Chapter 8   The Use of fMRI to Assess Multisensory Integration

Authors

Thomas W. James and Ryan A. Stevenson.

Although scientists have only recently had the tools available to noninvasively study the neural mechanisms of multisensory perceptual processes in humans (Calvert et al. 1999), the study of multisensory perception has had a long history in science (James 1890; Molyneux 1688). Before the advent of neuroimaging techniques, such as functional magnetic resonance imaging (fMRI) and high-density electrical recording, the study of neural mechanisms, using single-unit recording, was restricted to nonhuman animals such as monkeys and cats. These groundbreaking neurophysiological studies established many principles for understanding multisensory processing at the level of single neurons (Meredith and Stein 1983), and continue to improve our understanding of multisensory mechanisms at that level (Stein and Stanford 2008).

It is tempting to consider that neuroimaging measurements, like blood oxygenation level–dependent (BOLD) activation measured with fMRI, are directly comparable with findings from single-unit recordings. Although several studies have established clear links between BOLD activation and neural activity (Attwell and Iadecola 2002; Logothetis and Wandell 2004; Thompson et al. 2003), there remains a fundamental difference between BOLD activation and single-unit activity: BOLD activation is measured from the vasculature supplying a heterogeneous *population* of neurons, whereas single-unit measures are taken from *individual* neurons (Scannell and Young 1999). The ramifications of this difference are not inconsequential because the principles of multisensory phenomena established using single-unit recording may not apply to population-based neuroimaging data (Calvert et al. 2000). The established principles must be tested theoretically and empirically, and where they fail, they must be replaced with new principles that are specific to the new technique.

8.1. PRINCIPLES OF MULTISENSORY ENHANCEMENT

Although the definitions of unisensory and multisensory neurons may seem intuitive, for clarity, we will define three different types of neurons that are found in multisensory brain regions. The first class of neurons is unisensory. They produce significant neural activity (measured as an increase in spike count above spontaneous baseline) with only one modality of sensory input, and this response is not modulated by concurrent input from any other sensory modality. The second class of neurons is bimodal (or trimodal). They produce significant neural activity with two or more modalities of unisensory input (Meredith and Stein 1983; Stein and Stanford 2008). With single-unit recording, bimodal neurons can be identified by testing their response with unisensory stimuli from two different sensory modalities. The premise is simple: if the neuron produces significant activity with both modalities, then it is bimodal. However, bimodal activation only implies a convergence of sensory inputs, not the integration of those inputs (Stein et al. 2009). Bimodal neurons can be further tested for multisensory integration by using multisensory stimuli. When tested with a multisensory stimulus, most bimodal neurons produce activity that is greater than the maximum activity produced with either unisensory stimulus or multisensory enhancement. The criterion usually used to identify multisensory enhancement is called the maximum criterion or rule (AV > Max(A,V)). A minority of neurons produce activity that is lower than the maximum criterion, which is considered multisensory suppression. Whether the effect is enhancement or suppression, a change in activity of a neuron when the subject is stimulated through a second sensory channel only occurs if those sensory channels interact. Thus, multisensory enhancement and suppression are

indicators that information is being integrated. The third class of neurons is subthreshold. They have patterns of activity that look unisensory when they are tested with only unisensory stimuli, but when tested with multisensory stimuli, show multisensory enhancement (Allman and Meredith 2007; Allman et al. 2008; Meredith and Allman 2009). For example, a subthreshold neuron may produce significant activity with visual stimuli, but not with auditory stimuli. Because it does not respond significantly with both, it cannot be classified as bimodal. However, when tested with combined audiovisual stimuli, the neuron shows multisensory enhancement and thus integration. For graphical representations of each of these three classes of neurons, see Figure 8.1.

A majority of bimodal and subthreshold neurons show multisensory enhancement (i.e., exceed the maximum criterion when stimulated with a multisensory stimulus); however, neurons that show multisensory enhancement can be further subdivided into those that are superadditive and those that are subadditive. Superadditive neurons show multisensory activity that exceeds a criterion that is greater than the *sum* of the unisensory activities (AV > Sum(A,V); Stein and Meredith 1993). In the case of subthreshold neurons, neural activity is only elicited by a single unisensory modality; therefore, the criterion for superadditivity is the same as (or very similar to) the maximum criterion. However, in the case of bimodal neurons, the criterion for superadditivity is usually much greater than the maximum criterion. Thus, superadditive bimodal neurons can show extreme levels of multisensory enhancement. Although bimodal neurons that are superadditive are, by definition, multisensory (because they must also exceed the maximum criterion), the majority of multisensory enhancing neurons are not superadditive (Alvarado et al. 2007; Perrault et al. 2003; Stanford et al. 2007). To be clear, in single-unit studies, superadditivity is not a criterion for identifying multisensory enhancement, but instead is used to classify the *degree* of enhancement.

8.2. SUPERADDITIVITY AND BOLD fMRI

BOLD activation is measured from the vasculature that supplies blood to a heterogeneous population of neurons. When modeling (either formally or informally) the underlying activity that produces BOLD activation, it is tempting to consider that all of the neurons in that population have similar response properties. However, there is little evidence to support such an idea, especially within multisensory brain regions. Neuronal populations within multisensory brain regions contain a mixture of unisensory neurons from different sensory modalities in addition to bimodal and subthreshold multisensory neurons (Allman and Meredith 2007; Allman et al. 2008; Barraclough et al. 2005; Benevento et al. 1977; Bruce et al. 1981; Hikosaka et al. 1988; Meredith 2002; Meredith and Stein 1983, 1986; Stein and Meredith 1993; Stein and Stanford 2008). It is this mixture of neurons of different classes in multisensory brain regions that necessitates the development of new criteria for assessing multisensory interactions using BOLD fMRI.

