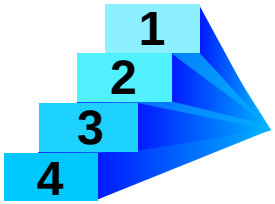
Statistical Techniques II

EXST7015

Treatment Arrangements



Treatment Arrangements

- **Sometimes the treatment simply consists of a list of levels that the investigator is interested in examining.**
- **We will term this type of treatment arrangement "*a priori*" an treatment arrangement.**
- **These are often fixed treatment levels that the investigator wants to examine, but they may be random.**

Tmt Arrangement *(continued)*

- There are several other possibilities.
  - Cross classified (factorial, two-way ANOVA)
    - Like treatments with blocks, two treatments can be cross-classified.
  - Nested treatment arrangement
-

Tmt Arrangement *(continued)*

- **A factorial arrangement of treatments occurs when we have two (or more) treatments of interest arranged such that each level of the first occurs with each level of the second. All possible combinations of the two treatments exist.**

Tmt Arrangement *(continued)*

- **Examples -**
  - **Examine the effect of three dietary supplements (a, b & c) on weight gain for males and females. Each sex gets the same three diets (6 combinations)**
  - **Examine the effectiveness of three pre-emergence herbicides and four post-emergence herbicides. All of the 12 combinations exist, each treatment may have a null treatment as a control.**

Tmt Arrangement *(continued)*

- **The other type of treatment arrangement is the nested treatment arrangement. Nested treatment arrangements occur when each level of some treatment occurs in combination with some other treatment, but the levels of the second treatment are not the same for each level of the first treatment.**

Tmt Arrangement *(continued)*

- **Examples -**
  - **Examine the effect of three dietary supplements on weight gain for males and females. Each sex gets three diets, but the diets are different for males (a, b & c) and females (d, e & f).**
  - **Examine the effectiveness of four post-emergence herbicides on three different crops. The approved post emergence herbicides are not the same for the three crops.**

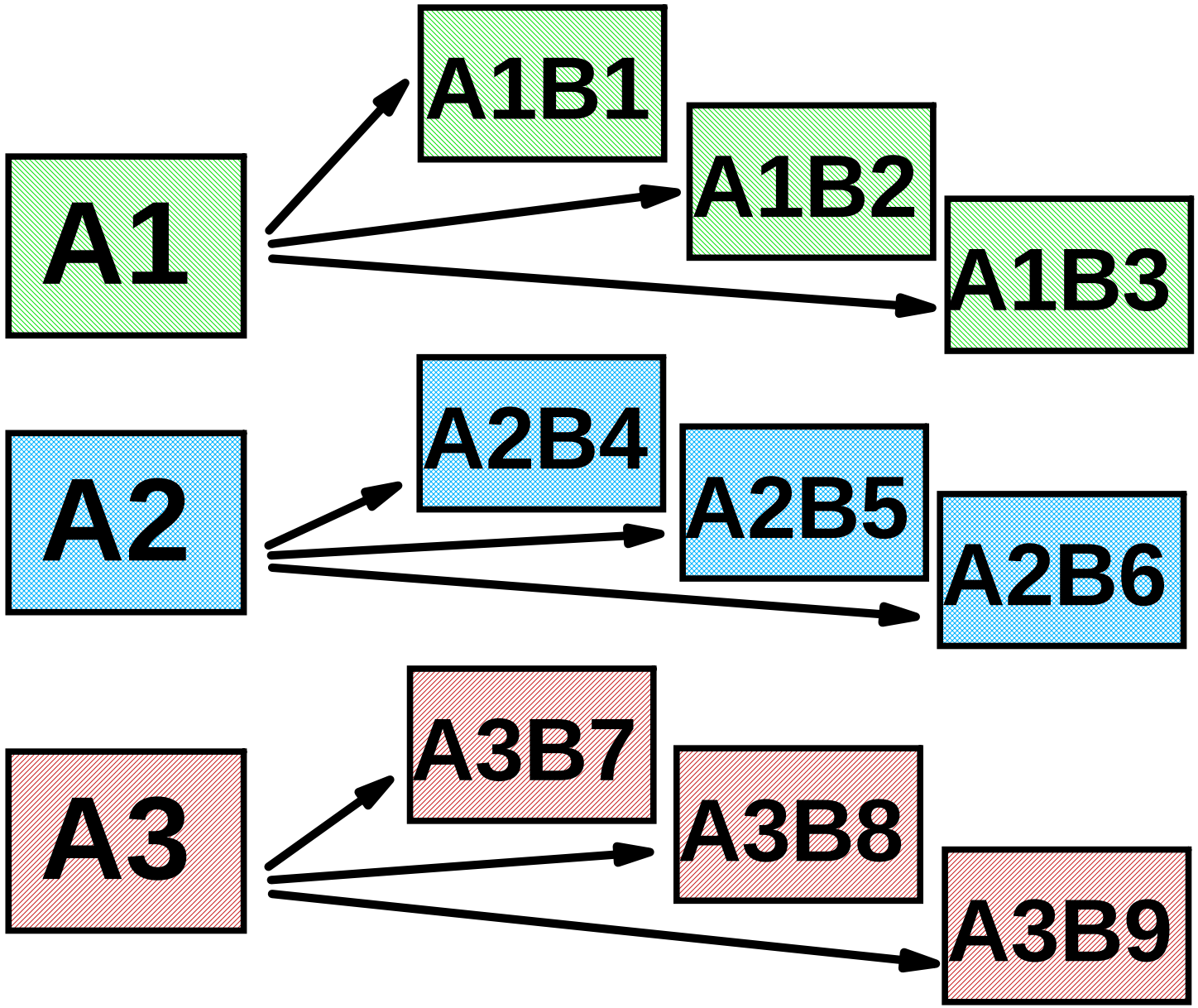
Tmt Arrangement *(continued)*

- **Factorial**

	A1	A2	A3
B1	a1b1	a2b1	a3b1
B2	a1b2	a2b2	a3b2
B3	a1b3	a2b3	a3b3
B4	a1b4	a2b4	a3b4

Tmt Arrangement *(continued)*

- **Nested**



Tmt Arrangement *(continued)*

- **Nested treatment arrangements are not too common. They can occur.**
- **For example, if we wanted to test for differences in attendance at State Parks. We choose 4 parks in TX, 5 in LA and 3 in MS. There is no "match" for the parks between states. We could measure attendance on randomly chosen dates and our model would be**
  - **MODEL Y = STATE PARK(STATE);**

Tmt Arrangement *(continued)*

- **Another example. Suppose we wanted to test for the effectiveness of various commonly used herbicide on major crops in LA by examining dollar value per acre. We choose crops (Cane, Rice, Soy and Corn). We select representative fields at random and treat with an appropriate herbicide.**

Tmt Arrangement *(continued)*

- **Unfortunately, the same herbicides are not used on these crops. Corn and Cane are grasses, and the herbicides target "broadleaf" plants. Soybean is a broadleaf plant, so it requires different herbicides. Rice is grown in water and requires special herbicides. So, each crop has it's own suite of herbicides.**
- **MODEL Y = CROP HERB(CROP);**

Tmt Arrangement *(continued)*

- **Factorial designs are VERY common, popular and highly recommended.**
- **This treatment arrangement also has some unique properties and interpretations (especially interactions)**
- **We will concentrate on this treatment arrangement.**

Treatment Interactions

- **The one really different thing about treatments is that we are interested in them (as opposed to blocks and nested error terms).**
- **We may want to test the individual levels. This will be our major topic following treatment arrangements.**
- **We are also likely to be interested in the INTERACTION!**
-

Tmt Interactions *(continued)*

- **This is new and VERY important. Block & treatment interactions are "error", and not of interest.**
- **However, treatment interactions measure how consistent one treatment is across the levels of another. This is interesting and important. It cannot be ignored.**
-

