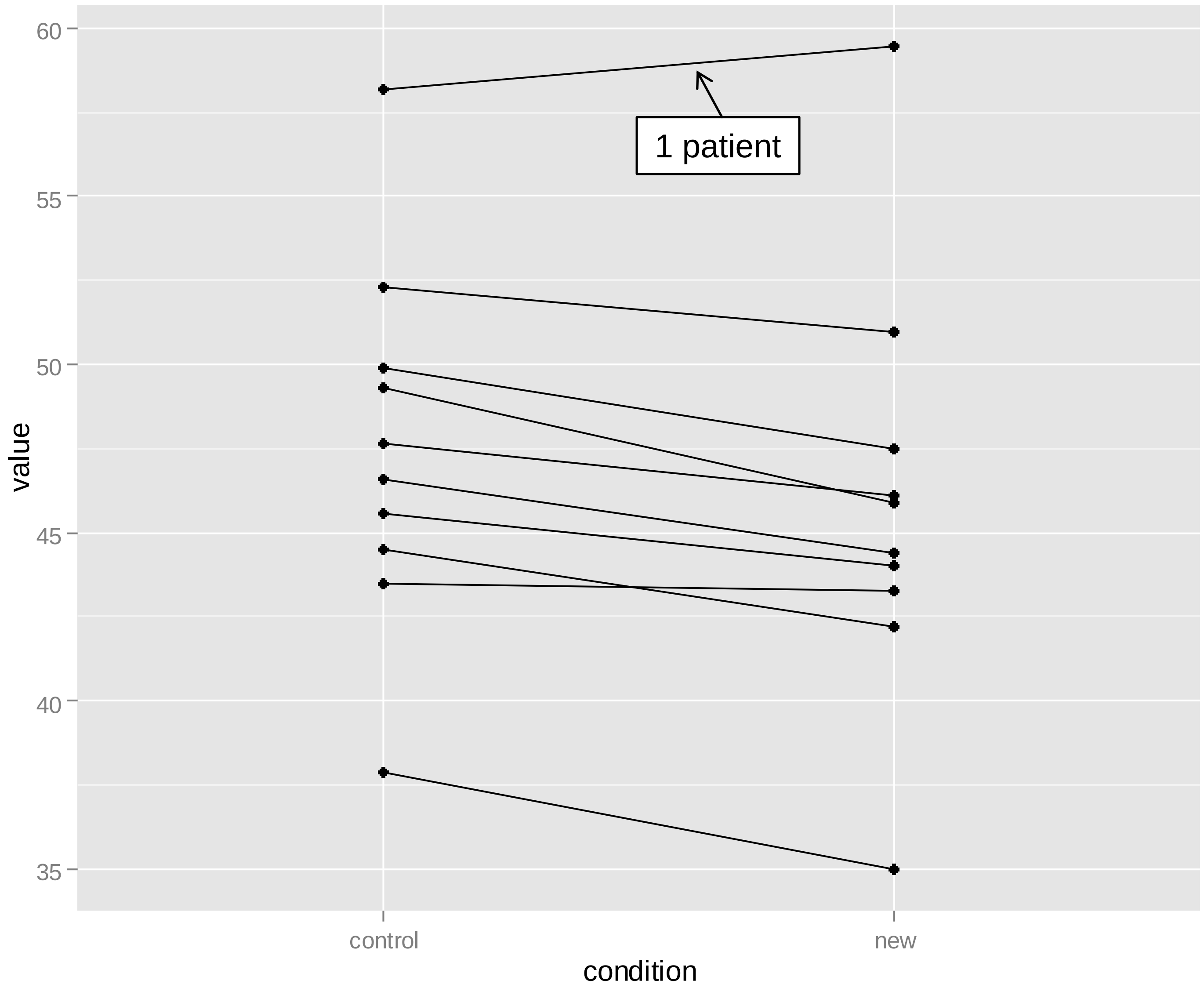
Complete Block Designs

Lukas Meier, Seminar für Statistik

Remember: Paired t-Test (Example from Elliott, 2006)

- Want to compare two different eye-drops ("new" vs. "control").
- **Every** subject gets **both treatments** (meaning: one per eye; at the **same** time).
- At the end, measure redness on quantitative scale in **every** eye.
- For every patient, calculate the difference "new - control".
- Perform standard one-sample t-test with these differences.