The first guideline established for studying multisensory phenomena specific to population-based BOLD fMRI measures was superadditivity (Calvert et al. 2000), which we will refer to here as the additive criterion to differentiate it from superadditivity in single units. In her original fMRI study, Calvert used audio and visual presentations of speech (talking heads) and isolated an area of the superior temporal sulcus that produced BOLD activation with a multisensory speech stimulus that was greater than the sum of the BOLD activations with the two unisensory stimuli (AV > Sum(A,V)). The use of this additive criterion was a departure from the established maximum criterion that was used in single-unit studies, but was based on two supportable premises. First, BOLD activation can be modeled as a time-invariant linear system, that is, activation produced by two stimuli presented together can be modeled by summing the activity produced by those same two stimuli presented alone (Boynton et al. 1996; Dale and Buckner 1997; Glover 1999; Heeger and Ress 2002). Second, the null hypothesis to be rejected is that the neuronal population does *not* contain multisensory neurons (Calvert et al. 2000, 2001; Meredith and Stein 1983). Using the additive criterion, the presence of multisensory neurons can be inferred (and the null hypothesis rejected) if activation with the multisensory stimulus exceeds the additive criterion (i.e., superadditivity).

The justification for an additive criterion as the null hypothesis is illustrated in Figure 8.2. Data in Figure 8.2 are

simulated based on single-unit recording statistics taken from Laurienti et al. (2005). Importantly, the data are modeled based on a brain region that does not contain multisensory neurons. A brain region that only contains unisensory neurons is *not* a site of integration, and therefore represents an appropriate null hypothesis. The heights of the two left bars indicate stimulated BOLD activation with unisensory auditory (A) and visual (V) stimulation. The next bar is the simulated BOLD activation with simultaneously presented auditory and visual stimuli (AV). The rightmost bar, Sum(A,V), represents the additive criterion. Assuming that the pools of unisensory neurons respond similarly under unisensory and multisensory stimulation (otherwise they would be classified as subthreshold neurons), the modeled AV activation is the same as the additive criterion.

For comparison, we include the maximum criterion (the Max(A,V) bar), which is the criterion used in single-unit recording, and sometimes used with BOLD fMRI (Beauchamp 2005; van Atteveldt et al. 2007). The maximum criterion is clearly much more liberal than the additive criterion, and the model in Figure 8.2 shows that the use of the maximum criterion with BOLD data could produce false-positives in brain regions containing only two pools of unisensory neurons and no multisensory neurons. That is, if a single voxel contained only unisensory neurons and no neurons with multisensory properties, the BOLD response will still exceed the maximum criterion. Thus, the simple model shown in Figure 8.2 demonstrates both the utility of the additive criterion for assessing multisensory interactions in populations containing a mixture of unisensory and multisensory neurons, and that the maximum criterion, which is sometimes used in place of the additive criterion, may inappropriately identify unisensory areas as multisensory.

It should be noted that the utility of the additive criterion applied to BOLD fMRI data is different conceptually from the superadditivity label used with single units. The additive criterion is used to identify multisensory interactions with BOLD activation. This is analogous to maximum criterion being used to identify multisensory interactions in single-unit activity. Thus, superadditivity with single units is not analogous to the additive criterion with BOLD fMRI. The term superadditivity is used with single-unit recordings as a label to describe a subclass of neurons that not only exceeded the maximum criterion, but also the superadditivity criterion.

8.3. PROBLEMS WITH ADDITIVE CRITERION

Although the additive criterion tests a more appropriate null hypothesis than the maximum criterion, in practice, the additive criterion has had only limited success. Some early studies successfully identified brain regions that met the additive criterion (Calvert et al. 2000, 2001), but subsequent studies did not find evidence for additivity even in known multisensory brain regions (Beauchamp 2005; Beauchamp et al. 2004a, 2004b; Laurienti et al. 2005; Stevenson et al. 2007). These findings prompted researchers to suggest that the additive criterion may be too strict and thus susceptible to false negatives. As such, some suggested using the more liberal maximum criterion (Beauchamp 2005), which, as shown in Figure 8.2, is susceptible to false-positives.

One possible reason for the discrepancy between theory and practice was described by Laurienti et al. (2005) and is demonstrated in Figure 8.3. The values in the bottom row of the table in Figure 8.3 are simulated BOLD activation. Each column in the table is a different stimulus condition, including unisensory auditory, unisensory visual, and multisensory audiovisual. The Sum(A,V) column is simply the sum of the audio and visual BOLD signals and represents the additive criterion (null hypothesis). The audiovisual stimulus conditions were simulated using five different models, the maximum model, the supermaximum model, the additive model, the superadditive model, and the Laurienti model. The first three rows of the table represent the contributions of different classes of neurons to BOLD activation, including auditory unisensory neurons (A cells), visual unisensory neurons (V cells), and audiovisual multisensory neurons (AV cells). To be clear, the BOLD value in the bottom-most row is the sum of the A, V, and AV cell's contributions. Summing these contributions is based on the assumption that voxels (or clusters of voxels) contain mixtures of unisensory and multisensory neurons, not a single class of neurons. Although the

"contributions" have no units, they are simulated based on the statistics of recorded impulse counts (spike counts) from neurons in the superior colliculus, as reported by Laurienti et al. (2005). Unisensory neurons were explicitly modeled to respond similarly under multisensory stimulation as they did under unisensory stimulation, otherwise they would be classified as subthreshold neurons, which were not considered in the models.