Tmt Interactions *(continued)*

- **Look at the table below. What value belongs in the missing cell?**

	T2 a	T2 b	T2 c
T1 a	3	5	7
T1 b	6	8	10
T1 c	2	????	6
T1 d	5	7	9

Treatment Interactions *(continued)*

- **The missing value is 4!!! How did you know?**

	T2 a	T2 b	T2 c	Mean	Effect
T1 a	3	5	7	5	-1
T1 b	6	8	10	8	2
T1 c	2	4.00	6	4	-2
T1 d	5	7	9	7	1
Mean	4	6	8	6	
Effect	-2	0	2		

Tmt Interactions *(continued)*

- **Could it be 16?**

	T2 a	T2 b	T2 c	Mean	Effect
T1 a	3	5	7	5	-2
T1 b	6	8	10	8	1
T1 c	2	**16**	6	8	1
T1 d	5	7	9	7	0
Mean	4	9	8	7	
Effect	-3	2	1		

Tmt Interactions *(continued)*

- **Could it be 1?**

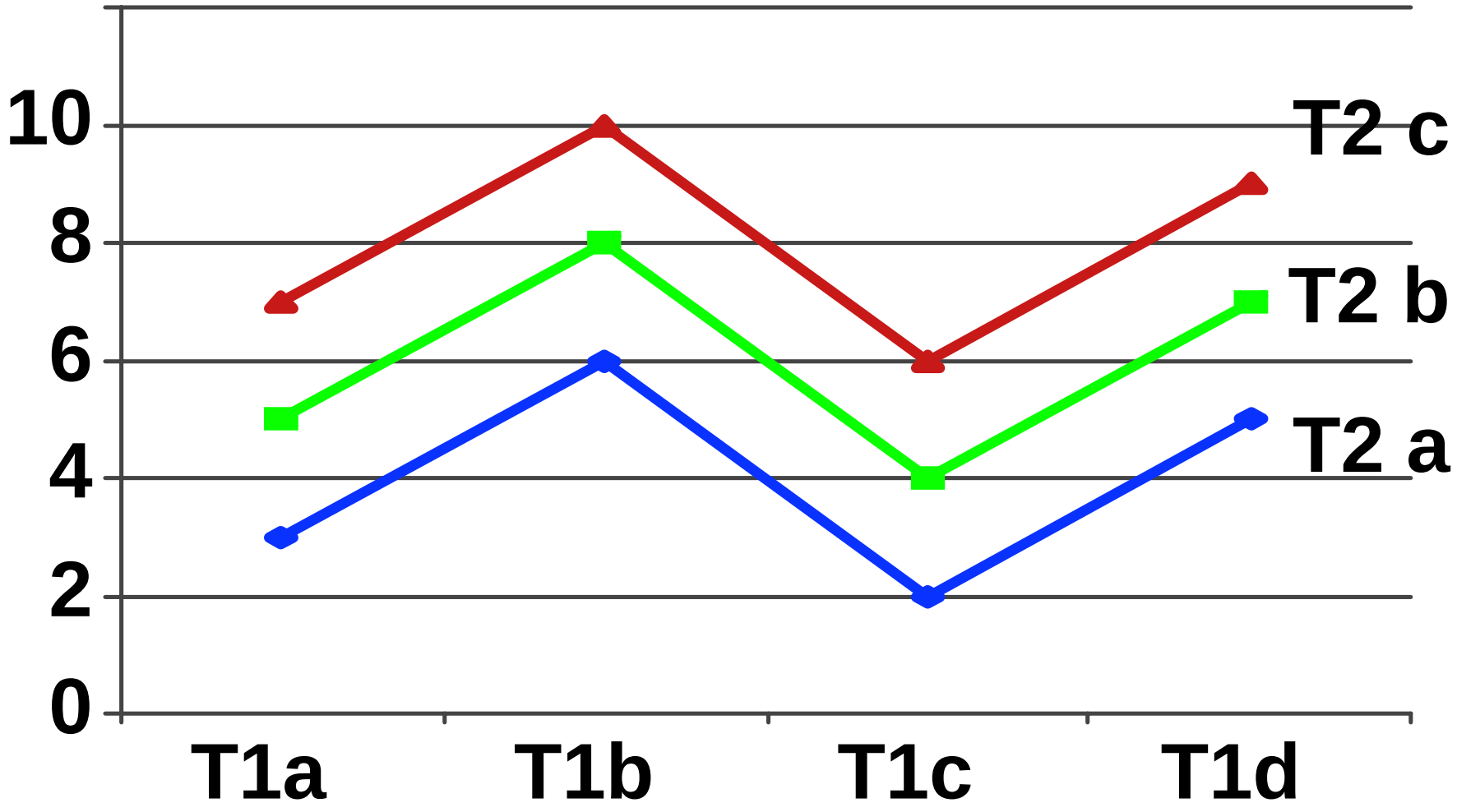
	T2 a	T2 b	T2 c	Mean	Effect
T1 a	3	5	7	5	-0.75
T1 b	6	8	10	8	2.25
T1 c	2	1	6	3	-2.75
T1 d	5	7	9	7	1.25
Mean	4	5.25	8	5.75	
Effect	-1.75	-0.5	2.25		

Tmt Interactions *(continued)*

- **Of course it can be any value it wants to be. There are no restrictions. However, if it is any value other than 4, then there is an interaction.**
- **If we plot the data and there is no interaction, the lines connecting the means should be parallel.**