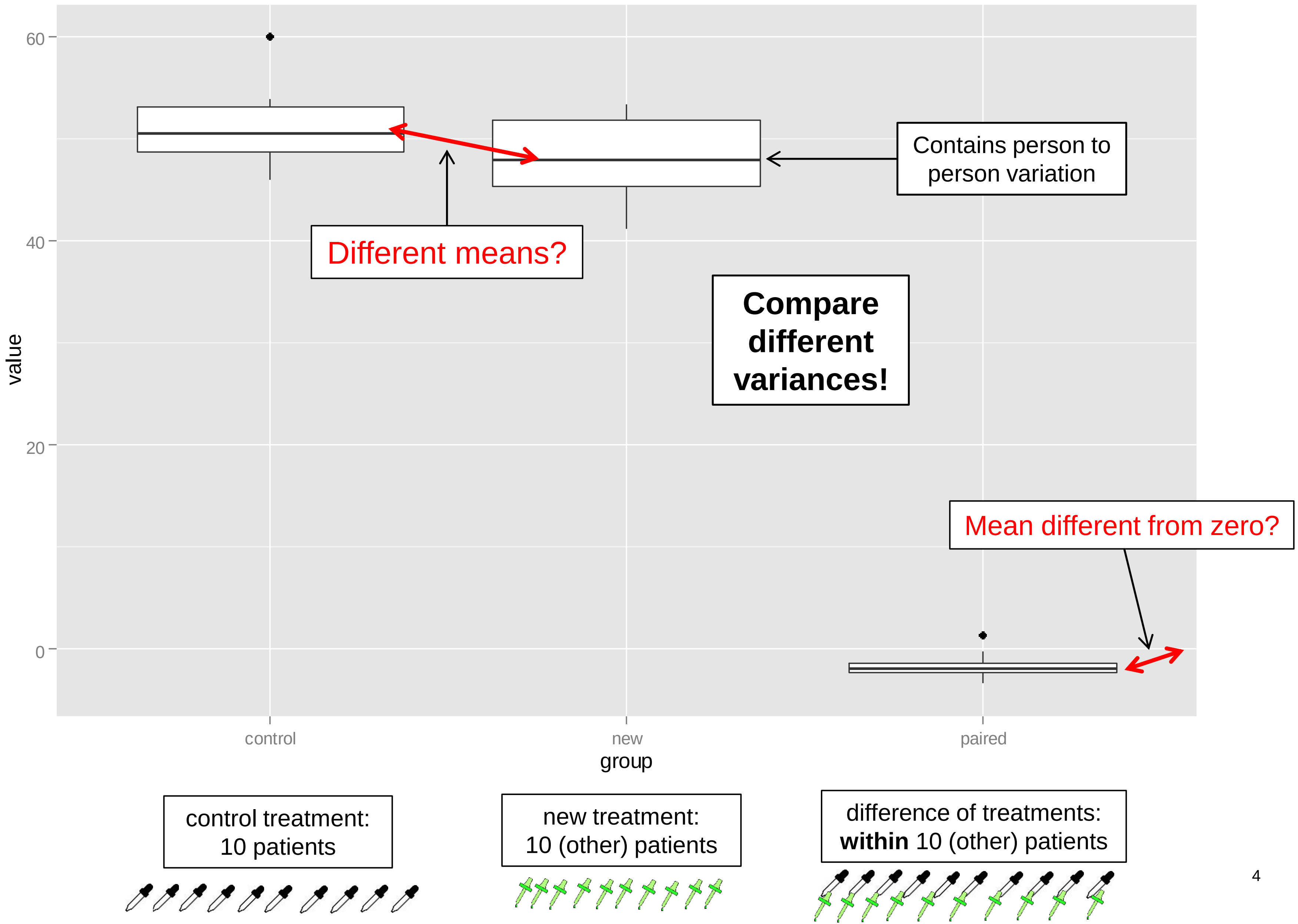
Fictional Data Set of 10 Patients



Paired t-Test

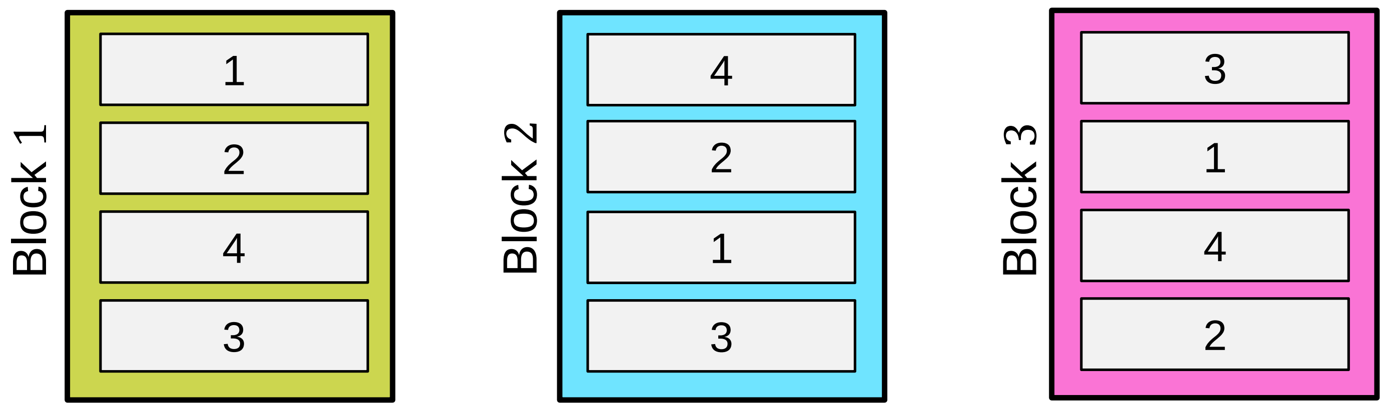
- Instead of using both eyes of 10 patients we could also do a similar experiment with
  - 10 patients getting the control treatment in one eye
  - 10 other patients getting the new treatment in one eye
- See next slide for potential data-sets.
- As mentioned in the first week, we can reduce **variance** by using **homogeneous experimental units**.
- A set of units that is homogeneous in some sense is called a **block**.
- In this example, a **block** is given by a **person**.

Paired vs. Unpaired Data



Randomized Complete Block Designs (RCB)

- A **Randomized Complete Block Design** (RCB) is the most basic blocking design.
- Assume we have r blocks containing g units each.



- Here, $r = 3$ blocks with $g = 4$ units.
- In every of the r blocks we randomly assign the g treatments to the g units, **independently** of the other blocks.

Randomized Complete Block Designs (RCB)