The five models of BOLD activation under audiovisual stimulation differed in the calculation of only one value: the contribution of the AV multisensory neurons. For the maximum model, the contribution of AV cells was calculated as the maximum of the AV cell contributions with visual and auditory unisensory stimuli. For the super-max model, the contribution of AV neurons was calculated as 150% of the AV cell contribution used for the maximum model. For the additive model, the contribution of AV cells was calculated as the sum of AV cell contributions with visual and auditory unisensory stimuli. For the superadditive model, the contribution of AV cells was calculated as 150% of the AV cell contribution used for the additive model. Finally, for the Laurienti model, the contribution of the AV cells was based on the statistics of recorded impulse counts. What the table makes clear is that, based on Laurienti's statistics, the additive criterion is too conservative, which is consistent with what has been found in practice (Beauchamp 2005; Beauchamp et al. 2004a, 2004b; Laurienti et al. 2005; Stevenson et al. 2007).

Laurienti and colleagues (2005) suggest three reasons why the simulated BOLD activation may not exceed the additive criterion based on the known neurophysiology: first, the proportion of AV neurons is small compared to unisensory neurons; second, of those multisensory neurons, only a small proportion are superadditive; and third, superadditive neurons have low impulse counts relative to other neurons. To exceed the additive criterion, the *average* impulse count of the pool of bimodal neurons must be significantly *superadditive* for population-based measurements to exceed the additive criterion. The presence of superadditive neurons in the pool is not enough by itself because those superadditive responses are averaged with other subadditive, and even suppressive, responses. According to Laurienti's statistics, the result of this averaging is a value somewhere between maximum and additive. Thus, even though the additive criterion is appropriate because it represents the correct null hypothesis, the statistical distribution of cell and impulse counts in multisensory brain regions may make it practically intractable as a criterion.

8.4. INVERSE EFFECTIVENESS

The Laurienti model is consistent with recent findings suggesting that the additive criterion is too conservative (Beauchamp 2005; Beauchamp et al. 2004a, 2004b; Laurienti et al. 2005; Stevenson et al. 2007); however, those recent studies used stimuli that were highly salient. Another established principle of multisensory single-unit recording is the law of inverse effectiveness. Effectiveness in this case refers to how well a stimulus drives the neurons in question. Multisensory neurons usually *increase* their proportional level of multisensory enhancement as the stimulus quality is degraded (Meredith and Stein 1986; Stein et al. 2008). That is, the multisensory gain increases as the "effectiveness" of the stimulus decreases. If the average level of multisensory enhancement of a pool of neurons increases when stimuli are degraded, then BOLD activation could exceed the additive criterion when degraded stimuli are used.

Figure 8.4 shows this effect using the simulated data from the Laurienti model (Figure 8.3). In the high stimulus quality condition, the simulated AV activation clearly does not exceed the additive criterion, indicated as Sum(A,V), and it can be seen that this is because of the subadditive contribution of the multisensory neurons. On the right in Figure 8.4, a similar situation is shown, but with less effective, degraded stimuli. In general, neurons in multisensory regions decrease their impulse counts when stimuli are less salient. However, the size of the decrease is different across different classes of neurons and different stimulus conditions (Alvarado et al. 2007). In our simulation, impulse counts of unisensory neurons were reduced by 30% from the values simulated by the Laurienti model. Impulse counts of bimodal neurons were reduced by 75% under unisensory stimulus conditions, and by 50% under multisensory stimulus conditions. This difference in reduction for bimodal neurons between unisensory and multisensory stimulus

conditions reflects inverse effectiveness, that is, the multisensory gain increases with decreasing stimulus effectiveness.

Using these reductions in activity with stimulus degradation, BOLD activation with the AV stimulus now exceeds the additive criterion. Admittedly, the reductions that were assigned to the different classes of neurons were chosen somewhat arbitrarily. There are definitely different combinations of reductions that would lead to AV activation that would not exceed the criterion. However, the reductions shown are based on statistics of impulse counts taken from single-unit recording data, and are consistent with the principle of inverse effectiveness reported routinely in the single-unit recording literature (Meredith and Stein 1986). Furthermore, there is empirical evidence from neuroimaging showing an increased likelihood of exceeding the additive criterion as stimulus quality is degraded (Stevenson and James 2009; Stevenson et al. 2007, 2009). Figure 8.5 compares AV activation with the additive criterion at multiple levels of stimulus quality. These are a subset of data from a study reported elsewhere (Stevenson and James 2009). Stimulus quality was degraded by parametrically varying the signal-to-noise ratio (SNR) of the stimuli until participants were able to correctly identify the stimuli at a given accuracy. This was done by embedding the audio and visual signals in constant external noise and lowering the root mean square contrast of the signals. AV activation exceeded the additive criterion at low SNR, but failed to exceed the criterion at high SNR.

Although there is significant empirical and theoretical evidence suggesting that the additive criterion is too conservative at high stimulus SNR, the data presented in Figure 8.5 suggest that the additive criterion may be a better criterion at low SNR. However, there are two possible problems with using low-SNR stimuli to assess multisensory integration with BOLD fMRI. First, based on the data in Figure 8.5, the change from failing to meet the additive criterion to exceeding the additive criterion is gradual, not a sudden jump at a particular level of SNR. Thus, the choice of SNR level(s) is extremely important for the interpretation of the results. Second, there may be problems with using the additive criterion with measurements that lack a natural zero, such as BOLD.