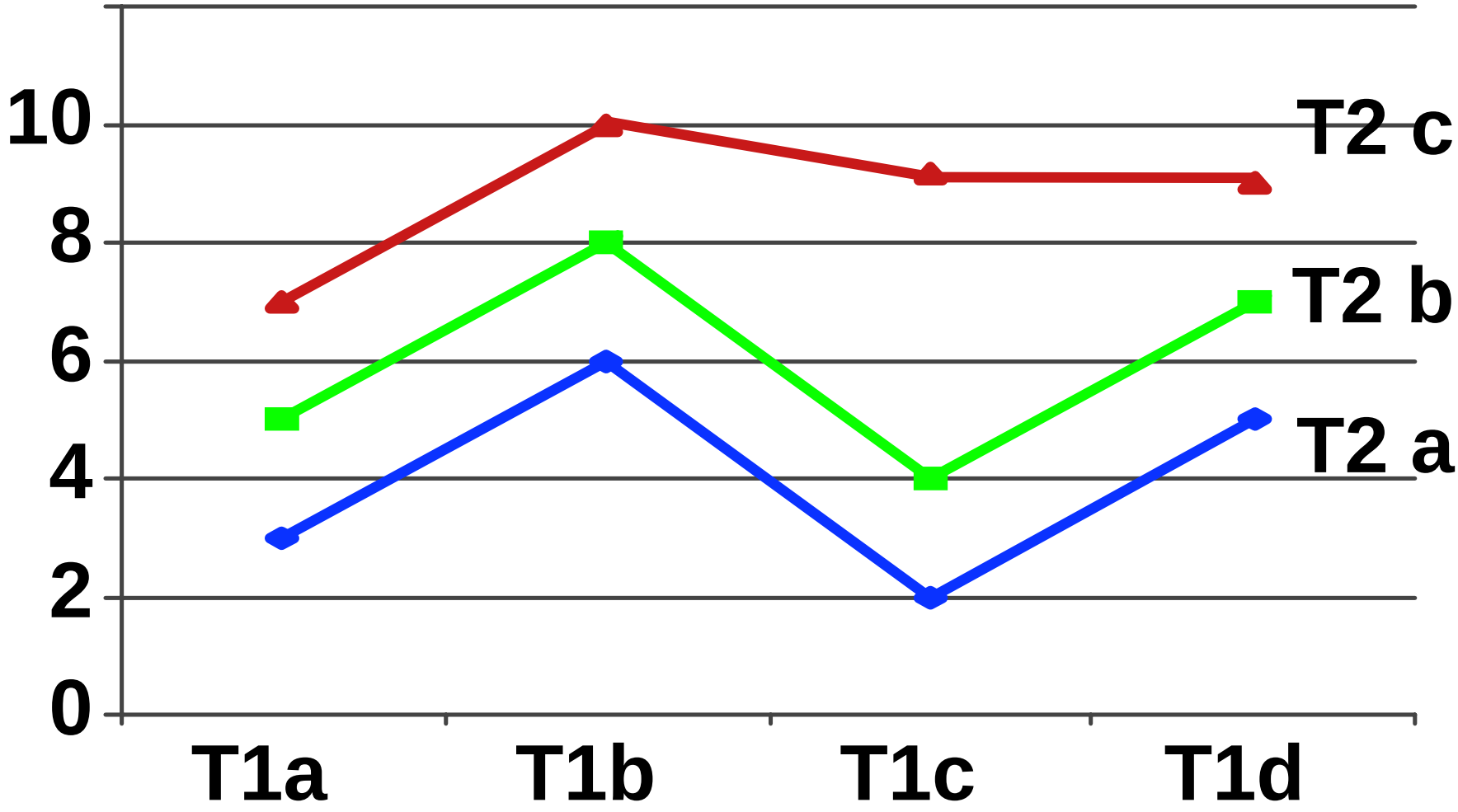
Tmt Interactions *(continued)*

- **No interaction.**



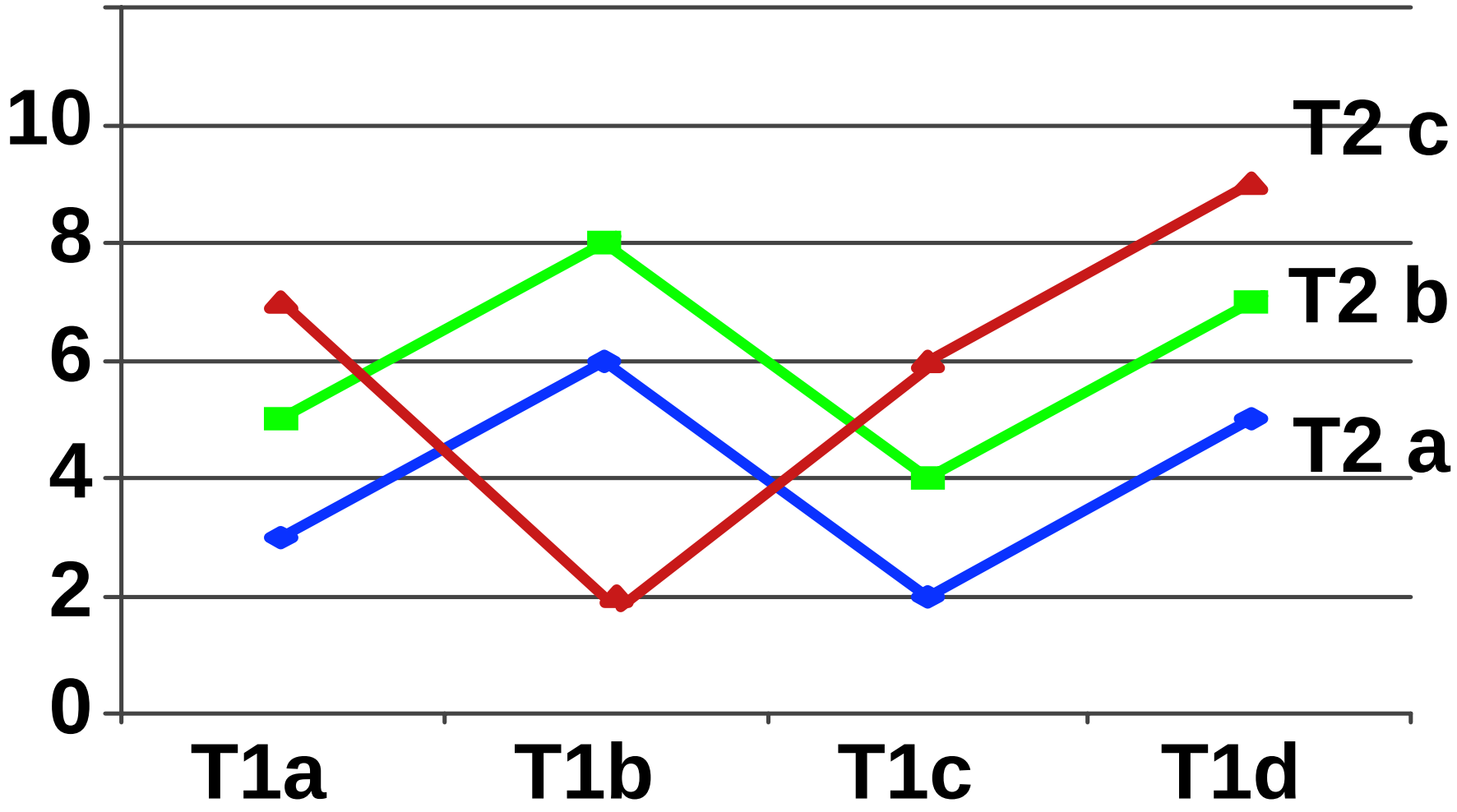
Tmt Interactions *(continued)*

- **If an interaction is present the lines are not parallel.**



Tmt Interactions *(continued)*

- **And may even cross.**



Tmt Interactions *(continued)*

- **So how do we interpret an interaction?**
- **If there is no interaction the behavior of the treatments is <u>consistent.</u> The means increase and decrease by the same amount.**
- **If there is an interaction, increases and decreased in the means are unpredictable and cannot be foreseen by the main effects.**

Tmt Interactions *(continued)*

- **Of course, in practice no lines are ever EXACTLY parallel. The means never increase and decrease by EXACTLY the same amount.**
- **So we need a statistical test to determine if the departure is statistically meaningful; if the interaction is "significant".**
- **No problem. We make the interaction a source in our model and test it.**

Tmt Interactions *(continued)*

- **But note one key factor. Blocks had interactions with treatments. We calculated those, and tested if we wanted**
- **However, interactions with blocks are usually not of interest, they are simply a measure of random error.**
- **Treatment interactions are of great interest, because if our treatments are not consistent we must know how they change to make our conclusions.**

Tmt Interactions *(continued)*

- **We will be especially concerned with factorial treatment arrangements. These are very common.**
- **R. A. Fisher pointed out that these designs had "hidden replication".**
- **For example, suppose we have a 4 by 5 factorial treatment arrangement with 2 replicate observations in each of the 20 treatment combinations.**

Tmt Interactions *(continued)*

- **Number of replicates per treatment combination. Note that treatment mean comparisons have more reps.**