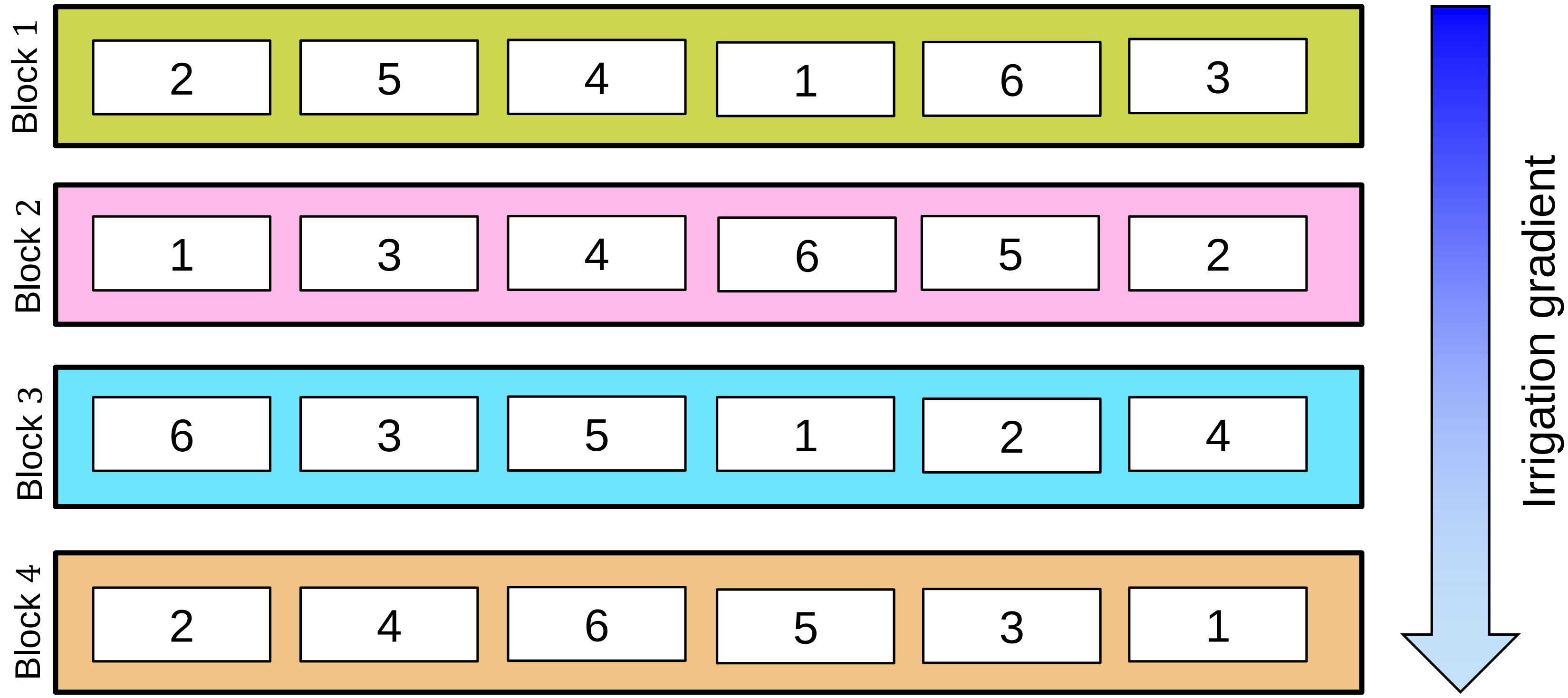
- Hence, a blocking design uses a **restricted randomization** scheme. Each block gets its **"own" randomization**.
- Blocking exists at the time of randomization!
- We call a blocking design **complete** if every treatment is used in every block.
- In the standard setup, we observe every treatment (only) **once** in every block, hence we have a total of r (the number of blocks) observations per treatment.
- Therefore, we have **no replicates** (for treatment and block combinations).

Example (Example 8.1 in Kuehl, 2000)

- Researchers wanted to evaluate the effect of several different fertilization timing schedules on stem tissue nitrate amounts.
- Treatment: Six different nitrogen application timing and rate schedules (including a control treatment of no nitrogen).
- Response: Stem tissue nitrate amount.
- Experiment design: irrigated field with a **water gradient** along one direction, see next slide.
- We already know:

  Available moisture will have an influence on the response.

Example: Layout of Experimental Design



- Any differences in plant responses caused by the **water gradient** will be associated with **blocks**.
- We also say: we **control for the water gradient.**

Example: Analysis

- $Y_{ij} = \mu + \underbrace{\alpha_i}_{\text{treatment effect}} + \underbrace{\beta_j}_{\text{block effect}} + \epsilon_{ij}$ with the usual assumptions for ϵ_{ij}.

- By only using main effects we implicitly assume that the effects are **additive**.