8.5. BOLD BASELINE: WHEN ZERO IS NOT ZERO

It is established procedure with fMRI data to transform raw BOLD values to percentage signal change values by subtracting the mean activation for the baseline condition and dividing by the baseline. Thus, for BOLD measurements, "zero" is not absolute, but is defined as the activation produced by the baseline condition chosen by the experimenter (Binder et al. 1999; Stark et al. 2001). Statistically, this means that BOLD measurements would be considered an interval scale at best (Stevens 1946). The use of an interval scale affects the interpretation of the additive criterion because of the fact that calculating the additive criterion is reliant on summing two unisensory activations and comparing with a single multisensory activation. Because the activation values are measured relative to an arbitrary baseline, the value of the baseline condition has a different effect on the summed unisensory activations than on the single multisensory activation. In short, the value of the baseline is subtracted from the additive criterion twice, but is subtracted from the multisensory activation only once (see Equation 8.3).

The additive criterion for audiovisual stimuli is described according to the following equation:

But, Equation 8.1 is more accurately described by

Equation 8.2 can be rewritten as

and then

Equation 8.4 clearly shows that the level of activation produced by the baseline condition influences the additive criterion. An increase in activation of the baseline condition causes the additive criterion to become more liberal (Figure 8.6). The fact that the additive criterion can be influenced by the activation of the experimenter-chosen baseline condition may explain why similar experiments from different laboratories produce different findings when

that criterion is used (Beauchamp 2005).

8.6. A DIFFERENCE-OF-BOLD MEASURE

We have provided a theoretical rationale for the inconsistency of the additive criterion for assessing multisensory integration using BOLD fMRI as well as a theoretical rationale for the inappropriateness of the maximum criterion as a null hypothesis for this same assessment. The maximum criterion is appropriate when used with single-unit recording data, but when used with BOLD fMRI data, which represent populations of neurons, cannot account for the contribution of unisensory neurons that are found in multisensory brain regions. Without being able to account for the heterogeneity of neuronal populations, the maximum criterion is likely to produce false-positives when used with a population-based measure such as fMRI.

Although the null hypothesis tested by the additive criterion is more appropriate than the maximum criterion, the additive criterion is not without issues. First, an implicit assumption with the additive criterion is that the average multisensory neuronal response shows a pattern that is superadditive, an assumption that is clearly not substantiated empirically. Second, absolute BOLD percentage signal change measurements are measured on an interval scale. An interval scale is one with no natural zero, and on which the absolute values are not meaningful (in a statistical sense). The relative *differences* between absolute values, however, are meaningful, even when the absolute values are measured on an interval scale. To specifically relate relative differences to the use of an additive criterion, imagine an experiment where A, V, and AV were not levels of a sensory modality factor, but instead A, V, and AV were three separate factors, each with at least two different levels (e.g., levels of stimulus quality). Rather than analyzing the absolute BOLD values associated with each condition, a relative difference measurement could be calculated between the levels of each factor, resulting in ΔA, ΔV, and ΔAV measurements. The use of relative differences alleviates the baseline problem because the baseline activations embedded in the measurements cancel out when a difference operation is performed across levels of a factor. If we replace the absolute BOLD values in Equation 8.1 with BOLD differences, the equation becomes

Note that the inequality sign is different in Equation 8.5 than in Equation 8.1. Equation 8.1 is used to test the directional hypothesis that AV activation exceeds the additive criterion. Subadditivity, the hypothesis that AV activation is less than the additive criterion, is rarely, if ever, used as a criterion by itself. It has used been used in combination with superadditivity, for instance, showing that a brain region exceeds the additive criterion with semantically congruent stimuli but does not exceed the additive criterion with semantically incongruent stimuli (Calvert et al. 2000). This example (using both superadditivity and subadditivity), however, is testing two directional hypotheses, rather than testing one nondirectional hypothesis. Equation 8.5 is used to test a nondirectional hypothesis, and we suggest that it should be nondirectional for two reasons. First, the order in which the two terms are subtracted to produce each delta is arbitrary. For each delta term, if the least effective stimulus condition is subtracted from the most effective condition, then Equation 8.5 can be rewritten as $\Delta AV < \Delta A + \Delta V$ to test for inverse effectiveness, that is, the multisensory difference should be less than the sum of the unisensory differences. If, however, the differences were taken in the opposite direction (i.e., most effective subtracted from least effective), Equation 8.5 would need to be rewritten with the inequality in the opposite direction (i.e., $\Delta AV > \Delta A + \Delta V$). Second, inverse effectiveness may not be the only meaningful effect that can be seen with difference measures, perhaps especially if the measures are used to assess function across the whole brain. This point is discussed further at the end of the chapter (Figure 8.9).

Each component of Equation 8.5 can be rewritten with the baseline activation made explicit. The equation for the audio component would be

where A_1 and A_2 represent auditory stimulus conditions with different levels of stimulus quality. When Equation 8.5 is rewritten by substituting Equation 8.6 for each of the three stimulus conditions, all baseline variables in both the

denominator and the numerator cancel out, producing the following equation:

The key importance of Equation 8.7 is that the baseline variable cancels out when relative differences are used instead of absolute values. Thus, the level of baseline activation has no influence on a criterion calculated from BOLD differences.

