Chapter 10 Supplement

Normalizing Scalp Distributions

As discussed in chapter 10, a difference in the amplitude of a single ERP generator between two conditions will lead to a multiplicative change in scalp distribution, which will appear as a condition × electrode site interaction in an ANOVA. To determine whether condition × electrode site interaction is caused by this multiplicative effect or is different from what would be expected from a pure multiplicative effect, McCarthy and Wood (1985) proposed *normalizing* the data to remove any differences in the overall amplitudes of the conditions. Once the overall amplitude differences are eliminated, any differences in scalp distribution that are caused by the multiplicative effect should be eliminated.

McCarthy and Wood (1985) proposed two different ways of doing this, but later research showed that one of them was incorrect. The correct way to do this is to divide the voltage at each electrode site by the *vector length* for that condition. To compute the vector length, you square the voltage from each electrode site in that condition, add the squared values together, and take the square root of this sum. In condition A of the experiment shown in figure 10.2A, for example, the vector length is sqrt($1.0^2 + 1.5^2 + 2.0^2$) = sqrt(1 + 2.25 + 4) = sqrt(7.25) = 2.6926. You then divide each voltage in a condition by the vector length for that condition. For example, the normalized voltage at the Fz electrode site in condition A is computed by dividing the original voltage (1.0 µV) by the vector length for this condition (2.6926 µV), giving us a normalized value of 0.3719 µV.

Figure 10.2B shows the results of normalizing the data in figure 10.2A. Notice that the normalization completely eliminates the main effect of condition. You will therefore use an analysis of the non-normalized data to determine whether a condition main effect is present. The normalization is designed solely to assess the condition × electrode interaction. After eliminating the overall differences in amplitude among conditions, we can see that the scalp distribution is exactly the same for conditions A and B. These conditions truly have the same scalp distribution, as would be expected if a single generator simply varied in magnitude between them. Condition C now has a different scalp distribution. If we conducted an ANOVA with the normalized data from conditions A and C, we would see a condition × electrode interaction (assuming we had sufficient statistical power), and that would conventionally be interpreted as meaning that the generators were not exactly the same in these two conditions.

Problems with Normalization

It turns out that there is a problem with normalization. The McCarthy and Wood (1985) normalization procedure was used by many labs for many years, and then Urbach and Kutas (2002) convincingly demonstrated that it doesn't really work under most realistic conditions. The most common failure occurs when more than one component contributes to the voltage. For example, imagine that you have two conditions, A′ and B′, which are just like conditions A and B in figure 10.2A, except that a second component is also present. Assume that this second component is equally large in both conditions and that it has an amplitude of –1 µV at all three electrode sites. The amplitudes observed in conditions A′ and B′ would be just like those shown in figure 10.2A, except that every voltage would be decreased by 1 µV. The differences between conditions A′ and B′ are no longer perfectly multiplicative, because only one of the two components increases in amplitude between the two conditions. If we normalized the data, the scalp distributions would still differ between the two conditions. The same problem happens if only a single component is present, but the baseline is contaminated by overlap or preparatory activity (see, e.g., figure 11.4 in online chapter 11).

These problems can sometimes be eliminated by measuring well-controlled difference waves, in which any prestimulus activity is subtracted away and only a single component remains in the difference. However, it is difficult to be certain that only a single component remains in the difference wave. And even if only a single component is present, differences in noise among conditions can distort the vector length that is used for the normalization (see Urbach & Kutas, 2006). Thus, a significant condition × electrode site interaction may be obtained after normalization even if there is no difference in the relative distribution of internal brain activity, and a real difference in the relative distribution of internal brain activity may not yield a significant condition × electrode site interaction in the normalized data (even with infinite statistical power).

There is also another problem with normalization, but it's conceptual rather than technical. Many researchers have looked at condition × electrode site interactions to determine whether or not the same brain areas are active in different experimental conditions. Even if normalization worked correctly, it would be impossible to make claims of this sort on the basis of a condition × electrode site interaction. The problem is that a significant interaction would be observed if exactly the same generators are active in both conditions but they differ in relative magnitude. That is, if areas A and B have amplitudes of 6 and 12 units in one condition and 8 and 9 units in another condition, this will lead to a change in scalp distribution that will produce a condition × electrode site interaction. Moreover, even if there is no difference in the magnitude of the generators across conditions, but there is a latency difference in one of the two generators across conditions, it is likely that the measured scalp distribution at any given time point will differ across conditions. Thus, you cannot draw strong conclusions about differences in generator sources on the basis of ANOVA results.

Recommendations

So, what should you do if you conduct a standard ANOVA and find a condition × electrode site interaction? In most cases, I would recommend simply reporting that you found the interaction and then saying very little about it. If the difference between conditions is largest at the scalp sites where the component is largest (e.g., if the difference between the difference waves in the bright and dim conditions was largest at the Pz electrode site), you can simply state that the pattern of results is approximately what would be expected if a single component varied in amplitude across conditions (without even performing a formal analysis on the normalized data). If reviewers press you to draw stronger conclusions, you could cite the Urbach and Kutas (2002, 2006) papers and say that strong conclusions cannot be drawn from condition × electrode site interactions. You may also want to do your analyses on difference waves. This should eliminate any contamination from the baseline period and minimize the number of components that are active in the data. For example, in the experiment shown in figure 10.1, you could measure P3 amplitude in rare-minus-frequent difference waves and look for an interaction between brightness and electrode site. But keep in mind that the use of difference waves does not eliminate all of the problems involved in normalization.

What if you want to conclude that the scalp distributions are truly different across conditions? If you are measuring the data from well-controlled difference waves, and the scalp distributions are markedly different after normalization, you are probably safe in concluding that the conditions differ in the relative strengths of multiple generators during the period of measurement. However, if the differences are subtle after normalization, it will be difficult to conclude that the generators differ. And no matter how large the differences are, it will always be difficult to draw very strong conclusions (e.g., that completely different generators are present in the different conditions) unless you have some additional source of evidence.