- Due to the **balanced design** we can use our standard estimates (one at a time) and sum of squares.

```
> fit <- aov(y ~ block + treatment, data = nitro)
> summary(fit)
            Df Sum Sq Mean Sq F value  Pr(>F)
block        3  197.0   65.67   9.120 0.00112 **
treatment    5  201.3   40.26   5.592 0.00419 **
Residuals   15  108.0    7.20
```

- Typically, we are **not** making inference about blocks (we already know that blocks are different!).

Interaction of Treatment with Block Factor

- The blocking may result in (very) large differences between units from different blocks (which is **ok**).
- In the model we **assumed** that the effects are additive.
- Meaning: the treatment effects are **constant** from block to block.
- If we only have **one observation** per treatment and block combination we can potentially only detect interaction effects of the multiplicative form.
- If we want to fit a model with interaction, we would need more than one observation per treatment and block combination. What does interaction mean?

Factorials in Complete Block Designs

- Conceptually it is straightforward to have (e.g.) a **two-factor factorial** in a randomized complete block design.

- The analysis is straightforward. In R we would just use the model formula `Y ~ Block + A * B`

Source	*df*	
Block	$r-1$	
A	$a-1$	
B	$b-1$	
AB	$(a-1)\cdot(b-1)$	
Error	$(ab-1)\cdot(r-1)$	← "Leftovers"
Total	$rab-1$	← # observations – 1

- We can test the interaction AB even if we only have one replicate per AB combination per block.

How Much Does Blocking Increase Precision?

- Squared standard errors for treatment means are
  - RCB design (what we've just done): $\frac{\sigma^2_{RCB}}{r}$
  - Completely randomized design: $\frac{\sigma^2_{CRD}}{n}$

Number of observations per treatment

- If we want to have the **same precision**, we have to ensure that

$$\frac{\sigma^2_{RCB}}{r} = \frac{\sigma^2_{CRD}}{n}.$$

If we know σ^2_{RCB} and σ^2_{CRD} than we have to use a ratio of

$$\frac{n}{r} = \frac{\sigma^2_{CRD}}{\sigma^2_{RCB}}.$$

How Much Does Blocking Increase Precision?

- σ^2_{RCB} is estimated by MS_E of our RCB.
- What about σ^2_{CRD}?
- Can be estimated using a properly **weighted average** of MS_E and MS_{Block}

$$\hat{\sigma}^2_{CRD} = w \cdot MS_{Block} + (1 - w) \cdot MS_E$$

  where w is some weight (see Oehlert, page 323).
- **Relative efficiency** is then defined as:

$$RE = \frac{\hat{\sigma}^2_{CRD}}{\hat{\sigma}^2_{RCB}}$$

  (sometimes multiply with correction factor for df’s).
- RE gives us the ratio $\frac{n}{r}$.

How Much Does Blocking Increase Precision?

- In our example: relative efficiency ≈ 2.
- Meaning: A CRD would need twice as many experimental units to achieve the **same efficiency** (precision).
- Here: 8 replications per treatment (instead of 4).
- **Easier for a quick check:** Have a look at the ratio $\frac{MS_{Block}}{MS_E}$

$$\boxed{\frac{MS_{Block}}{MS_E} > 1 \iff \text{Relative Efficiency} > 1}$$

More than One Blocking Factor

- Up to now: one blocking factor involved, i.e. we can block on a **single source** of variation.
- Sometimes: need to block on **more than one** source.
- We will discuss some special cases.
  - Latin Squares
  - Graeco-Latin Squares

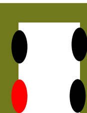
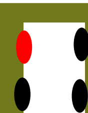
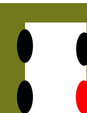
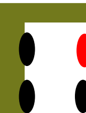
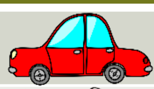
Example: Car Tires (Kuehl, 2000, Example 8.2)

- An experiment tests **4** car tire **treatments** (A, B, C, D) on **4 cars**. Response: Wear of a tire.

- Each treatment appears on one of the **4 positions** of each **car**.