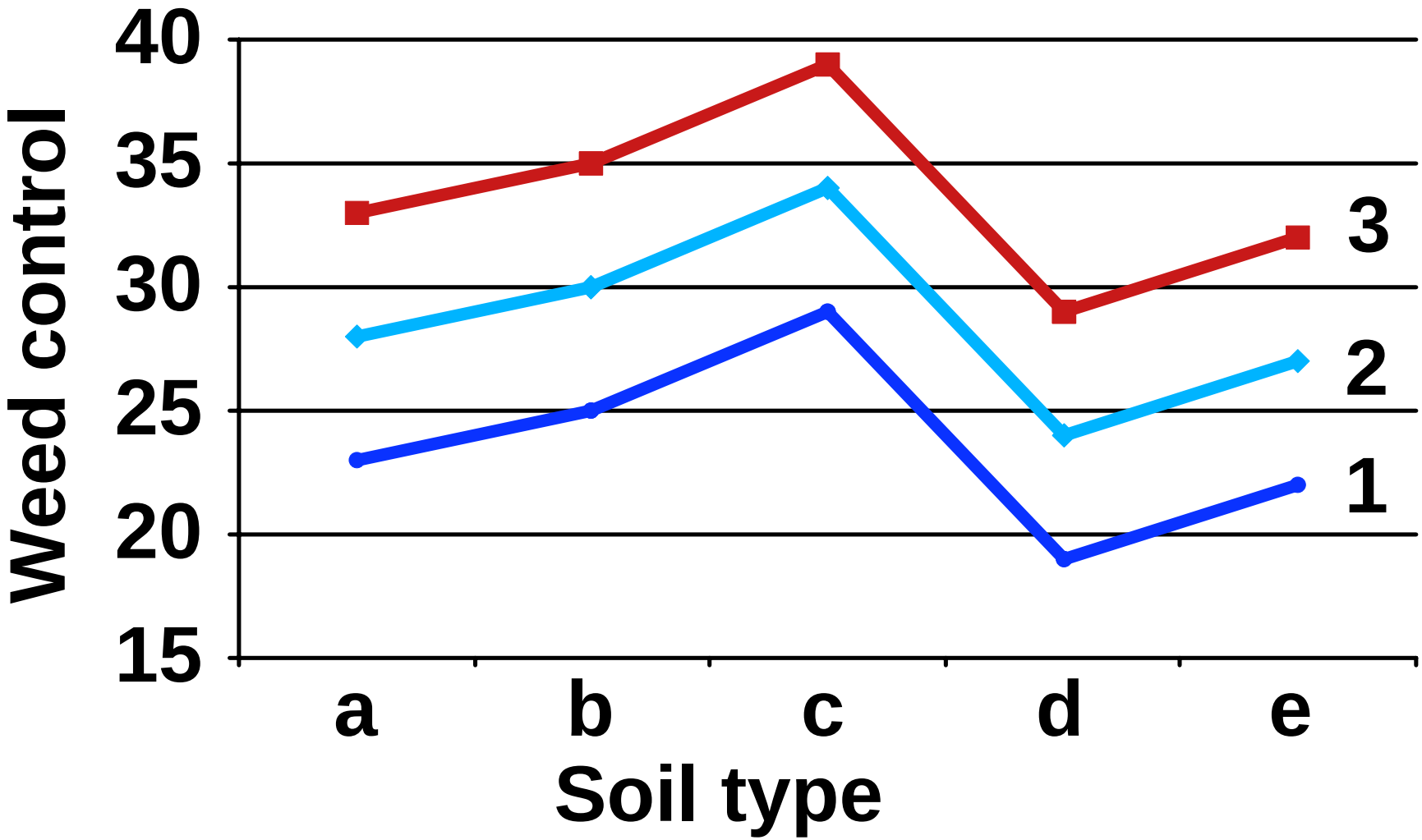
	T2 a	T2 b	T2 c	T2 d	Sum
T1 a	2	2	2	2	8
T1 b	2	2	2	2	8
T1 c	2	2	2	2	8
T1 d	2	2	2	2	8
T1 e	2	2	2	2	8
Sum	10	10	10	10	40

Tmt Interactions *(continued)*

- **How important are interactions? If we have significant main effects and significant interactions, can we ignore one?**
- **Lets examine some graphs for an experiment. Suppose we are trying to determine the best of three herbicides (1, 2, 3) to control weeds on five soil types (a, b, c, d and e).**