The null hypothesis represented by Equation 8.5 is similar to the additive criterion in that the sum of two unisensory values is compared to a multisensory value. Those values, however, are relative differences instead of absolute BOLD percentage signal changes. If the multisensory difference is less (or greater) than the additive difference criterion, one can infer an interaction between sensory channels, most likely in the form of a third pool of multisensory neurons in addition to unisensory neurons. The rationale for using additive differences is illustrated in Figure 8.7. The simulated data for the null hypothesis reflect the contributions of neurons in a brain region that contains only unisensory auditory and visual neurons (Figure 8.7a). In the top panel, the horizontal axis represents the stimulus condition, either unisensory auditory (A) or visual (V), or multisensory audiovisual (AV). The subscripts 1 and 2 represent different levels of stimulus quality. For example, A_1 is high-quality audio and A_2 is low-quality audio. To relate these simulated data to the data in Figure 8.2 and the absolute additive criterion, the height of the stacked bar for AV_1 is the absolute additive criterion (or null hypothesis) for the high-quality stimuli, and the height of the AV_2 stacked bar is the absolute additive criterion for the low-quality stimuli. Those absolute additive criteria, however, suffer from the issues discussed above. Evaluating the absolute criterion at multiple levels of stimulus quality provides the experimenter with more information than evaluating it at only one level, but a potentially better way of assessing multisensory integration is to use a criterion based on *differences* between the high- and low-quality stimulus conditions. The null hypothesis for this additive differences criterion is illustrated in the bottom panel of Figure 8.7a. The horizontal axis shows the difference in auditory (ΔA), visual (ΔV), and audiovisual (ΔAV) stimuli, all calculated as differences in the heights of the stacked bars in the top panel. The additive differences criterion, labeled Sum(ΔA, ΔV), is also shown, and is the same as the difference in multisensory activation (ΔAV). Thus, for a brain region containing only two pools of unisensory neurons, the appropriate null hypothesis to be tested is provided by Equation 8.5.

The data in Figure 8.7b apply the additive differences criterion to the simulated BOLD activation data shown in Figure 8.4. Recall from Figure 8.4 that the average contribution of the multisensory neurons is subadditive for high-quality stimuli (A_1, V_1, AV_1), but is superadditive with low-quality stimuli (A_2, V_2, AV_2). In other words, the multisensory pool shows inverse effectiveness. The data in the bottom panel of Figure 8.7b are similar to the bottom panel of Figure 8.7a, but with the addition of this third pool of multisensory neurons to the population. Adding the third pool makes ΔAV (the difference in multisensory activation) significantly less than the additive differences criterion (Sum(ΔA, ΔV)), and rejects the null hypothesis of only two pools of unisensory neurons.

Figure 8.8 shows the same additive differences analysis performed on the empirical data from Figure 8.5 (Stevenson and James 2009; Stevenson et al. 2009). The empirical data show the same pattern as the simulated data. With both the simulated and empirical data, ΔAV was less than Sum(ΔA, ΔV), a pattern of activation similar to inverse effectiveness seen in single units. In singleunit recording, there is a positive relation between stimulus quality and impulse count (or effectiveness). This same relation was seen between stimulus quality and BOLD activation. Although most neurons show this relation, the multisensory neurons tend to show smaller decreases (proportionately) than the unisensory neurons. Thus, as the effectiveness of the stimuli decreases, the multisensory gain increases. Decreases in stimulus quality also had a smaller effect on multisensory BOLD activation than on unisensory BOLD activation, suggesting that the results in Figure 8.8 could (but do not necessarily) reflect the influence of inversely-effective neurons.

In summary, we have demonstrated some important theoretical limitations of the criteria commonly used in BOLD

fMRI studies to assess multisensory integration. First, the additive criterion is susceptible to variations in baseline. Second, the additive criterion is sensitive only if the average activity profile of the multisensory neurons in the neuronal population is superadditive, which, empirically, only occurs with very low-quality stimuli. A combination of these two issues may explain the inconsistency in empirical findings using the additive criterion (Beauchamp 2005; Calvert et al. 2000; Stevenson et al. 2007). Third, the maximum criterion tests a null hypothesis that is based on a homogeneous population of only multisensory neurons. Existing single-unit recording data suggest that multisensory brain regions have heterogeneous populations containing unisensory, bimodal, and sometimes, subthreshold neurons. Thus, the null hypothesis tested with the maximum criterion is likely to produce false-positive results in unisensory brain regions.

As a potential solution to these concerns, we have developed a new criterion for assessing multisensory integration using relative BOLD differences instead of absolute BOLD measurements. Relative differences are not influenced by changes in baseline, protecting the criterion from inconsistencies across studies. The null hypothesis to be tested is the sum of unisensory differences (additive differences), which is based on the assumption of a heterogeneous population of neurons. In addition to the appropriateness of the null hypothesis tested, the additive differences criterion produced positive results in known multisensory brain regions when tested empirically (Stevenson et al. 2009). Evidence for inverse effectiveness with audiovisual stimuli was found in known multisensory brain regions such as the superior temporal gyrus and inferior parietal lobule, but also in regions that have garnered less attention from the multisensory community, such as the medial frontal gyrus and parahippocampal gyrus (Figure 8.9). These results were found across different pairings of sensory modalities and with different experimental designs, suggesting the use of additive differences may be of general use for assessing integration across sensory channels. A number of different brain regions, such as the insula and caudate nucleus, also showed an effect that appeared to be the opposite of inverse effectiveness (Figure 8.9). BOLD activation in these brain regions showed the opposite relation with stimulus quality as sensory brain regions, that is, highquality stimuli produced less activation than low-quality stimuli. Because of this opposite relation, we termed the effect observed in these regions indirect inverse effectiveness. More research will be needed to assess the contribution of indirect inverse effectiveness to multisensory neural processing and behavior.

