Data from a study of food preference:

3 types of protein supplement, control, new liquid, new solid
randomly assigned to 75 men and 75 women, each tasted one. 25 per type.
response is a measure of preference (-3 = absolutely dislike, 3 = wonderful)
6 treatments (3 types, 2 sexes).

	Type			
Sex	old	liquid	solid	average
F	0.24	1.12	1.04	0.80
M	0.20	1.24	1.08	0.84
average	0.22	1.18	1.06	0.82

Notice that:

- Could describe treatments as 2 factors: type with 3 levels and sex with 2 levels
- This is a complete 2 way factorial: all 6 combinations used in the study

Goal: understand and use the Two-way factorial ANOVA table:

Source	d.f.	SS	MS	F	p-value
Sex	1	0.06	0.06	0.04	0.83
Type	2	27.36	13.68	10.12	< 0.0001
Sex*type	2	0.16	0.08	0.06	0.94
Error	144	194.56	1.35		
c.Total	149				

Could ask about differences between the 6 treatments (1 way ANOVA) or ask specific questions:

Treatments have a structure that suggests specific questions:

Averaged over types, is there a difference between sexes?
Averaged over sexes, is there a difference between types?
  Is there a difference between new liquid and new solid?
  Is there a difference between control (old) and average of the new types?
Do men and women react similarly to the types? i.e.
  Is the difference between liquid and solid the same in men and women?
  Is the difference between control and new the same in men and women?

When each cell has the same number of observations (here, 25),
3 ways to construct F tests to answer these questions

1) computing contrast SS
2) comparing models
3) using formulae

You've seen using formulae and comparing models when we discussed 1 way ANOVA; you've also seen estimating answers to specific questions. Contrast SS is a new idea. Each estimate has an associated SS, and these can be added up to get SS for comparisons that include more than one estimate.

Looking ahead:

When equal numbers per treatment, as here:
All 3 methods give exactly same answers
When unequal numbers per treatment:
(1) works, (2) works when done right, (3) fails miserably
When missing cells (e.g. old product not given to men)
(1) works if done carefully, (2) and (3) fail miserably.

2 way factorial by Contrasts:

1 way ANOVA for these data:

Source	d.f.	SS	MS	F	p-value
Treatments	5	27.58	5.52	4.08	0.0017
Error	144	194.56	1.35		
c.total	149	222.14			

But, the 5 d.f. F test for treatments does not answer any of specific questions!
Answer the specific questions using contrasts.

Review of linear contrasts = linear combinations of means
$\gamma = \sum_{ij} c_{ij}\,\mu_{ij}, \qquad g = \sum_{ij} c_{ij}\,\overline{Y}_{ij.}, \qquad$ s.e. $g = s_p\sqrt{\sum_{ij}(c_{ij})^2/n_{ij}}$

	μ_{FC}	μ_{FL}	μ_{FS}	μ_{MC}	μ_{ML}	μ_{MS}	g	s.e.
M - F, av. over type	-1/3	-1/3	-1/3	1/3	1/3	1/3	0.04	0.19
L - S, av over sex	0	1/2	-1/2	0	1/2	-1/2	0.12	0.23
C - (L+S)/2, av over sex	1/2	-1/4	-1/4	1/2	-1/4	-1/4	-0.90	0.20
L-S, same in M and F	0	1	-1	0	-1	1	-0.08	0.46
C-(L,S), same in M and F	1	-1/2	-1/2	-1	1/2	1/2	0.12	0.40

Each linear contrast gives an estimate and se, which can be used in a t-test or a confidence interval.

When two contrasts are orthogonal, they represent statistically unrelated pieces of information

$\sum_i l_i\mu_i$ and $\sum_i k_i\mu_i$ are orthogonal when $\sum_i \frac{l_i\, k_i}{n_i} = 0$

Are L-S and C-(L+S)/2 orthogonal? (remember all cells have 25 people, $n_i = 25$)

L - S, av over sex (l_i)	0	1/2	-1/2	0	1/2	-1/2
C - (L+S)/2, av over sex (k_i)	1/2	-1/4	-1/4	1/2	-1/4	-1/4
product	0	-1/8	1/8	0	1/8	-1/8

$\sum_{ij} l_{ij}\, k_{ij}/n_{ij} = (0 + -1/8 + 1/8 + 0 + -1/8 + 1/8)/25 = 0/25 = 0$, so yes.