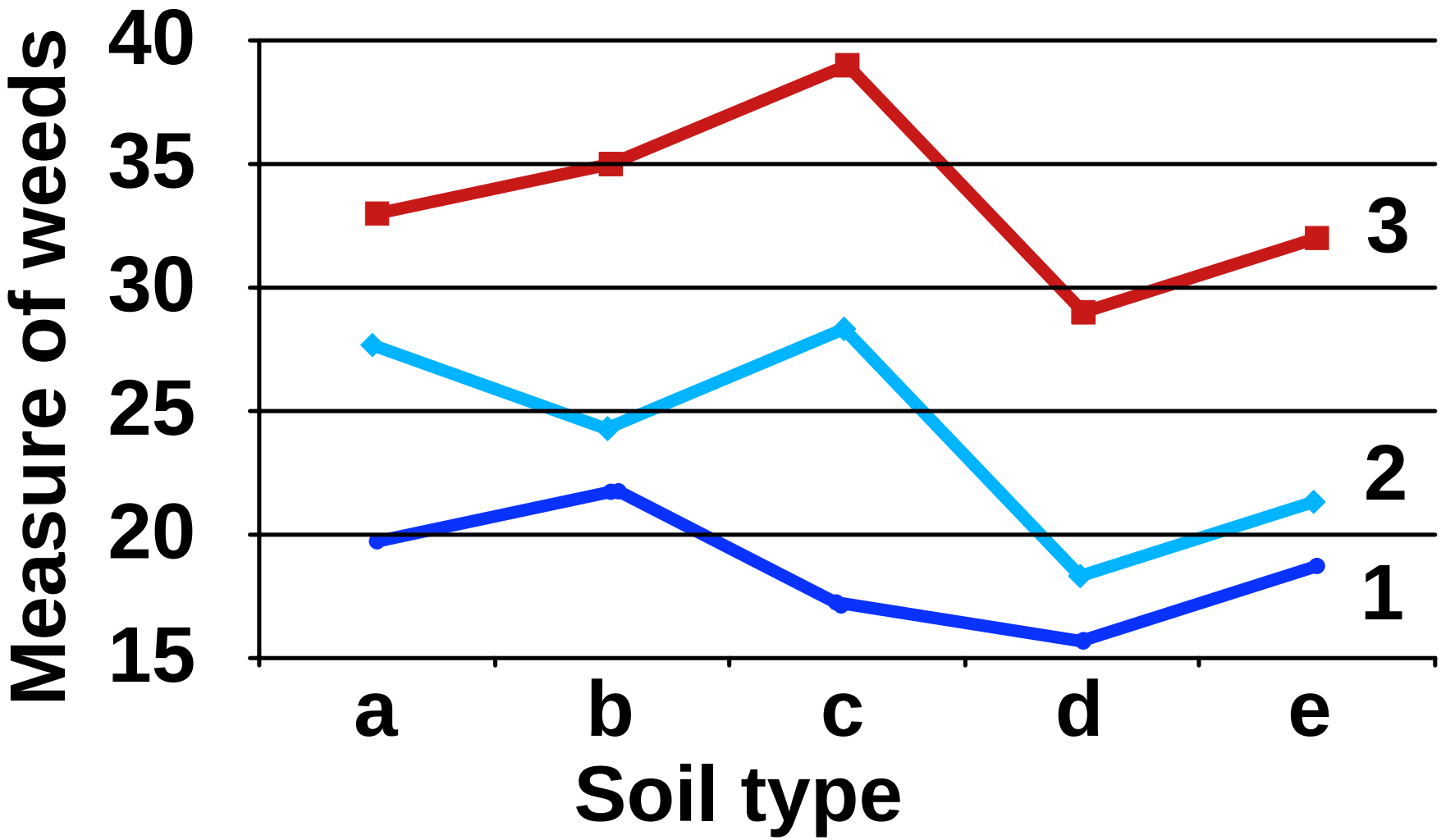
Tmt Interactions *(continued)*

- **No interaction. Herbicide 3 is best on every soil type.**



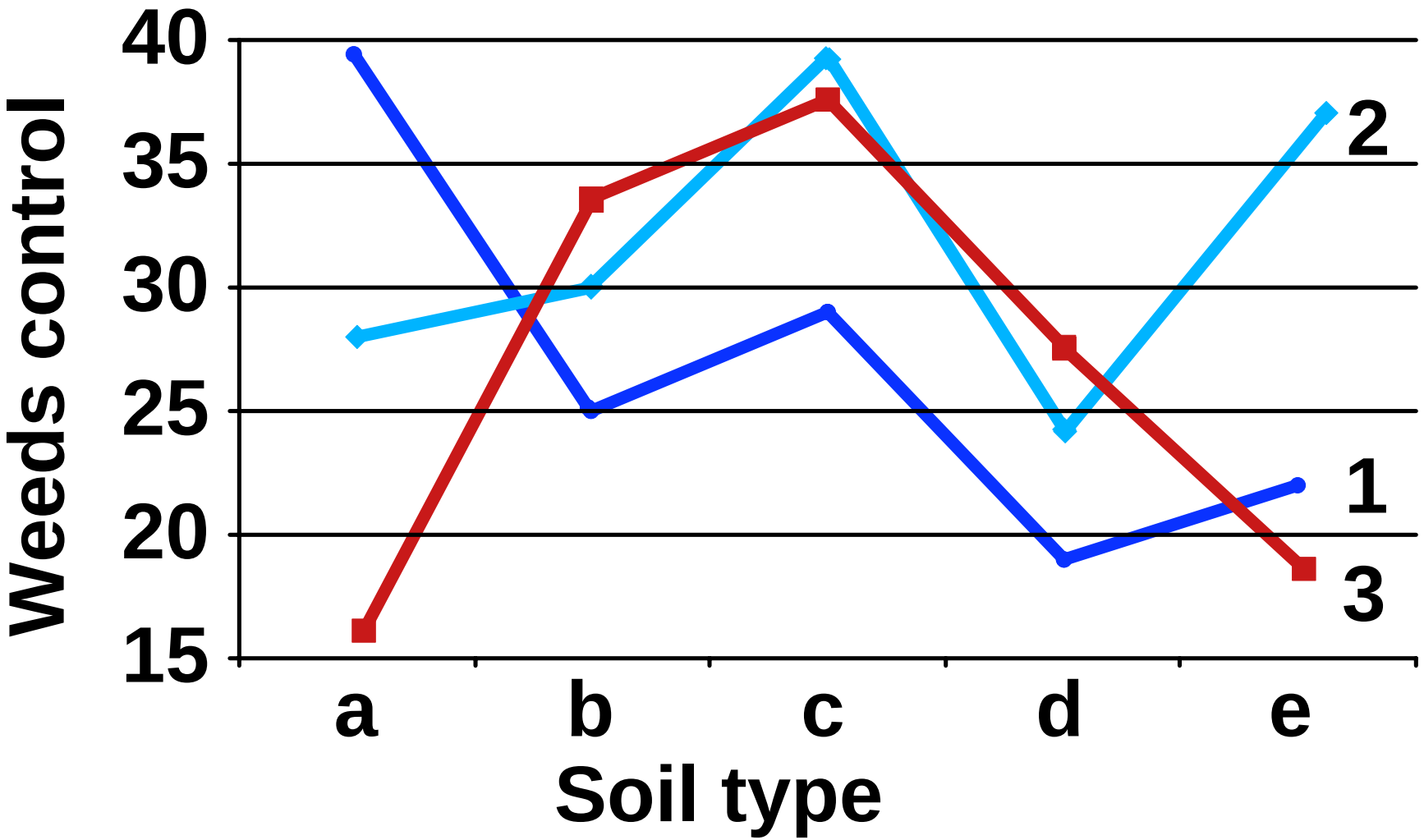
Tmt Interactions *(continued)*

- **Interaction present. Herbicide 3 is still best on every soil type.**



Tmt Interactions *(continued)*

- **Interaction. Which herbicide is best?**

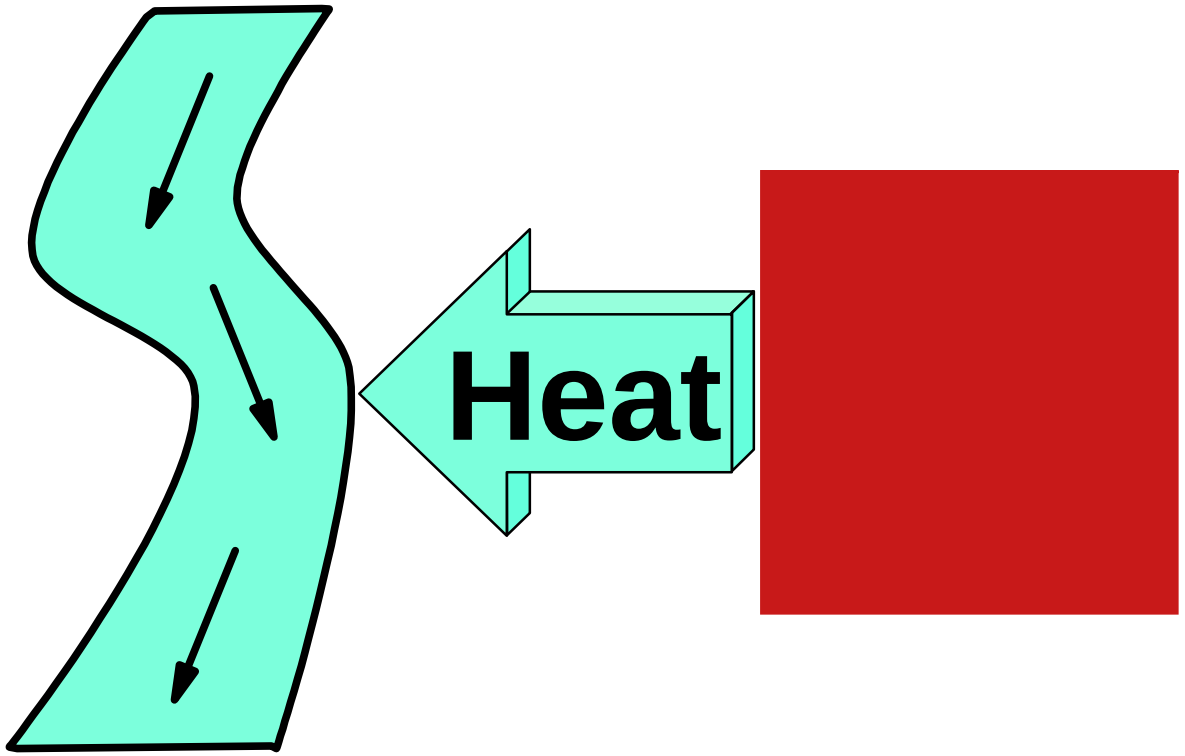


Tmt Interactions *(continued)*

- **Sometimes the interaction is significant, but one main effect stands out anyway.**
- **Other times the interaction is so strong that that the best results for each treatment 1 depends on the combination with treatment 2.**
- **The bottom line. Unlike treatment and block interactions, treatment interactions are not "assumed" away! Test them, and be prepared to examine them.**
- **Lets look at another case of interaction.**

