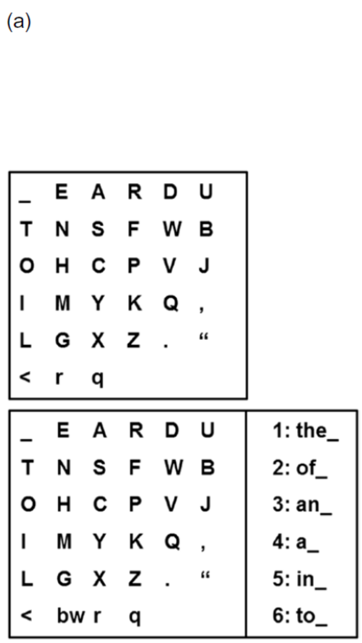
I. Scott MacKenzie: ***Human-Computer Interaction. An Empirical Research Perspective***, Elsevier, 2013.

Chapter 5: Designing HCI Experiments

5.12 Group effects and asymmetric skill transfer

If the learning effect is the same from condition to condition in a within-subjects design, then the group means on a dependent variable should be approximately equal.[1] This was demonstrated above (see Figure 5.9). In other words, the advantage due to practice for a condition tested later in the experiment is offset equally by the disadvantage when the same condition is tested earlier in the experiment. That's the point of counterbalancing. However, there are occasions where different effects appear for one order (e.g., A→B) compared to another (e.g., B→A). In such cases there may be a group effect—differences across groups in the mean scores on a dependent variable. When this occurs, it is a problem. In essence, counterbalancing did not work. A group effect is typically due to asymmetric skill transfer—differences in the amount of improvement, depending on the order of testing.



(b)

Testing Half		Group
First (Trials 1-10)	**Second** (Trials 11-20)	
20.42	27.12	1
22.68	28.39	
23.41	32.50	
25.22	32.12	
26.62	35.94	
28.82	37.66	
30.38	39.07	
31.66	35.64	
32.11	42.76	
34.31	41.06	
19.47	**24.97**	2
19.42	**27.27**	
22.05	**29.34**	
23.03	**31.45**	
24.82	**33.46**	
26.53	**33.08**	
28.59	**34.30**	
26.78	**35.82**	
31.09	**36.57**	
31.07	**37.43**	

FIGURE 5.13

Experiment comparing two scanning keyboards: (a) Letters-only keyboard (LO, *top*) and letters plus word prediction keyboard (L + WP, *bottom*). (b) Results for entry speed in characters per minute (cpm). Shaded cells are for the LO keyboard.

We could develop an example of asymmetric skill transfer with hypothetical data, as with the counterbalancing example above; however, there is an example data set in a research report where an asymmetric transfer effect is evident. The example provides a nice visualization of the effect, plus an opportunity to understand why asymmetric skill transfer occurs. So we'll

---

[1] There is likely some difference, but the difference should not be statistically significant.

use that data. The experiment compared two types of scanning keyboards for text entry (Koester and Levine, 1994a). Scanning keyboards use an on-screen virtual keyboard and a single key or switch for input. Rows of keys are highlighted one by one (scanned). When the row bearing the desired letter is highlighted, it is selected. Scanning enters the row and advances left to right. When the key bearing the desired letter is highlighted it is selected and the letter is added to the text message. Scanning keyboards provided a convenient text entry method for many users with a physical disability.

The experiment compared a letters-only (LO) scanning keyboard with a similar keyboard that added word prediction (L + WP). The keyboards are shown in Figure 5.13a. Six participants entered 20 phrases of text, 10 with one keyboard, followed by 10 with the other. To compensate for learning effects, counterbalancing was used. Participants were divided into two groups. Group 1 entering text with the LO keyboard first, then with the L + WP keyboard. Group 2 used the keyboards in the reverse order. Although not usually provided in a report, the results were given in a table showing the entry speed in characters per minute (cpm). The data are reproduced in Figure 5.13b as they appeared in the original report (Koester and Levine, 1994a, Table 2). The two columns show the sequence of testing: first half, then second half. The shaded and un-shaded cells show the results for the LO and L + WP keyboards respectively, thus revealing the counterbalanced order.

There are at least three ways to summarize the data in Figure 5.13b. The overall result showing the difference between the LO and L + WP keyboards is shown in the left-side chart in Figure 5.14. Clearly, there was very little difference between the two keyboards: 30.0 cpm for the LO keyboard versus 30.3 cpm for the L + WP keyboard. The L + WP keyboard was just 1 percent faster. The error bars are large, mostly due to the improvement from trial to trial, as seen in Figure 5.13b.