Each contrast has an associated SS, $\text{SS}_{\text{contrast}} = \frac{g^2}{\sum_{ij}(c_{ij})^2/n_{ij}} = \frac{g^2\, s_p^2}{s.e.^2}$

Contrast	g	$\sum_{ij}(c_{ij})^2/n_{ij}$	SS
M - F, av. over type	0.04	2/75	0.06
L - S, av over sex	0.12	1/25	0.36
C - (L+S)/2, av over sex	-0.90	3/100	27.00
L-S, same M and F	-0.08	4/25	0.04
C-(L,S), same M and F	0.12	3/25	0.12

Contrast SS give you tests of multiple questions simultaneously. For example:
"Averaged over sexes, is there a difference between types?" compares 3 types
No differences among the three types, implies
that three marginal means are equal $\mu_C = \mu_L = \mu_S$
this implies that L-S = 0 and C-(L+S)/2 = 0

When contrasts are orthogonal, can add SS to simultaneously test both components
SS for L-S = 0.36, SS for C-(L+S)/2 = 27, sum = 27.36

There are multiple sets of two orthogonal contrasts for 3 groups.
The SS depend on which contrasts are computed:
L-C: SS = 23.04, S - (L+C)/2: SS = 4.32.
But no matter what set you use, when they are orthogonal, the sum is the same:
sum=27.36

The sum of non-orthogonal contrasts is wrong:
L-C: SS = 23.04, S-C: SS = 17.64, sum=40.68
or, L-C: SS = 23.04, L-S: SS = 0.36, sum = 23.40

What about more than 2 contrasts?
A set of 3, 4, 5, $\cdots$ contrasts are orthogonal if all pairs are orthogonal
All pairs above are orthogonal. What do the SS add up to?
0.06 + 0.36 + 27.0 + 0.04 + 0.12 = 27.58 = trt SS from 1 way ANOVA

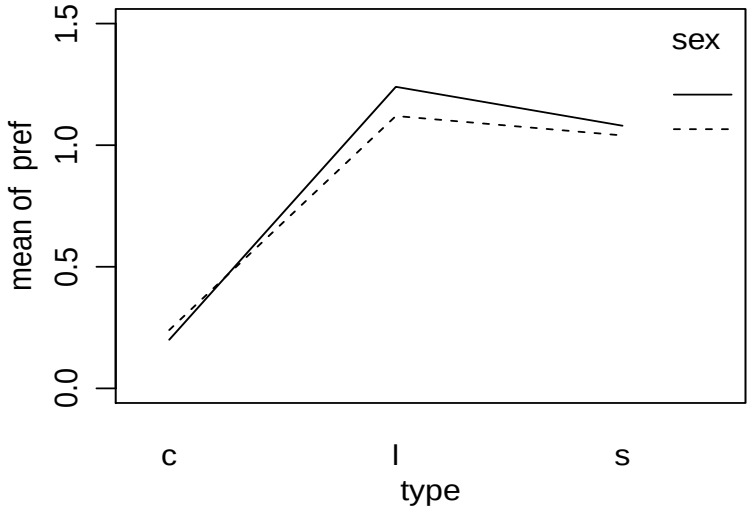
The traditional lines in the ANOVA table for a 2 way ANOVA represent one specific subdivision of the SS between all $I \times J$ groups.

2 way ANOVA for these data:

Source	d.f.	SS	MS	F	p-value
Sex	1	0.06	0.06	0.04	0.83
Type	2	27.36	13.68	10.12	< 0.0001
Sex*type	2	0.16	0.08	0.06	0.94
Error	144	194.56	1.35		
c.Total	149				

Conclusions from the 2 way ANOVA:
No evidence of a difference between men and women, averaged over types
Very strong evidence of at least one difference between types, averaged over sexes
individual contrasts tell you:
difference between liquid and solid is small, estimate = -0.12 with s.e. = 0.23
big difference between old and new types, averaged over sexes.
The estimate is that the new types are preferred by 0.90 units with s.e. = 0.20.
No evidence of interaction between sex and type.