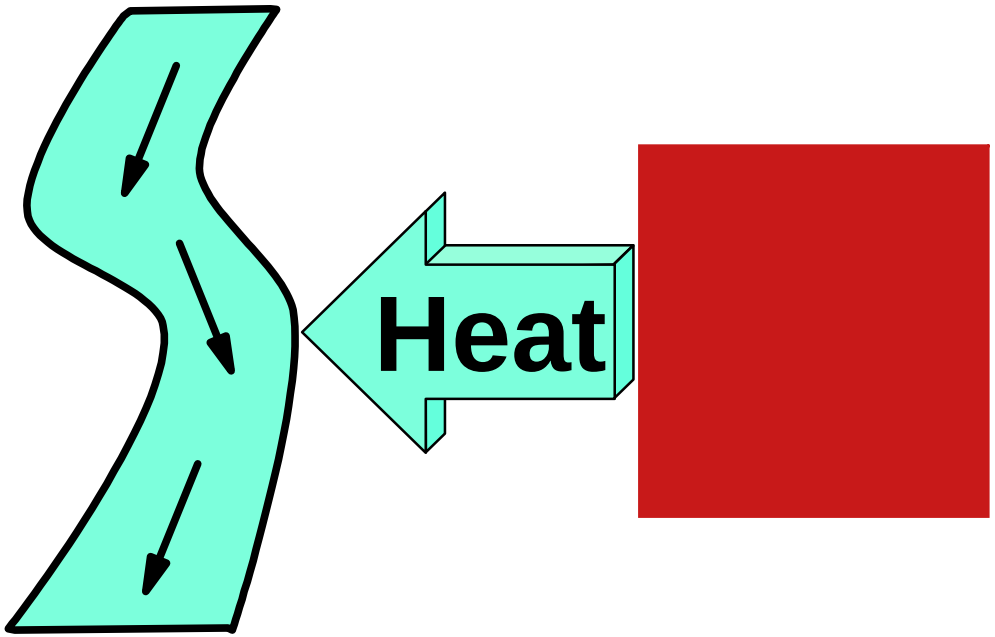
Tmt Interactions *(continued)*

- **Environmental impact measurement. Suppose we are constructing a power plant, and plan to dump cooling water into a river.**



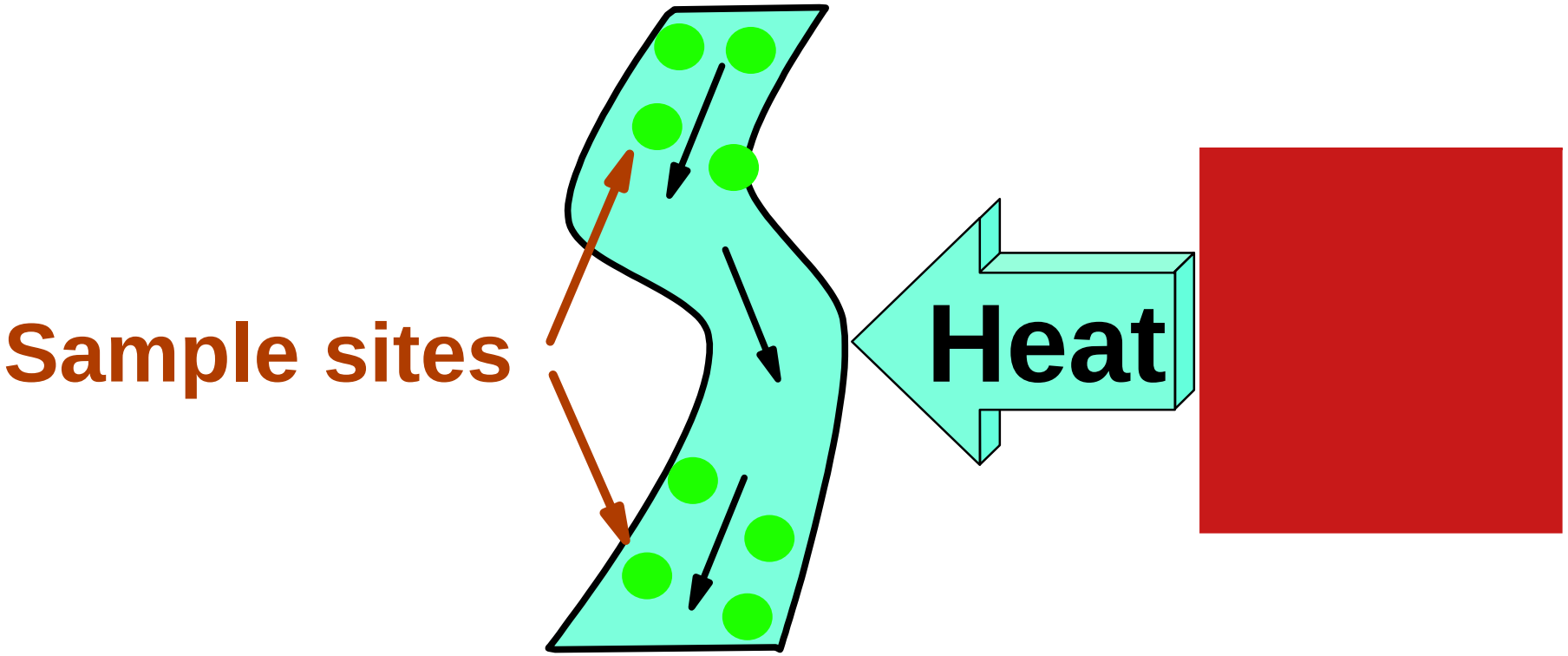
Tmt Interactions *(continued)*

- **We want to determine if there is an impact on the growth of Channel catfish in the river. We measure growth by sampling otoliths from small catfish.**



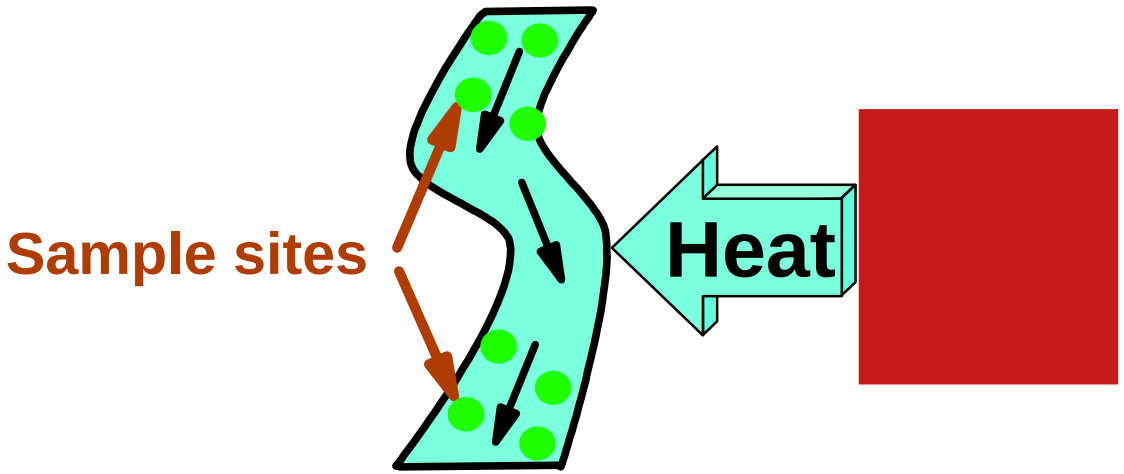
Tmt Interactions *(continued)*

- **We sample above the power plant and below the power plant to see if the growth is different. "Upstream - downstream" should detect impact.**