Block factors

- Experiment set-up was as follows:

Tire position				
	A	B	C	D
	B	C	D	A
	C	D	A	B
	D	A	B	C

Latin Squares

- This design is a so called **Latin Square.**
- Each treatment (the Latin letters) appears exactly **once** in **each row** and exactly **once** in **each column**.
- A Latin Square **blocks** on both **rows** and **columns simultaneously**.
- The design is very **restrictive**. A Latin Square needs to have
  - g treatments (the Latin letters)
  - Two block factors each having g levels (the rows and the columns)
  - Hence, a total of g^2 experimental units
- We're only seeing g^2 out of g^3 possible combinations (but the subset we see is selected in a smart, balanced way).

Latin Squares

- A Latin Square is nothing else than an assignment of treatments to units with the **side constraints**
  - each treatment appears **exactly once in each row**.
  - each treatment appears **exactly once in each column**.
- Picking a random Latin Square isn’t trivial: Fisher-Yates algorithm (see book for details).

Analysis of Latin Squares

- Use **main effects model** with **treatment**, **row** and **column** effects.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \epsilon_{ijk}$$

treatment

Block factor 1 (rows)

Block factor 2 (columns)

- The design is **balanced** having the effect that our usual estimators and sums of squares are “working”.
- As in an RCB we do **not** test for the block effects.
- Latin Squares can have **few degrees** of freedom for error if g is small, making detection of treatment effects difficult:

g	df of MS_E
3	2
4	6
5	12

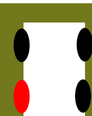
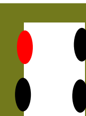
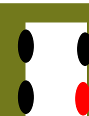
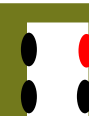
Latin Squares

- Just because the design contains the word “square” doesn’t mean that the physical layout of the experiment has to be a square.
- Often, one blocking factor is **time**: Think of testing 5 different machines (A, B, C, D, E) on 5 days with 5 operators (response: yield of machine):

Operator					
Mon	E	B	C	A	D
Tue	B	D	E	C	A
Wed	A	C	D	B	E
Thu	C	E	A	D	B
Fri	D	A	B	E	C

Graeco Latin Squares

- What if we have one more blocking criterion?
- Use so called **Graeco Latin Squares** (if applicable).
- Take a Latin Square and superimpose it with another block factor, denoted by Greek letters (here: think of driver)

Car				
	$A\alpha$	$B\gamma$	$C\delta$	$D\beta$
	$B\beta$	$A\delta$	$D\gamma$	$C\alpha$
	$C\gamma$	$D\alpha$	$A\beta$	$B\delta$
	$D\delta$	$C\beta$	$B\alpha$	$A\gamma$

Graeco Latin Squares

- The Latin letters occur once in each row and column
- The Greek letters occur once in each row and column

Latin squares

- In addition: each Latin letter occurs exactly **once** with each Greek letter.
- Use **main effects model**

$$Y_{ijkl} = \mu + \underbrace{\alpha_i}_{\text{treatment}} + \underbrace{\beta_j}_{\substack{\text{Block factor 1} \\ \text{(rows)}}} + \underbrace{\gamma_k}_{\substack{\text{Block factor 2} \\ \text{(columns)}}} + \underbrace{\delta_l}_{\substack{\text{Block factor 3} \\ \text{(Greek letters)}}} + \epsilon_{ijkl}$$

to analyze data.

More General Situations

- In practice, (Graeco) Latin Squares are often **impractical** due to the **very restrictive assumptions** on the **number of levels** of the involved treatment and block factors.
- E.g., think of the car tire example with 7 instead of 4 tire treatments.
- Or going back to the intro example: What if we wanted to compare **three** different eye-drops?
- This will lead us to **balanced incomplete block designs (BIBD)**, see later.

General Rules for Analyzing Block Designs

- As we have seen, we treat block factors just as other factors in our model formulas.
- Typically, a block effect is assumed to be **additive** (i.e., main effects only).
- Block factors are **not** tested but they can be examined with respect to efficiency gain.
- ANOVA table and df's are "as usual".
- A possible interaction between block and treatment factor(s) is difficult to sell.