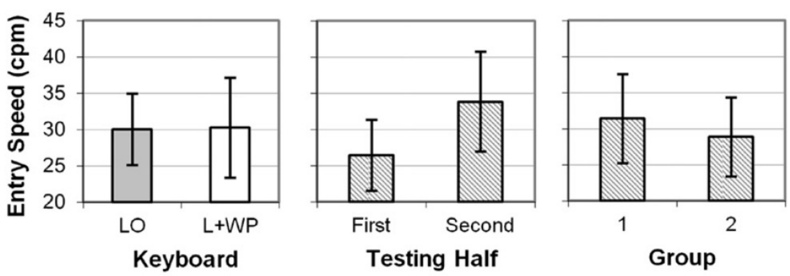


FIGURE 5.14

Three ways to summarize the results in Figure 5.13b, by keyboard (*left*), by testing half (*center*), and by group (*right*). Error bars show ±1 *SD*.

The center chart in Figure 5.14 shows another view of the results, comparing the first half and second half of testing. A learning effect is clearly seen. The overall entry speed was 26.4 cpm in the first half of testing (trials 1 to 10) and 33.8 cpm, or 28 percent higher, in the second half of testing (trials 11 to 20). Learning is fully expected, so this result is not surprising.

Now consider the right-side chart in Figure 5.14. Counterbalancing only works if the order effects are the same or similar. This implies that the performance benefit of an LO→L + WP order is the same as the performance benefit of an L + WP→LO order. If so, the group means will be approximately equal. (This was demonstrated earlier in the counterbalancing example; see Figure 5.9). The right-side chart in Figure 5.14 reveals a different story. The mean for Group 1 was 31.4 cpm. The mean for Group 2 was lower at 28.8 cpm. For some reason, there was an 8 percent performance disadvantage for Group 2. This is an example of asymmetric skill transfer. Figure 5.15 illustrates. The figure reduces the data in Figure 5.13b to four points, one for each quadrant of 10 trials. Asymmetry is clearly seen in the cross-over of the lines connecting the LO points and L + WP points between the first half and second half of testing.

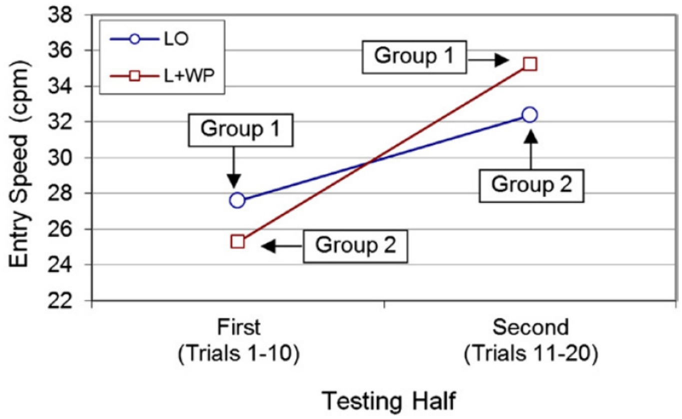


FIGURE 5.15

Demonstration of asymmetric skill transfer. The chart uses the data in Figure 5.13b.

If counterbalancing had worked, the lines in Figure 5.15 would be approximately parallel. They are not parallel because of the asymmetry in the LO→L + WP order versus the L + WP→LO order. Asymmetric skill transfer is usually explainable by considering the test conditions or the experimental procedure. For this experiment, the effect occurs because of the inherent differences in entering text with the letters-only (LO) keyboard versus entering text with the letters plus word prediction (L + WP) keyboard. In fact, this example provides an excellent opportunity to understand why asymmetric skill transfer sometimes occurs. Here is the explanation. The L + WP keyboard is an enhanced version of the LO keyboard. The basic method of entering letters is the same with both keyboards; however, the L + WP keyboard adds word-prediction, allowing words to be entered before all letters in the word are entered. It is very likely that entering text first with the LO keyboard served as excellent practice for the more difficult subsequent task of entering text with the L + WP keyboard. To appreciate this, examine the two points labeled Group 1 in Figure 5.15. Group 1 participants performed better overall because they were tested initially with the easier LO keyboard before moving on the enhanced L + WP keyboard. Group 2 participants fared less well because they were tested initially on the more difficult L + WP keyboard.

The simplest way to avoid asymmetric skill transfer is to use a between-subjects design. Clearly, if participants are exposed to only one test condition, they cannot experience skill transfer from another test condition. There are other possibilities, such as having participants practice on a condition prior to data collection. The practice trials seek to overcome the benefit of practice in the earlier condition, so that the measured performance accurately reflects the inherent properties of the test condition. It is not clear that this would work in the example. Participants cannot "unlearn." In the end, the performance difference between the LO and L + WP keyboards remains an outstanding research question. The practice effect (28%) was much greater than the group effect (8%), so it is difficult to say whether word prediction in the L + WP keyboard offers a performance advantage. Clearly, there is a benefit with the L + WP keyboard, because words can be entered before all the letters are entered. However, there is also a cost, since users must attend to the on-going prediction process, and this slows entry. To determine whether the costs outweigh the benefits in the long run, a longitudinal study is required. This is examined in the next section.