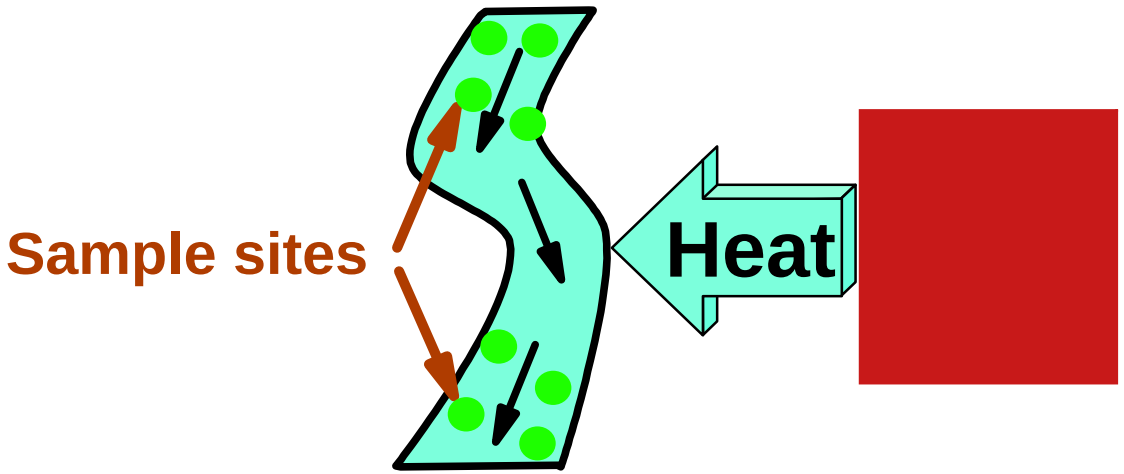
Tmt Interactions *(continued)*

- **But then we are told that this will mean nothing. Growth downstream has always been different from growth upstream. Better habitat, nutrition, etc.**



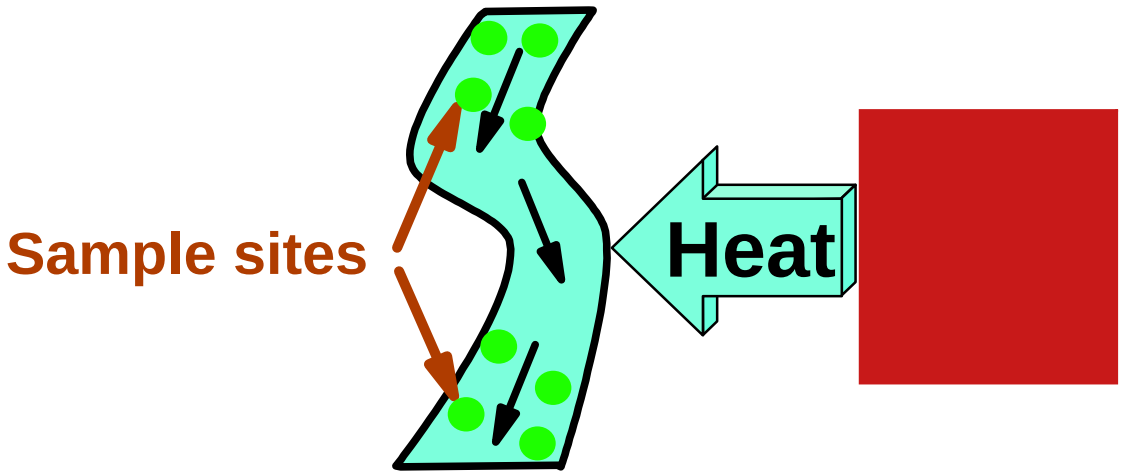
Tmt Interactions *(continued)*

- **So we try another tactic, we sample for a few years before the plant goes into operation and for a few years after the plant goes into operation. Surely "before-after" will detect impact.**



Tmt Interactions *(continued)*

- **Not necessarily. Maybe the years before were wet "El Niño" years and the years after were dry "La Niña" years. Or maybe something happens way upstream at the same time our power plant is finished! Then any observed changes would not be due to our power plant.**



Tmt Interactions *(continued)*

- **So how do we sample impact? We must detect an interaction.**
- **In this case the ONLY term of real interest for detecting impact is the interaction. The main effects are not useful in detecting impact!!**

	Before	After
Upstream	24	27
Downstream	32	25

Tmt Interactions *(continued)*

- **Terminology.**
  - **Additivity - Take a cell in a factorial treatment arrangement with an overall mean of 10.**
  - **If the EFFECT for treatment 1 is "5" and the effect for treatment 2 is "-2", the value in the cell should be overall mean + effect 1 + effect 2 = 10 + 5 - 3 = 12.**
  - **We get the cell by ADDing the effects. This is additivity.**
  - **This will not work if there is an interaction.**

Tmt Interactions *(continued)*

- **Interactions are sometimes referred to as <u>tests of additivity</u>.**
- **For the model**
  - $Y_{ij} = \mu + \tau_{1i} + \tau_{2j} + \tau_1\tau_{2ij} + \varepsilon_{ijk}$
- **There is no interaction if**
  - $Y_{ij} - \bar{Y}_{i.} - \bar{Y}_{.j} + \bar{Y}_{..} = 0$
    - **Note that this calculation was done on means ($\bar{Y}$), not effects (τ).**

Tmt Interactions *(continued)*

- For cell T1c, T2b
- 4-6-4+6 = 0, no interaction

	T2 a	T2 b	T2 c	Mean	Effect
T1 a	3	5	7	5	-1
T1 b	6	8	10	8	2
T1 c	2	4	6	4	-2
T1 d	5	7	9	7	1
Mean	4	6	8	6	
Effect	-2	0	2		

Tmt Interactions *(continued)*

- **Multiplicative models (Chi square analysis and log-linear models).**
- **Drug A saves 50 percent of fish with a certain fungus.**
- **Drug B saves 50 percent of fish with the same fungus.**
- **Giving Drug A and Drug B together should save what percent?**
  - **100%, 75%, 50%, 25%, 0%?**

Tmt Interactions *(continued)*

- **For an additive model, the answer is 100%.**
- **If we have 100 fish and Drug A and Drug B both save 50 of 100, then all fish will be saved.**
- **In a proportional or multiplicative model, Drug A saves 50%, adding Drug B will save 50% of the remaining fish for a total of 75%.**
-

Tmt Interactions *(continued)*

- **We will not be working with these models, but you should be aware of them.**
- **Chi square tests of independence test for proportional interactions, not additive interactions.**
- **Log-linear models (which we saw for regression) can be applied to ANOVA (by taking the log of Y_i), and test for multiplicative effects.**

MS for Treatments

- **Expected mean squares for treatments, nested or cross-classified, work exactly the same as for nested error terms or cross-classified blocks.**
- **The only difference is that treatments may well be both fixed, while blocks are random. This will be the only real new consideration.**
-

EMS for Tmts *(continued)*

- **The source table for this CRD, with sources, degrees of freedom and EMS is given below. The treatments are *a priori*, either fixed or random.**

Source	d.f.	SS	MS	EMS
Tmt	t-1	SSTmt	MSTmt	$\sigma^2+n\sigma^2_\tau$
Error	t(n-1)	SSE	MSE	σ^2
Total	tn-1	SSTotal		

EMS for Tmts *(continued)*

- **CRD with fixed effect treatments.**

Source	d.f.	SS	EMS
Tmt	t-1	SSTmt	$\sigma^2 + n\Sigma\tau^2_i/_{(t-1)}$
Error	t(n-1)	SSE	σ^2
Total	tn-1	SSTotal	

EMS for Tmts *(continued)*

- The design below has 4 nested levels. The top line is a treatment, the bottom an error. The two others could be either.
- Nested treatments are not common.