These match the visual impression from plotting cell means:



Vocabulary:
Cell mean: average Y for one combination of all factors (e.g. Men, Liquid)
Marginal mean: ave. of cell means, averaged over all other factors, e.g.:
marginal mean for Liquid = average of Men/Liquid and Women/Liquid
marginal mean for Men = average of Men/Control, Men/Liquid and Men/Solid
LSmean: "Least squares" mean = either cell mean or marginal
Depending on what you're talking about, a cell or an average of cells

Simple effect: difference or linear contrast between two cell means,
e.g. Men, Liquid - Women, Liquid
Main effect: difference or linear contrast between two marginal means, e.g. Men - Women

Interactions:

Interaction exists when simple effects are not the same. Equivalent to non-parallel lines in a plot of means. Interactions answer the questions:

Do men and women react similarly to the types? i.e.
Is the difference between liquid and solid the same in men and women?
Is the difference between control and new the same in men and women?

Interpretation of interactions:
Sometimes (GxE study in plant/animal breeding): interactions are the goal of the study. Test of interactions is a key result.

Usually, focus on main effects or simple effects. Interaction test is produced automatically. How to interpret main effects depends size of interaction.

When interaction not significant:
Nice, easy interpretation of main effects. Interpret main effects as estimates of each simple effect.
Here, report that palatability of the liquid product is 0.96 units larger than the control.
This estimates simple effect in men and simple effect in women. Customary practice is to use the main effect here because

- main effects are more precise than simple effects (see below)
- There is no evidence of a difference in simple effects, so we will assume they are equal.

When interaction is significant:

1. Dogma (common in texts, journal reviewers): split data e.g. separately consider men (N=75) and women (N=75). analyze each group separately, report simple effects and tests within each group.

   If more interested in men-women, would split into three groups (control, liquid, solid)

2. My approach (1): Remember that the marginal mean is an average and the difference in marginal means is an average of simple effects.

   Do these averages "make sense"? If so, then report the marginal means and their differences.

   Don't forget that the simple effects could be quite different

   Examples (assuming a significant interaction):
   A situation where the main effect makes sense:
   Will sell product to a population of 50% men, 50% women.

Marginal mean is the average palatability in this population.
Note: can change the 50/50 in SAS and R/emmeans. Default is 50/50

Makes sense with some tweaking:
Will sell product to a population of 80% men, 20% women.
Concept of a marginal mean makes sense, but don't want 50/50.
Want 80/20 instead.
SAS and R/emmeans allow you to specify proportions other than 50/50

Doesn't make sense:
Fertilizer x Corn variety. Can only plant a field with one corn variety.
Want to know which fertilizer is best for that variety.
Average doesn't make sense, so need simple effects.

3. My approach (2): Look at how large the interaction effects are, not just whether they are significant. Esp. useful if the error is small or n is large. Sometimes will decide to report main effects and ignore the interaction, even if statistically significant

   Example: fertilizer x corn variety.
   diff between fertilizers for variety A: 5.2 bu/ac (se = 0.01),
   diff between fertilizers for variety B: 5.3 bu/ac (se = 0.01).
   Test is significant, but may decide to ignore.
   Don't ignore qualitative interactions!

Precision of marginal and cell means and their contrasts or differences:

Need estimate of within-group variability = $s = \sqrt{\text{MSE}}$
Food palatability study, $s = \sqrt{1.35} = 1.16$

s.e. of a row marginal mean = $s\sqrt{1/nJ} = s\sqrt{1/\#\text{ men in study}} = 1.16/\sqrt{75} = 0.13$.
s.e. of a col marginal mean = $s\sqrt{1/nI} = s\sqrt{1/\#\text{ liquid values in study}} = 1.16/\sqrt{50} = 0.16$.
s.e. of a cell mean = $s\sqrt{1/n} = s\sqrt{1/\#\text{ men,liquid in study}} = 1.16/\sqrt{25} = 0.23$.

s.e. of a diff. between XXX = $\sqrt{2}$ s.e. of XXX
substitute row mean, col mean or cell mean for XXX, as appropriate.
s.e. diff. in row means = 0.19, s.e. diff in col means = 0.23, s.e. diff in cell means = 0.33

s.e. of an interaction, e.g. $(\mu_{11} - \mu_{12}) - (\mu_{21} - \mu_{22}) = s\sqrt{4/n} = 0.46$

Some marginal means estimated more precisely,
all marginal means estimated more precisely than any cell mean.

Interaction effects are the least precisely estimated!

Difference between Liquid and Control in men:
two possible estimates: marginal means: 1.18 -0.22 = 0.96, s.e. = 0.23

simple effect: 1.24 - 0.20 = 1.04, s.e. = 0.33
which to use?

