

**Keppel, G. & Wickens, T. D. *Design and Analysis***  
**Chapter 10: Introduction to Factorial Designs**

**10.1 Basic Information from a Factorial Design**

- “...a factorial design contains within it a set of separate *single*-factor experiments.” That’s why it’s so important to really understand the single-factor design!
- K&W illustrate the component nature of the factorial design in Table 10.1. Thus, in the example they’re discussing, you would be designing the format for a series of reading books for elementary schools. You could examine the impact of the length of the printed lines and the contrast between the print and the paper. The length of the printed lines of text could be 3, 5, or 7 inches. The difference in lightness between the print and the paper could also be manipulated (low, medium, or high). The dependent variable is reading speed.

	3 in.	5 in.	7 in.
Low Contrast			
Medium Contrast			
High Contrast			

- You could think of this experiment as a combination of two separate single-factor independent group designs. Thus, one factor (independent variable) would be line length. The other factor would be contrast. Let’s first consider the experiment on line length. Assume that we’re dealing only with the three levels of the factor and with  $n = 30$ . Thus, we would be looking at an ANOVA with a source table that might look like this:

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Line Length	62.18	2	31.09	7.08
Error	381.82	87	4.39	
Total	444	89		

Alternatively, we could think of the contrast single-factor experiment, which should again make sense to you in the context of the prior chapters. That is, we would analyze the data from this single-factor independent groups design with an ANOVA that might look like this:

Source	<i>SS</i>	<i>df</i>	<i>MS</i>	<i>F</i>
Contrast	201.98	2	100.99	36.33
Error	242.02	87	2.78	
Total	444	89		

When we now consider this study as a two-factor design, you should be able to predict the *df* that would be involved in the experiment. With the same data (i.e., 90 total scores) presumably distributed equally across the 9 unique conditions,  $n = 10$ . You should also be

able to enter the  $SS_{\text{Total}}$ . In fact, if you think about it a bit, you should also be able to predict  $SS_{\text{Line Length}}$  and  $SS_{\text{Contrast}}$ .

Source	SS	df	MS	F
Line Length				
Contrast				
Interaction				
Error				
Total				

- In considering the two-factor design, we would describe the independent impact of line length as a **main effect**. That is, even though contrast is a factor in the study, we could ignore contrast by averaging over the effects of the three levels of contrast to arrive at the main effect of line length. Similarly, the independent effect of contrast would also be a main effect. In essence, analysis of these main effects is what you've been doing with single-factor designs.
- Another way to think of these data as a single-factor design is to look at the effects of a factor (e.g. line length) at only one level of the other factor (e.g., low contrast). In so doing, we wouldn't be looking at a main effect of line length, but instead the **simple effect** of line length under Low Contrast. We could also examine the simple effect of line length under Medium Contrast. Finally, we could examine the simple effect of line length under High Contrast.
- Note that we could turn the experiment on end and think of the 3 simple effects of contrast at varying levels of line length. We could also consider the main effect of contrast (averaging over all levels of line length).
- Aside from economy, the real advantage of factorial designs is the ability to examine interactions. In essence, an **interaction** is present when the simple effects differ. Thus, if the effects of line length were equivalent regardless of the level of contrast, we would have no evidence of an interaction. When an interaction is present, interpretation of a main effect becomes complicated.

## 10.2 The Concept of Interaction

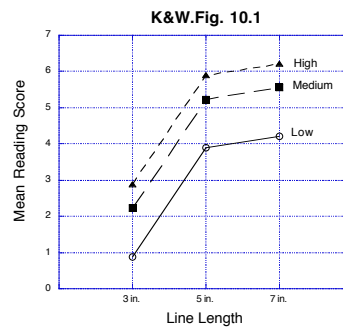
- The main effects of a factorial design are computed in a fashion identical to the computation of treatment effects in a single factor design. (See, I told you it was really crucial to understand the single-factor ANOVA!) Thus, the only new computation is the interaction term.