8.7. LIMITATIONS AND FUTURE DIRECTIONS

All of the simulations above made the assumption that BOLD activation could be described by a time-invariant linear system. Although there is clearly evidence supporting this assumption (Boynton et al. 1996; Dale and Buckner 1997; Glover 1999; Heeger and Ress 2002), studies using serial presentation of visual stimuli suggest that nonlinearities in BOLD activation may exist when stimuli are presented closely together in time, that is, closer than a few seconds (Boynton and Finney 2003; Friston et al. 1999). Simultaneous presentation could be considered just a serial presentation with the shortest asynchrony possible. In that case, the deviations from linearity with simultaneous presentation may be substantial. A careful examination of unisensory integration and a comparison of unisensory with multisensory integration could provide valuable insights about the linearity assumption of BOLD responses.

The simulations above were also based on only one class of multisensory neuron, the bimodal neurons, which respond with two or more sensory modalities. Another class of multisensory neurons has recently been discovered, which was not used in the simulations presented here. Subthreshold neurons respond to only one sensory modality when stimulated with unisensory stimuli. However, when stimulated with multisensory stimuli, these neurons show multisensory enhancement (Allman and Meredith 2007; Allman et al. 2008; Meredith and Allman 2009). Adding this class of neurons to the simulations may increase the precision of the predictions for population models with more than two populations of neurons. The goal of the simulations presented here, however, was to develop null hypotheses based on neuronal populations composed of only two unisensory pools of neurons. Rejecting the null hypothesis then implies the presence of at least one other pool of neurons besides the unisensory pools. In our simulations, we modeled that pool as bimodal; however, we could have also modeled subthreshold neurons or a combination of

bimodal and subthreshold neurons. Our impression is that the addition of subthreshold neurons to the simulations would not qualitatively change the results, because subthreshold neurons are found in relatively small numbers (less than the number of subadditive bimodal neurons), and their impulse counts are low compared to other classes of neurons (Allman and Meredith 2007).

The simulations above made predictions about levels of BOLD activation, but were based on principles of multisensory processing that were largely derived from spike (action potential) count data collected using single-unit recording. BOLD activation reflects a hemodynamic response, which itself is the result of local neural activity. The exact relationship, however, between neural activity and BOLD activation is unclear. There is evidence that increased spiking produces small brief local reductions in tissue oxygenation, followed by large sustained increases in tissue oxygenation (Thompson et al. 2003). Neural spike count, however, is not the only predictor of BOLD activation levels nor is it the best predictor. The correlation of BOLD activation with local field potentials is stronger than the correlation of BOLD with spike count (Heeger et al. 2000; Heeger and Ress 2002; Logothetis and Wandell 2004). Whereas spikes reflect the output of neurons, local field potentials are thought to reflect the postsynaptic potentials or input to neurons. This distinction between input and output and the relationship with BOLD activation raises some concerns about the relating studies using BOLD fMRI to studies using single-unit recording. Of course, spike count is also highly correlated with local field potentials, suggesting that spike count, local field potentials, and BOLD activation are all interrelated and, in fact, that the correlations among them may be related to another variable that is responsible for producing all of the phenomena (Attwell and Iadecola 2002).

Multisensory single-unit recordings are mostly performed in monkey and cat superior colliculus and monkey superior temporal sulcus or cat posterolateral lateral suprasylvian area (Allman and Meredith 2007; Allman et al. 2008; Barraclough et al. 2005; Benevento et al. 1977; Bruce et al. 1981; Hikosaka et al. 1988; Meredith 2002; Meredith and Stein 1983, 1986; Stein and Meredith 1993; Stein and Stanford 2008). With BOLD fMRI, whole-brain imaging is routine, which allows for exploration of the entire cortex. The principles that are derived from investigation of specific brain areas may not always apply to other areas of the brain. Thus, whole-brain investigation has the distinct promise of producing unexpected results. The unexpected results could be because of the different proportions of known classes of neurons, or the presence of other classes of multisensory neurons that have not yet been found with single-unit recording. It is possible that the indirect inverse effectiveness effect described above (Figure 8.9) may reflect the combined activity of types of multisensory neurons with response profiles that have not yet been discovered with single-unit recording.

8.8. CONCLUSIONS

We must stress that each method used to investigate multisensory interactions has a unique set of limitations and assumptions, whether the method is fMRI, high-density recording, single-unit recording, behavioral reaction time, or others. Differences between methods can have a great impact on how multisensory interactions are assessed. Thus, it should not be assumed that a criterion that is empirically tested and theoretically sound when used with one method will be similarly sound when applied to another method. We have developed a method for assessing multisensory integration using BOLD fMRI that makes fewer assumptions than established methods. Because BOLD measurements have an arbitrary baseline, a criterion that is based on relative BOLD differences instead of absolute BOLD values is more interpretable and reliable. Also, the use of BOLD differences is not limited to comparing across multisensory channels, but should be equally effective when comparing across unisensory channels. Finally, it is also possible that the use of relative differences may be useful with other types of measures, such as EEG, which also use an arbitrary baseline. However, before using the additive differences criterion with other measurement methods, it should be tested both theoretically and empirically, as we have done here with BOLD fMRI.