My suggested approach:
plot the cell means to show the pattern,
use interaction to decide whether to report marginal or simple effects

lack of fit tests using contrasts

Sometimes useful to think backwards with contrasts ⇒ answers to new questions.

Food palatability study: there are 6 trts, so 5 orthogonal contrasts.
You might have 2 important questions, e.g. diff between types: L-S, C-(L+S/2).
Can compute SS for those contrasts.
If other questions are very much less important,
may only want to ask 'is anything else different?'. I.e.,
(5 d.f. for diff btwn trt) = (2 df for btwn types) + (3 df for everything else).

Could figure out 3 orthog. contrasts for (everything else), but don't need to.
Logic: There are 5 orthog. contrasts. L-S and C-(L+S)/2 are two of them.
Get the SS for the other 3 by subtraction.
SS for btwn trts (5 d.f.) = SS for types (2 d.f.) + SS for everything else (3 d.f.), so
SS for everything else (3 d.f.) = SS for btwn trts (5 d.f.) - SS for types (2 d.f.)

Source	d.f.	SS	MS	F	p
Treatments	5	27.58			
L-S	1	0.36			
C-(L+S)/2	1	27.00			
rest	3	0.22	0.07	0.054	>0.5
		= 27.58-27.36			
Error	144	194.56	1.35		

Interpretation:
No evidence of any difference in palatability, other than the differences between types.

Sometimes called a lack of fit test or 'leftover SS' test.

You used this in 301/587 to test for lack of fit to a linear regression when you had multiple observations per group.

E.g. insulating fluid example in the Statistical Sleuth.
X = voltage, Y = log breakdown time.
7 voltage groups, different sample sizes for each voltage. 76 total observations

Source	d.f.	SS	MS	F	p-value
Voltage	6	196.48	32.74	13.00	< 0.0001
lin. reg.	1	190.15	190.15	75.51	< 0.0001
lack of fit	5	6.33	1.26	0.50	0.78
Error	69	173.75	2.51		

Model for 2 way anova (effects form):

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

Ignore variance, σ^2. Focus on the parameters defining the means for each group: μ, α_i, β_j, and γ_{ij}.

Same problem as with 1-way ANOVA: too many parameters.
In total, 12 parameters to be estimated from 6 cell means.
Way too many!

Solutions:
1) impose constraints on the parameters
e.g. force some parameters to be 0 or to sum to 0.
with 6 constraints, can estimate remaining 6 parameters.
Choice of constraint is arbitrary.
2) Use a generalized inverse
provides a solution for 12 parameters from 6 cell means
SAS and JMP do this; R does not
Many generalized inverses, choice made for you
SAS's choice equivalent to one specific choice of constraints

Illustration with made up data with three sets of constraints:

μ	α_M	α_W	β_C	β_S	β_L	γ_{MC}	γ_{ML}	γ_{MS}	γ_{WC}	γ_{WS}	γ_{WL}
1	2	**0**	-1	2	**0**	1	2	**0**	**0**	**0**	**0**
2	**0**	-2	**0**	3	1	1	2	**0**	**0**	**0**	**0**
2.33	**1**	**-1**	**-1.33**	**1.66**	**-0.33**	1	2	**0**	**0**	**0**	**0**

All three sets of parameters fit the data equally well!

My reaction when I first saw this:
What! One data set, many different answers.
That's too confusing! Which one is right?

They're all right, but you don't care
you're not interested in specific values for these parameters.
Interested in things like:
treatment means, marginal means, simple effects, and main effects.

Statistical theory:
the things you're really interested in have the same values
no matter which set of parameters you use.
because you're interested in estimable functions of the parameters

Estimable function:
a quantity that does not depend on the arbitrary choice of constraint.
Examples:

$\mu_{ML} = \mu + \alpha_M + \beta_L + \gamma_{ML}$
$\mu_{ML} - \mu_{WL} = (\mu + \alpha_M + \beta_L + \gamma_{ML}) - (\mu + \alpha_W + \beta_L + \gamma_{WL}) = (\alpha_M - \alpha_W) + (\gamma_{ML} - \gamma_{WL})$
$\mu_{M.} - \mu_{W.}$
$= \frac{1}{3}\Sigma_{C,L,S}(\mu + \alpha_M + \beta_i + \gamma_{Mi}) - \frac{1}{3}\Sigma_{C,L,S}(\mu + \alpha_W + \beta_i + \gamma_{Wi})$
$= (\alpha_M - \alpha_W) + \frac{1}{3}\Sigma_{C,L,S}(\gamma_{Mi} - \gamma_{Wi})$