- K&W first present an example of an outcome that illustrates no interaction:

	3 in.	5 in.	7 in.	Marginal Mean
Low Contrast	.89	2.22	2.89	2.00
Medium Contrast	3.89	5.22	5.89	5.00
High Contrast	4.22	5.55	6.22	5.33
Marginal Mean	3.00	4.33	5.00	4.11

Why is this outcome consistent with no interaction? Let's focus on the simple effects of line length. (We could just as easily have chosen to focus on the simple effects of contrast.) For the Low Contrast condition, note that the increase in reading speed from the 3-inch condition to the 5-inch condition is 1.33, and the increase from the 5-inch condition to the 7-inch condition is .67. Although overall reading speed increases when the contrast is medium, the increase in reading speed from the 3-inch condition to the 5-inch condition remains 1.33, and the increase from the 5-inch condition to the 7-inch condition remains .67. Finally, when the contrast is high, reading speeds show further improvement, but the increase in reading speed from the 3-inch condition to the 5-inch condition remains 1.33, and the increase in reading speed from the 5-inch condition to the 7-inch condition remains .67. The fact that the changes in reading speed remain consistent for all the simple effects is indicative of a lack of interaction.

- You could also take a graphical approach to interaction, which is to look for a lack of parallel lines in a graph of the condition means. As you can see in the graph below, the lines connecting the means are parallel.

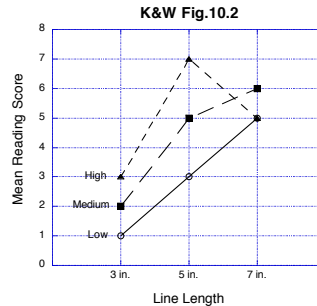


- In the absence of a significant interaction, you would focus your attention on the main effects. That is, because the simple effects are all telling you basically the same story, there's little reason to examine them separately.
- K&W next present an example of an outcome that illustrates an interaction:

	3 in.	5 in.	7 in.	Marginal Mean
Low Contrast	1.00	2.00	3.00	2.00
Medium Contrast	3.00	5.00	7.00	5.00
High Contrast	5.00	6.00	5.00	5.33
Marginal Mean	3.00	4.33	5.00	4.11

(Note that the marginal means remain the same as those in the example showing no interaction. Thus, the main effects are identical.)

- In this example, you can quickly see that the simple effects are not the same. This outcome is particularly evident when looking at a figure of the data. (Note the non-parallel lines.)



- When an interaction is present, we are most often less interested in the main effects. Instead, we typically focus on interpreting/understanding the interaction.

### 10.3 The Definition of an Interaction

- K&W provide a number of definitions of an interaction. In case any one of them may help you to understand the concept, I'll simply copy them for you here:

An interaction is present when the effects of one independent variable on behavior change at the different levels of the second independent variable. (N.B. Keppel's admonition to keep in mind that the two independent variables do not influence one another!)

An interaction is present when the values of one or more contrasts in one independent variable change at the different levels of the other independent variable.

An interaction is present when the simple effects of one independent variable are not the same at all levels of the second independent variable.

An interaction is present when the main effect of an independent variable is not representative of the simple effects of that variable.

An interaction is present when the differences among the cell means representing the effect of Factor A at one level of Factor B do not equal the corresponding differences at another level of Factor B.

