Q: What is the right error?

Review questions:
  What's the treatment design?
  What's the observational unit?
  What's the experimental unit?

Key issue: 2 sizes of eu
  pasture: randomly assigned to grazing system
  heifer: randomly assigned to implant type

Appropriate analyses need to account for both eu's

Previous analysis was wrong
  only has one size (heifer)
  Effectively ignores any variation among pastures
  Said another way: Assumes heifers are independent, even those in the same pasture
    Nope: ICC = 0.65

Vocabulary

Reminder:
  eu and ou
    experimental unit: "thing" randomly (and independently) assigned to a treatment
    observational unit: "thing" contributing one row of data
  treatment and experimental designs
    treatment: what is done to an eu. Often combination of factor levels
    treatment design: Choice of what to be done to eu's, examples:
      2 way complete factorial
      "L" design
    experimental design: how treatments are randomly assigned to eu's, examples:
      CRD, i.e. randomly assigned to collection of eu's
      RCBD, i.e., randomly assigned within blocks
      Latin Square: randomly assigned subject to row and column restrictions

Split plot design:
  A common name for a design with two sizes of eu.
  Split plot is especially common in agronomic / biological situations
  And for randomized experiments
    Comes from RA Fisher's agricultural background
    Main plot: "larger" eu, e.g. pasture, that is then divided into smaller pieces
    Split plots: "smaller" eu. e.g. heifer

In social sciences, especially with observational studies, called multilevel studies
The book uses this vocabulary. Traditional to start with the "smallest" unit
Level 1 = split plot = observations
Level 2 = main plot = groups of observations

In engineering, called "hard-to-change factor" designs.
some factors (hard-to-change) require extended time to change the level
other factors can be changed quickly.
In the Ag Eng. example below, changing the machinery takes perhaps an hour.
Changing the combine speed can be done nearly instantaneously.
Randomly assigning all combinations of machinery and speed
would require many changes of machinery.
A natural way to reduce the total study time:
is to run multiple speeds with one set of machinery,
then change the machinery and run all speeds with the new machinery.
Make sure to replicate machinery

Examples of studies with 2 sizes of eu:

Application area	Main plot / treatment	Split plot / treatment
Ag. Eng. (combines)	15 minute run machinery	3 minute run speed
Agronomy	Field Irrigation	row variety
Biochemistry	96 well plate incubation time	individual well dose of chemical
Nutrition	person ethnicity, gender	period diet
Horticulture	water bath root temperature	pot species
Meat science	10lb batch of meat rosemary oil	package of hot dogs radiation dose
Education	class teaching method	student gender

Each size of eu has its own experimental design.

pastures: CRD
heifers: CRD

The most common split plot design
(in agronomy / animal science / ecology / natural resources):
main plot: RCBD
split plot: CRD
You will find ANOVA tables and model equations for this design lots of places

But, many, many other possible combinations
So how can you figure out an appropriate model for all the data from your specific study?

Constructing a model for a split plot design

We will construct two models:
one for the main plots (pastures),
the other for the split plots (heifers),
then knit them together.

Main plots:
ignore split plot treatment (implant),
observations are now subsamples so mentally average heifers within a pasture.
12 rows of data.
ou is now the pasture.
Write out the design for 12 pastures:
2 components: grazing treatment, and error = pasture(grazing).
pastures are nested in grazing
no connection between pasture 1 in continuous and pasture 1 in rotation
Here's the skeleton ANOVA table for the main plot part of the design.

Source	df
Grazing	2
Error = pasture(Grazing)	9
c.total	11

You need to figure out what identifies each unique main plot
The error is the pooled variability of main plots within each main plot treatment
If you analyze main plot means, the error is included automatically.
We need to name it in terms of other main plot characteristics
because we will work with all observations,
but still need to identify the main plot error
pasture(Grazing) is a random effect
Because it's the error and error terms are random effects

Split plot part of the design:
Each of the 12 pastures is a block w.r.t implants and heifers
I call these "mini block"s
Write out the split part of the design.
Here CRD with one treatment factor (implant):

Source	df	
"mini block"	11	= 12 - 1
Implant	2	
Error	94	= 108 - (1 + 11 + 2)

Now combine main and split parts of the design.
the main plot df sum to 11, the same df as mini-blocks.
If they don't equal, there is a mistake somewhere.

Source	df
Grazing	2
Error = pasture(grazing)	9
Implant	2
Error	94

Add the treatment interactions.
These only exist in the combined model.

The result is the skeleton anova for the analysis.

Source	df	
Grazing	2	
Pasture(grazing)	9	Main plot error
Implant	2	
Grazing*Implant	4	
Error	90	Split plot error

The model equation corresponds to this skeleton ANOVA

$$\begin{aligned} Y_{ijkl} &= \mu + \alpha_j + \gamma_{ij} + \delta_k + \alpha\delta_{jk} + \varepsilon_{ijkl}, \\ \gamma_{ij} &\sim N(0, \sigma^2_{pasture}), \\ \varepsilon_{ijkl} &\sim N(0, \sigma^2_{heifer}), \end{aligned} \tag{1}$$

where:
j is the grazing treatment and α_j is its effect,
ij identifies each pasture, and γ_{ij} is the variability between pastures.
k is the type of implant,
$\alpha\delta_{jk}$ is the interaction between grazing and implant,
l identifies the heifer within a pasture, so $ijkl$ identifies each heifer
and ε_{ijkl} is the variability between heifers within a pasture and implant

Does accounting for the split plot change the results? Yes

Variance components:

Original MSE	0.028
Pastures	0.020
Heifers w/i pastures	0.011
Total variance	0.031

F tests using split-plot model:

Source	F	p value
Grazing	0.61	0.56
Implant	4.22	0.018
G*I	0.60	0.66

Mean (se), compared with original analysis

Factor	Group	Original	Split plot
Grazing	C	1.59 (0.028)	1.59 (0.074)
	R	1.69 (0.028)	1.69 (0.074)
	S	1.58 (0.028)	1.58 (0.074)
Difference		(0.039)	(0.104)
Implant	N	1.58 (0.028)	1.58 (0.045)
	A	1.65 (0.028)	1.65 (0.045)
	B	1.62 (0.028)	1.62 (0.045)
Difference		(0.039)	(0.025)