Illustration with made up data with three sets of constraints:

μ	α_M	α_W	β_C	β_S	β_L	γ_{MC}	γ_{ML}	γ_{MS}	γ_{WC}	γ_{WS}	γ_{WL}	μ_{ML}	$\mu_{ML} - \mu_{WL}$	$\mu_{M.} - \mu_{W.}$
1	2	0	-1	2	0	1	2	0	0	0	0	5	4	3
2	0	-2	0	3	1	1	2	0	0	0	0	5	4	3
2.33	1	-1	-1.33	1.66	-0.33	1	2	0	0	0	0	5	4	3

If your software complains about "non-est" or "not estimable"
you have asked for a quantity that depends on the choice of constraints.
Your software is telling you that what you asked for multiple possible answers.
The most likely reason is that you asked for the wrong thing.

SS by Model comparison:

Reminder: 1 way ANOVA, SS can be computed by comparing two models:
Full: $Y_{ij} = \mu_i + e_{ij}$
Reduced: $Y_{ij} = \mu + e_{ij}$
The difference in error SS = the SS for "groups"

Model comparison the hard way:
Fit the reduced model:

```
 lm(y ~ 1, data=food)
proc glm data=food; model Y = ; run;
```

SS_{error} = 222.14, with 149 d.f. (= 150 - 1)

Fit the full model:

```
 lm(y ~ group, data=food)
proc glm data=food; class group; model Y = group; run;
```

SS_{error} = 194.56, with 144 d.f. (= 150 - 6)

SS for groups by subtraction: SS_{groups} = 222.14 - 194.56 = 27.58 with 149 - 144 = 5 d.f.
Exactly the same SS_{groups} as by contrasts or formulae!

Model comparison for a 2 way factorial:

Idea: Compute SS for Sex by comparing model with sex effect ($\alpha's$) to one without
Fit the reduced model:

```
 lm(y ~ 1, data=food)
proc glm data=food; class sex; model Y = ; run;
```

SS_{error} = 222.14, with 149 d.f. (= 150 - 1)

Fit the full model:

```
lm(y ~ sex, data=food)
proc glm data=food; class sex; model Y = sex; run;
```

SS_{error} = 222.08, with 148 d.f. (= 150 - 2)

SS for sex by subtraction: SS_{sex} = 222.14 - 222.08 = 0.06 with 149 - 148 = 1 d.f.
Exactly the same SS_{sex} as by contrasts or formulae!

But, which pair of models?

Effect		model	error SS	SS for effect
Sex	Full	$\mu + \alpha_i$	222.08	
	Red.	μ	222.14	0.06
Sex	Full	$\mu + \alpha_i + \beta_j$	194.72	
	Red.	$\mu + \beta_j$	194.78	0.06
Sex	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	194.56	
	Red.	$\mu + \beta_j + \gamma_{ij}$	194.62	0.06
Type	Full	$\mu + \beta_j$	194.78	
	Red.	μ	222.14	27.36
Type	Full	$\mu + \alpha_i + \beta_j$	194.72	
	Red.	$\mu + \alpha_i$	222.08	27.36
Type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	194.56	
	Red.	$\mu + \alpha_i + \gamma_{ij}$	221.92	27.36
Sex*type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	194.56	
	Red.	$\mu + \alpha_i + \beta_j$	194.72	0.16

Very nice consequence of equal sample sizes (also called balanced data): When sample sizes are equal (balanced), choice of model pair doesn't matter. consequence of orthogonality

When sample sizes are not equal, choice does matter.

Types of SS, illustrated using unbalanced data.
Food palatability study with sample sizes ranging from 21 to 25 people per group.