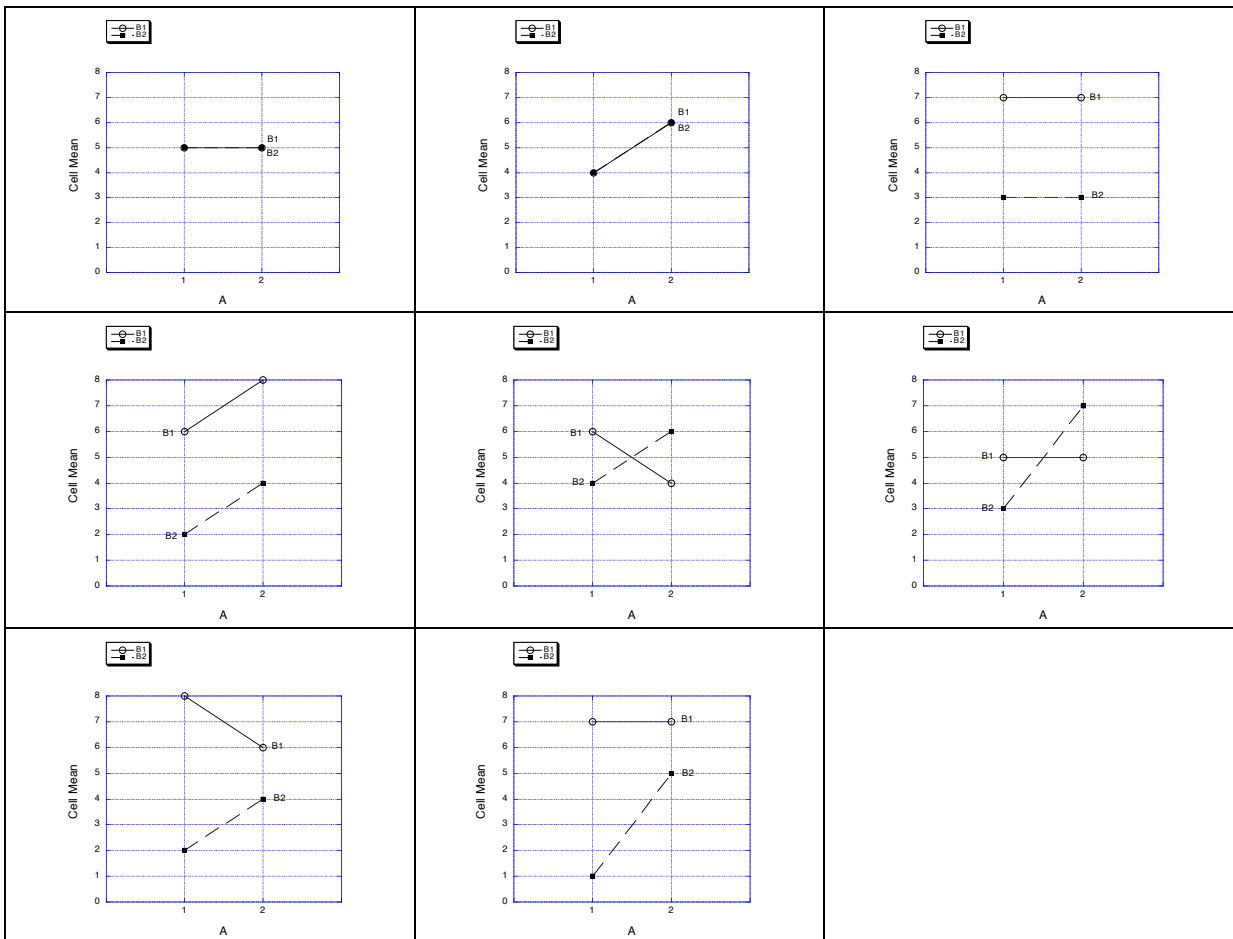
An interaction is present when the effects of one of the independent variables are conditionally related to the levels of the other independent variable.