Source	d.f.	EMS
Tmt	t-1	$\sigma^2+n\sigma^2_\gamma+ns\sigma^2_\delta+nsp\Sigma\tau^2_i/_{(t-1)}$
B(Tmt)	t(p-1)	$\sigma^2+n\sigma^2_\gamma+ns\sigma^2_\delta$
C(B*Tmt)	tp(s-1)	$\sigma^2+n\sigma^2_\gamma$
Rep(C*B*tmt)	tps(n-1)	σ^2
Total	tpsn-1	

EMS for Tmts *(continued)*

- **The source table for an RBD.**

Source	d.f.	SS	MS	EMS
Tmt	t-1	SSTmt	MSTmt	$\sigma^2+n\sigma^2_\tau$
Block	b-1	SSBlk	MSBlk	$\sigma^2+n\sigma^2_\beta$
Error	(t-1) (b-1)	SSE	MSE	σ^2
Total	tn-1	SSTotal		

EMS for Tmts *(continued)*

- **The source table for an Factorial. Do not ever do the experiment below!!! There is no test of the interaction because there is no error term!**

Source	d.f.	SS	EMS
Tmt 1	t_1-1	SSTmt1	$\sigma^2+n\sigma^2_{\tau 1}$
Tmt 2	t_2-1	SSTmt2	$\sigma^2+n\sigma^2_{\tau 2}$
T1 * T2	$(t_1-1)(t_2-1)$	SSInter	$\sigma^2_{\tau 1\tau 2}$
Total	tn-1	SSTotal	

EMS for Tmts *(continued)*

- **We can do experiments with one block and one treatment, because the interaction is an error term. We cannot do experiments with just two treatments. We need replicate experimental units within treatments to test for interactions.**
- **The previous "bad" model would be**
  - $Y_{ij} = \mu + \tau_{1i} + \tau_{2j} + \tau_{1i}\tau_{2j}$
- **The "good" model would be**
  - $Y_{ij} = \mu + \tau_{1i} + \tau_{2j} + \tau_{1i}\tau_{2j} + \varepsilon_{ijk}$

EMS for Tmts *(continued)*

Source	d.f.	SS	EMS
Tmt 1	t1-1	SSTmt1	$\sigma^2+n\sigma^2_{\tau1\tau2}+nt_2\sigma^2_{\tau1}$
Tmt 2	t2-1	SSTmt2	$\sigma^2+n\sigma^2_{\tau1\tau2}+nt_1\sigma^2_{\tau2}$
T1 * T2	(t1-1) (t2-1)	SST1T2	$\sigma^2+n\sigma^2_{\tau1\tau2}$
Error	tb(n-1)	SSE	σ^2
Total	tbn-1	SSTotal	

EMS for Tmts *(continued)*

- **The EMS for treatments work the same as for blocks and treatments.**
- **There is however one really big consideration remaining.**
- **Blocks are random, treatments are either random or fixed, and this will affect our tests.**
- **Note that on the preceding page the treatment interaction was actually the error term for the treatment main effects.**

EMS for Tmts *(continued)*

- **This is true, and it is not a problem. SAS GLM and MIXED will do the appropriate tests as long as you specify that the treatments are random.**
- **If one treatment is fixed and one is random, nothing changes for this example, since the interaction of a random effect and a fixed effect is still random.**
- **The test of the main effects is still done with the interaction.**

EMS for Tmts *(continued)*

Source	d.f.	SS	EMS
Tmt 1	t1-1	SSTmt1	$\sigma^2+n\sigma^2_{\tau1\tau2}+nt_2\Sigma\tau^2_i/(\tau1-1)$
Tmt 2	t2-1	SSTmt2	$\sigma^2+n\sigma^2_{\tau1\tau2}+nt_1\sigma^2_{\tau2}$
T1*T2	(t1-1) (t2-1)	SST1T2	$\sigma^2+n\sigma^2_{\tau1\tau2}$
Error	tb(n-1)	SSE	σ^2
Total	tbn-1	SSTotal	

EMS for Tmts *(continued)*

- **HOWEVER, if BOTH effects are fixed, then the interaction is also FIXED.**
- **Fixed effects occur only on their own source line, not in any other sources!**
- **This makes a BIG difference!!!**
-

EMS for Tmts *(continued)*

- **Note that when all treatments are fixed we do not have the interaction as part of the main effect EMS.**

Source	d.f.	SS	EMS
Tmt 1	t1-1	SSTmt1	$\sigma^2 + \cancel{n\sigma^2_{\tau 1 \tau 2}} + nt_2 \Sigma \tau^2_{1i} / (\tau 1 - 1)$
Tmt 2	t2-1	SSTmt2	$\sigma^2 + \cancel{n\sigma^2_{\tau 1 \tau 2}} + nt_1 \Sigma \tau^2_{2j} / (\tau 2 - 1)$
T1 * T2	(t1-1) (t2-1)	SST1T2	$\sigma^2 + n\Sigma(\tau_1 \tau_2)^2_{ij} / (\tau 1 - 1)(\tau 2 - 1)$
Error	tb(n-1)	SSE	σ^2
Total	tbn-1	SSTotal	

EMS for Tmts *(continued)*

- **Now, what is the correct error term for the treatments and interactions?**

Source	d.f.	SS	EMS
Tmt 1	t1-1	SSTmt1	$\sigma^2 + nt_2\Sigma\tau^2_{1i}/(\tau 1-1)$
Tmt 2	t2-1	SSTmt2	$\sigma^2 + nt_1\Sigma\tau^2_{2j}/(\tau 2-1)$
T1 * T2	(t1-1) (t2-1)	SST1T2	$\sigma^2+n\Sigma(\tau_1\tau_2)^2_{ij}/(\tau 1-1)(\tau 2-1)$
Error	tb(n-1)	SSE	σ^2
Total	tbn-1	SSTotal	

EMS for Tmts *(continued)*

- **Right, the experimental error term!**
- **With a random effects model or mixed model, interactions are error terms. However with all effects fixed, the experimental error is the error for both main effects and interactions.**
- **Note that this is what SAS does by default in PROC GLM, so these tests are available by default in some models.**

EMS for Tmts *(continued)*

- **What if you have both experimental error and sampling error?**
- **Now the experimental error must be used, and the "sampling error" is the residual error. You may specify a TEST statement to make the tests, or rely on the PROC GLM random statement with the /test option, or PROC MIXED.**

EMS for Tmts *(continued)*

- **Now, what is the correct error term for the treatments and interactions?**

Source	d.f.	EMS
Tmt 1	t1-1	$\sigma^2 + n\sigma^2_\gamma + nt_2\Sigma\tau^2_{1i}/(\tau 1-1)$
Tmt 2	t2-1	$\sigma^2 + n\sigma^2_\gamma + nt_1\Sigma\tau^2_{2j}/(\tau 2-1)$
T1*T2	(t1-1) (t2-1)	$\sigma^2 + n\sigma^2_\gamma + n\Sigma(\tau_1\tau_2)^2_{ij}/(\tau 1-1)(\tau 2-1)$
E. Error	tb(s-1)	$\sigma^2 + n\sigma^2_\gamma$
S. Error	tbs(n-1)	σ^2
Total	tbn-1	

Missing cells

- **Factorial with missing cells (don't ever do have missing cells, and if you do, don't use Type IV SS in SAS unless you really know what you are doing)!**

	A1	A2	A3
B1	a1b1	a2b1	a3b1
B2	.	a2b2	a3b2
B3	a1b3	a2b3	.
B4	a1b4	a2b4	a3b4

Missing cells *(continued)*

- **And if you are using SAS TYPE IV SS, you probably do not know what you are doing.**
- **Missing cells are not an issue IF THERE IS KNOWN TO BE NO INTERACTION. SAS TYPE III SS basically gives the result assuming no treatment interaction.**
- **If there is an interaction, there is no proper test, the treatments cannot be separated.**

Missing cells *(continued)*

- **If you have missing cells you can take all treatment combinations as a single treatment and do selected contrasts. We will discuss contrasts soon.**
- **I hate experiments with missing cells.**

Tmt Arrangement Examples

- **See SAS output handout**

Summary

- **There are three types of treatment arrangement.**
  - **A priori - very common and relatively simple**
  - **Factorial - the most common and important.**
  - **Nested - not so common, but can occur**
- **A major new development with factorial treatment arrangements is the consideration of interactions.**

Summary *(continued)*

- **We now have a more serious consideration of whether a treatment is FIXED or RANDOM. Selecting the appropriate error term depends on this determination.**
- **Missing cells are a no-no.**
-