Type I SS
also called sequential SS
each term compared to model **with only 'earlier' terms**

Model	Effect		model	error SS	SS for effect
Sex Type Sex*type	Sex	Full	$\mu + \alpha_i$	209.87	
		Red.	μ	210.00	0.13
	Type	Full	$\mu + \alpha_i + \beta_j$	186.46	
		Red.	$\mu + \alpha_i$	209.87	23.41
	Sex*type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \beta_j$	186.46	0.34
Type Sex Sex*Type	Type	Full	$\mu + \beta_j$	186.60	
		Red.	μ	210.00	23.40
	Sex	Full	$\mu + \alpha_i + \beta_j$	186.46	
		Red.	$\mu + \beta_j$	186.60	0.14
	Sex*type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \beta_j$	186.46	0.34

Type III SS: also called partial SS: Each term compared to model with **all other terms except term of interest**

SAS model	Effect		model	error SS	SS for effect
sex type sex*type	Sex	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \beta_j + \gamma_{ij}$	186.26	0.14
	Type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \gamma_{ij}$	209.44	23.32
	Sex*type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \beta_j$	186.46	0.34
type sex sex*type	Sex	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \beta_j + \gamma_{ij}$	186.26	0.14
	Type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \gamma_{ij}$	209.44	23.32
	Sex*type	Full	$\mu + \alpha_i + \beta_j + \gamma_{ij}$	186.12	
		Red.	$\mu + \alpha_i + \beta_j$	186.46	0.34

Type I SS depend on order of terms in the model.
- when data are unbalanced (most unequal sample sizes)
  - different orders give different type I SS
  - Here, differences are small; can be huge.
- when data are balanced (equal sample sizes)
  - different orders give the same type I SS
- Why the difference?
  - Contrasts that are orthogonal for equal sample sizes
    - are not orthogonal for unequal sample sizes

Type III SS (and F tests) are the same for any orders.

Big advantage to type III tests.
Same as SS derived using contrasts among cell means.

Type II SS (and F tests) similar to type III, but assume no interaction.
Type IV SS: proposed by SAS in the 1970's, no longer used

Missing cells

No observations for one or more combinations of levels. E.g. Men only tasted old and solid products, no data for men/liquid ⇒ missing information on that cell

		Type		
Sex	old	liquid	solid	average
F	0.24	1.12	1.04	0.80
M	0.20		1.08	??
average	0.22	??	1.06	??

1 way ANOVA has 4 d.f. (only 5 groups). Divide into usual 2 way ANOVA quantities:

Source	d.f.	
Sex	1	
Type	2	
Sex*Type	**1**	should be 2. lose the d.f. here, in the interaction!

If you have missing cells, and the model includes an interaction
SAS (and some other programs) will report type I and type III SS.
Main effect tests (sex, type) are **meaningless!**,
because those tests correspond to meaningless comparisons among models.
Can interpret interaction (but limited to a subset of the data).
Often the first clue: LSMEANS are non-est.
Remember, row average is average of three cell means.
Can't estimate Men, liquid,
so can't calculate men marginal mean
or liquid marginal means.

Solutions:
Best: Write your own contrasts.
Choose them to answer the questions that can be answered from the data
OK: Test interaction, if very ns. e.g. $p > 0.20$, drop interaction from model
Main effect tests do make sense if no interaction in model.
Make sure you know what you're doing!

Estimates of marginal means

Balanced data easy, unbalanced data requires some care
Type III approach with all interactions in the model usually makes most sense

Estimates of sex effect (difference between men and women)

Same ideas for type, but have to deal with 3 levels

Balanced data: (25 per cell)

Model	M	W	diff.	s.e.
sex only, ignore type	0.840	0.800	-0.40	0.20
"type II", ignore interaction	0.840	0.800	-0.40	0.120
type III	0.840	0.800	-0.40	0.120
type III without interaction in model	0.840	0.800	-0.40	0.19

Unbalanced data: (21-25 obs per cell)

Model	M	W	diff.	s.e.
sex only, ignore type	0.863	0.803	-0.060	0.203
"type II", ignore interaction:	0.859	0.795	-0.064	0.194
type III:	0.858	0.795	-0.062	0.194
type III without interaction in model:	0.859	0.795	-0.064	0.192

Differences are small here. Can be large.

My view:
- Experimental studies:
  - Type III with interaction makes most "sense":
  - Cell (combination of factors) is what you manipulate
    - Factors are made-up constructs
  - Tests compare averages of cell means
    - equivalent to averages of simple effects.
- Observational studies:
  - Analyses often don't include interactions
    - often more interaction terms than there are observations
  - Type III without interaction makes most "sense"
  - Unless there is a clear ordering of factors

But, assumes population of interest has equal amounts of each group.
- Usually reasonable in a designed experiment.
- May be a problem in an observational study.