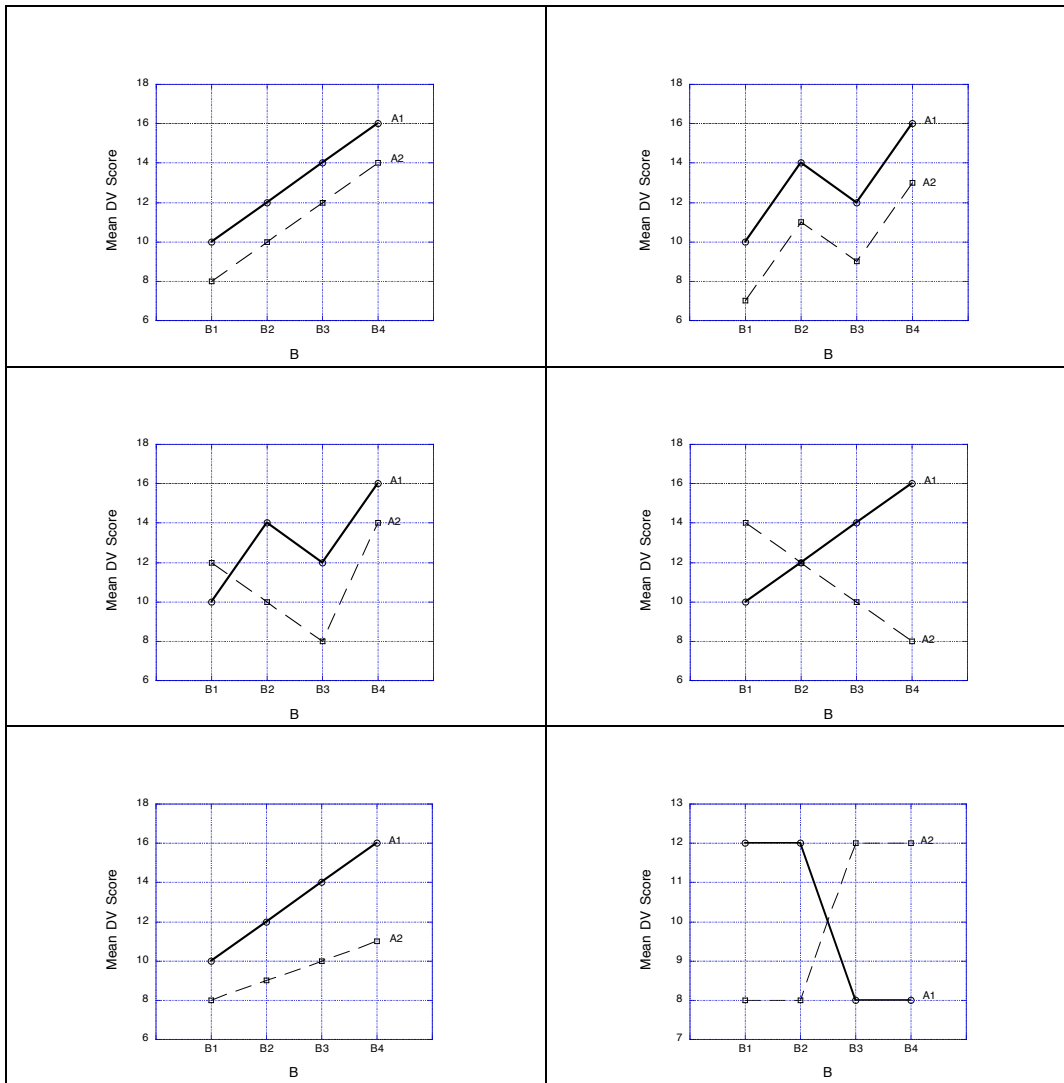
- Although the presence of an interaction may complicate the explication of the results of an experiment, you should realize that many experiments actually predict an interaction. Thus, you may well be looking to find an interaction in an experiment. K&W use Garcia's research to illustrate such an instance. You should be aware of Garcia's seminal research and the

example that K&W provide is a nice illustration of a simple experiment that predicts an interaction.

### 10.4 Further Examples of an Interaction

• K&W provide a number of 2x2 outcomes to illustrate the presence or absence of an interaction. You should be sure that you can easily determine that an interaction appears to be present or absent in an experiment. Of course, the final arbiter of the presence of an interaction is not a graph, but the statistical test that you will find in the source table. Even though lines may appear to be non-parallel, if the  $F$  for the interaction is small (i.e., the  $p$  value is above .05), then the lines are statistically parallel and there is no significant interaction.





### 10.5 Measurement of the Dependent Variable

• K&W discuss the notion that as researchers, we have a good deal of control over the types of measures that we choose to use in our research. (Just as we control the IVs and their levels.) As they illustrate, if you were interested in measuring statistical knowledge (or any other knowledge domain), there are all sorts of tests you could implement to assess a person's level of knowledge. As a researcher, you would be able to envision the advantages and disadvantages of a range of dependent variables.

- K&W provide an example of a situation in which the way you measure your DV may cloud the extent to which an interaction is present. They talk about a difficult test for children (Younger and Older) some of whom work on the test for 10 min and others for 20 min. Their hypothetical results are:

	10 min	20 min
Younger	8	14
Older	30	36

As you look at the numbers, you should be struck by the fact that there is no interaction in the data. At the same time, as K&W note, the performance of the younger children probably increases appreciably more than that of the older children with more time. That is, from a base of 8 correct, an increase of 6 is substantial. At the same time, an increase from 30 to 36 is not equally impressive. But with this DV, you couldn't conclude that there is an interaction.

- "...for the size of an interaction to be precisely interpretable, we need to know that the simple effects—in this case the differences between means—mean the same thing regardless of where they lie in the range of the dependent variable. Strictly speaking, the following principle must hold:

*The meaning of the difference between two values of the dependent variable does not depend on the size of the scores."*

- K&W describe some interactions as **removable** and others as **non-removable**. You'll need to plot the interaction with each variable on the x-axis, and if the lines cross in either plot, then the interaction is non-removable (or a **crossover interaction**). If the interaction is removable, then a transformation on the DV may well remove the interaction. "...many researchers consider non-removable interactions more compelling because they are resistant to every possible change in measurement scale."