ACKNOWLEDGMENTS

REFERENCES

1. Allman B.L, Meredith M.A. Multisensory processing in "unimodal" neurons: Cross-modal subthreshold auditory effects in cat extrastriate visual cortex. Journal of Neurophysiology. 2007;98:545–9. [PubMed: 17475717]

2. Allman B.L, Keniston L.P, Meredith M.A. Subthreshold auditory inputs to extrastriate visual neurons are responsive to parametric changes in stimulus quality: Sensory-specific versus non-specific coding. Brain Research. 2008;1242:95–101. [PMC free article: PMC2645081] [PubMed: 18479671]

3. Alvarado J.C, Vaughan J.W, Stanford T.R, Stein B.E. Multisensory versus unisensory integration: Contrasting modes in the superior colliculus. Journal of Neurophysiology. 2007;97:3193–205. [PubMed: 17329632]

4. Attwell D, Iadecola C. The neural basis of functional brain imaging signals. Trends in Neurosciences. 2002;25:621–5. [PubMed: 12446129]

5. Barraclough N.E, Xiao D, Baker C.I, Oram M.W, Perrett D.I. Integration of visual and auditory information by superior temporal sulcus neurons responsive to the sight of actions. Journal of Cognitive Neuroscience. 2005;17:377–91. [PubMed: 15813999]

6. Beauchamp M.S. Statistical criteria in FMRI studies of multisensory integration. Neuroinformatics. 2005;3:93–113. [PMC free article: PMC2843559] [PubMed: 15988040]

7. Beauchamp M.S, Argall B.D, Bodurka J, Duyn J.H, Martin A. Unraveling multisensory integration: Patchy organization within human STS multisensory cortex. Nature Neuroscience. 2004a;7:1190–2. [PubMed: 15475952]

8. Beauchamp M.S, Lee K.E, Argall B.D, Martin A. Integration of auditory and visual information about objects in superior temporal sulcus. Neuron. 2004b;41:809–23. [PubMed: 15003179]

9. Benevento L.A, Fallon J, Davis B.J, Rezak M. Auditory–visual interaction in single cells in the cortex of the superior temporal sulcus and the orbital frontal cortex of the macaque monkey. Experimental Neurology. 1977;57:849–72. [PubMed: 411682]

10. Binder J.R, Frost J.A, Hammeke T.A, editors. et al. Conceptual processing during the conscious resting state. A functional MRI study. Journal of Cognitive Neuroscience. 1999;11:80–95. [PubMed: 9950716]

11. Boynton G.M, Engel S.A, Glover G.H, Heeger D.J. Linear systems analysis of functional magnetic resonance imaging in human V1. Journal of Neuroscience. 1996;16:4207–21. [PubMed: 8753882]

12. Boynton G.M, Finney E.M. Orientation-specific adaptation in human visual cortex. The Journal of Neuroscience. 2003;23:8781–7. [PubMed: 14507978]

13. Bruce C, Desimone R, Gross C.G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. Journal of Neurophysiology. 1981;46:369–84. [PubMed: 6267219]

14. Calvert G.A, Brammer M.J, Bullmore E.T, editors. et al. Response amplification in sensory-specific cortices during crossmodal binding. NeuroReport. 1999;10:2619–23. [PubMed: 10574380]

15. Calvert G.A, Campbell R, Brammer M.J. Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. Current Biology. 2000;10:649–57. [PubMed: 10837246]

16. Calvert G.A, Hansen P.C, Iversen S.D, Brammer M.J. Detection of audio-visual integration sites in humans by application of electrophysiological criteria to the BOLD effect. NeuroImage. 2001;14:427–38. [PubMed: 11467916]

17. Dale A.M, Buckner R.L. Selective averaging of rapidly presented individual trials using fMRI. Human Brain Mapping. 1997;5:329–40. [PubMed: 20408237]

18. Friston K.J, Zarahn E, Josephs O, Henson R.N, Dale A.M. Stochastic designs in event-related fMRI. NeuroImage. 1999;10:607–19. [PubMed: 10547338]

19. Glover G.H. Deconvolution of impulse response in event-related BOLD fMRI. NeuroImage. 1999;9:416–29. [PubMed: 10191170]

20. Heeger D.J, Huk A.C, Geisler W.S, Albrecht D.G. Spikes versus BOLD: What does neuroimaging tell us about neuronal activity? Nature Neuroscience. 2000;3:631–3. [PubMed: 10862687]

21. Heeger D.J, Ress D. What does fMRI tell us about neuronal activity? Nature Reviews Neuroscience. 2002;3:142–51. [PubMed: 11836522]

22. Hikosaka K, Iwai E, Saito H, Tanaka K. Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. Journal of Neurophysiology. 1988;60:1615–37. [PubMed: 2462027]

23. James W. The Principles of Psychology. New York: Henry Holt & Co; 1890.

24. Laurienti P.J, Perrault T.J, Stanford T.R, Wallace M.T, Stein B.E. On the use of superadditivity as a metric for characterizing multisensory integration in functional neuroimaging studies. Experimental Brain Research. 2005;166:289–97. [PubMed: 15988597]

25. Logothetis N.K, Wandell B.A. Interpreting the BOLD signal. Annual Review of Physiology. 2004;66:735–69. [PubMed: 14977420]

26. Meredith M.A. On the neuronal basis for multisensory convergence: A brief overview. Brain Research. Cognitive Brain Research. 2002;14:31–40. [PubMed: 12063128]

27. Meredith M.A, Allman B.L. Subthreshold multisensory processing in cat auditory cortex. NeuroReport. 2009;20:126–31. [PMC free article: PMC2839368] [PubMed: 19057421]

28. Meredith M.A, Stein B.E. Interactions among converging sensory inputs in the superior colliculus. Science. 1983;221:389–91. [PubMed: 6867718]

29. Meredith M.A, Stein B.E. Visual, auditory, and somatosensory convergence on cells in superior colliculus results in multisensory integration. Journal of Neurophysiology. 1986;56:640–62. [PubMed: 3537225]

30. Molyneux W. Letter to John Locke. In: de Beer E.S, editor. The correspondence of John Locke. Oxford: Clarendon Press; 1688.