Sometimes this isn't (may not be) appropriate:
- see me for more details if interested

Beware:
- Dropping interaction from model gives you type II estimates

(even though labelled type III).
Moral: always include all interactions in your model
(unless you have good reasons to do otherwise, and you know what you're doing!

Putting the pieces together; Doing the analysis:

Start with the ANOVA table (1 way or 2 way, depending on question(s).

Use F tests based on type III SS to answer "standard" factorial questions.
F tests are the start, not the end of the analysis
What are the means, differences/contrasts that answer important questions?
How precise are means, differences, contrasts?
Most useful number in the ANOVA table: often the MSE
Plot the means in a way that communicates the key results

Some useful things to check:
Check residuals to make sure model reasonable
Especially if important effects have large s.e.'s
Look for equal variances, additive effects
If not reasonable, correcting often increases power
If you believe you have balanced data:
Check whether Type I SS = Type III SS. Should be same when balanced.
Often find they're not! especially useful when many factors or levels
Student forgot about the missing observation(s).
One or more lines of data accidently left out.
R/JMP/SAS didn't read data correctly.
Check d.f. for highest order interaction.
Should be product of main effect d.f.
If not, you're missing one or more cells. Stop and think hard!

Study design: Choosing a sample size:

Easy using t-statistics. Generalize the approach from 1way ANOVA.
Choose the quantity of greatest interest or least precisely estimated

Specify the difference of interest and error s.d.
Calculate the s.e. of the quantity of interest.
Plug into power calculation.
Can often use software by using contrasts and 1-way ANOVA approach.
If not, will probably have to do by hand.

Example: want 80% power to detect a difference of 0.5 in palatibility, s.d. =1.16.
Evaluate main effect of sex, main effect of food (liquid-solid), one simple effect, and interaction (M l-s - W l-s)

Effect	s.e.	n per group
Sex	$\sqrt{2/3n}$	29
Type: L - S	$\sqrt{2/2n}$	43
Simple effect: L-S in M	$\sqrt{2/n}$	85
Interaction	$\sqrt{4/n}$	170

SS in ANOVA table by averaging observations and using formulae:

Notation: I rows (here I=2, sex), J columns (here J=3, type), n obs per cell

Y_{ijk}: observation k for sex i, type j.
$Y_{ij.}$: average of observations from sex i, type j. (n=25)
$Y_{i..}$: average of observations from sex i (nJ=75)
$Y_{.j.}$: average of observations from type j (nI=50)
$Y_{...}$: average of all observations (nIJ=150)

SS as variability between averages (works when data are balanced)

Source	d.f.	here	Sum of Squares	here
Treatment	$IJ-1$	5	$n\sum_{ij}\left(\overline{Y}_{ij.}-\overline{Y}_{...}\right)^2$	27.58
Error	$IJ(n-1)$	144	$\sum_{ijk}\left(Y_{ijk}-\overline{Y}_{ij.}\right)^2$	194.56
c.total	$IJn-1$	149	$\sum_{ijk}\left(Y_{ijk}-\overline{Y}_{...}\right)^2$	222.14

Source	d.f.	here	Sum of Squares	here
Sex	$I-1$	1	$nJ\sum_{i}\left(\overline{Y}_{i..}-\overline{Y}_{...}\right)^2$	0.06
Type	$J-1$	2	$nI\sum_{j}\left(\overline{Y}_{.j.}-\overline{Y}_{...}\right)^2$	27.36
Sex*Type	$(I-1)(J-1)$	2	$n\sum_{i}\left(\overline{Y}_{ij.}-\overline{Y}_{i..}-\overline{Y}_{.j.}+\overline{Y}_{...}\right)^2$	0.16
Error	$IJ(n-1)$	144	$\sum_{ijk}\left(Y_{ijk}-\overline{Y}_{ij.}\right)^2$	194.56
c.total	$IJn-1$	149	$\sum_{ijk}\left(Y_{ijk}-\overline{Y}_{...}\right)^2$	222.14

Notice:

Sex SS is variability between averages for each sex, 0 when all sexes have same average
Type SS is variability between averages for each type, 0 when all types have same average
Will come back to Sex*Type
Error same in both ANOVA's: pooled variability between obs in the same treatment.
df for Sex, Type and Sex*type add up to df for trt in 1 way ANOVA
  quick algebra, always so
SS for Sex, Type and Sex*type add up to SS for trt in 1 way ANOVA
  tedious algebra, always so when balanced

Approach completely falls apart if sample sizes are unequal.