31. Perrault T.J Jr., Vaughan J.W, Stein B.E, Wallace M.T. Neuron-specific response characteristics predict the magnitude of multisensory integration. Journal of Neurophysiology. 2003;90:4022–6. [PubMed: 12930816]

32. Scannell J.W, Young M.P. Neuronal population activity and functional imaging. Proceedings of the Royal Society of London. Series B. Biological Sciences. 1999;266:875–81. [PMC free article: PMC1689920] [PubMed: 10380677]

33. Stanford T.R, Stein B.E. Superadditivity in multisensory integration: Putting the computation in context. NeuroReport. 2007;18:787–92. [PubMed: 17471067]

34. Stark C.E, Squire L.R. When zero is not zero: The problem of ambiguous baseline conditions in fMRI. Proceedings of the National Academy of Sciences of the United States of America. 2001;98:12760–6. [PMC free article: PMC60127] [PubMed: 11592989]

35. Stein B.E, Meredith M.A. The Merging of the Senses. Cambridge, MA: The MIT Press; 1993.

36. Stein B.E, Stanford T.R. Multisensory integration: Current issues from the perspective of the single neuron. Nature Reviews Neuroscience. 2008;9:255–66. [PubMed: 18354398]

37. Stein B.E, Stanford T.R, Ramachandran R, Perrault T.J Jr., Rowland B.A. Challenges in quantifying multisensory integration: Alternative criteria, models, and inverse effectiveness. Experimental Brain Research. 2009;198:113–26. [PMC free article: PMC3056521] [PubMed: 19551377]

38. Stevens S.S. On the theory of scales of measurement. Science. 1946;103:677–80. [PubMed: 17750512]
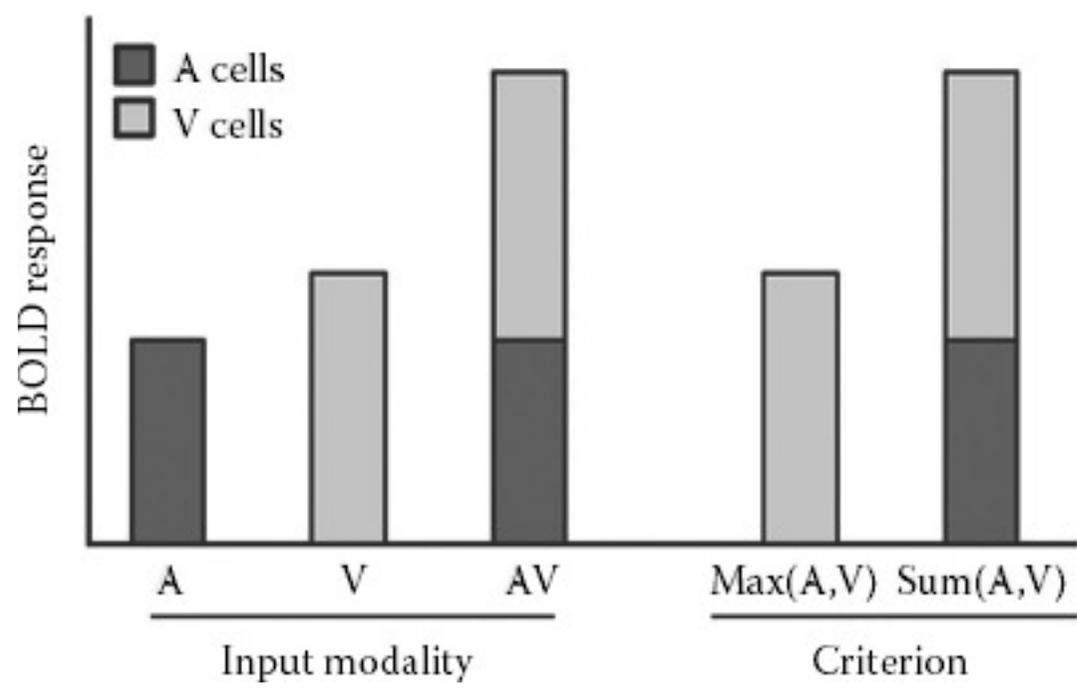
39. Stevenson R.A, James T.W. Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. NeuroImage. 2009;44:1210–23. [PubMed: 18973818]

40. Stevenson R.A, Geoghegan M.L, James T.W. Superadditive BOLD activation in superior temporal sulcus with threshold non-speech objects. Experimental Brain Research. 2007;179:85–95. [PubMed: 17109108]

41. Stevenson R.A, Kim S, James T.W. An additive-factors design to disambiguate neuronal and areal convergence: Measuring multisensory interactions between audio, visual, and haptic sensory streams using fMRI. Experimental Brain Research. 2009;198:183–94. [PubMed: 19352638]

42. Thompson J.K, Peterson M.R, Freeman R.D. Single-neuron activity and tissue oxygenation in the cerebral cortex. Science. 2003;299:1070–2. [PubMed: 12586942]

43. van Atteveldt N.M, Formisano E, Blomert L, Goebel R. The effect of temporal asynchrony on the multisensory integration of letters and speech sounds. Cerebral Cortex. 2007;17:962–74. [PubMed: 16751298]
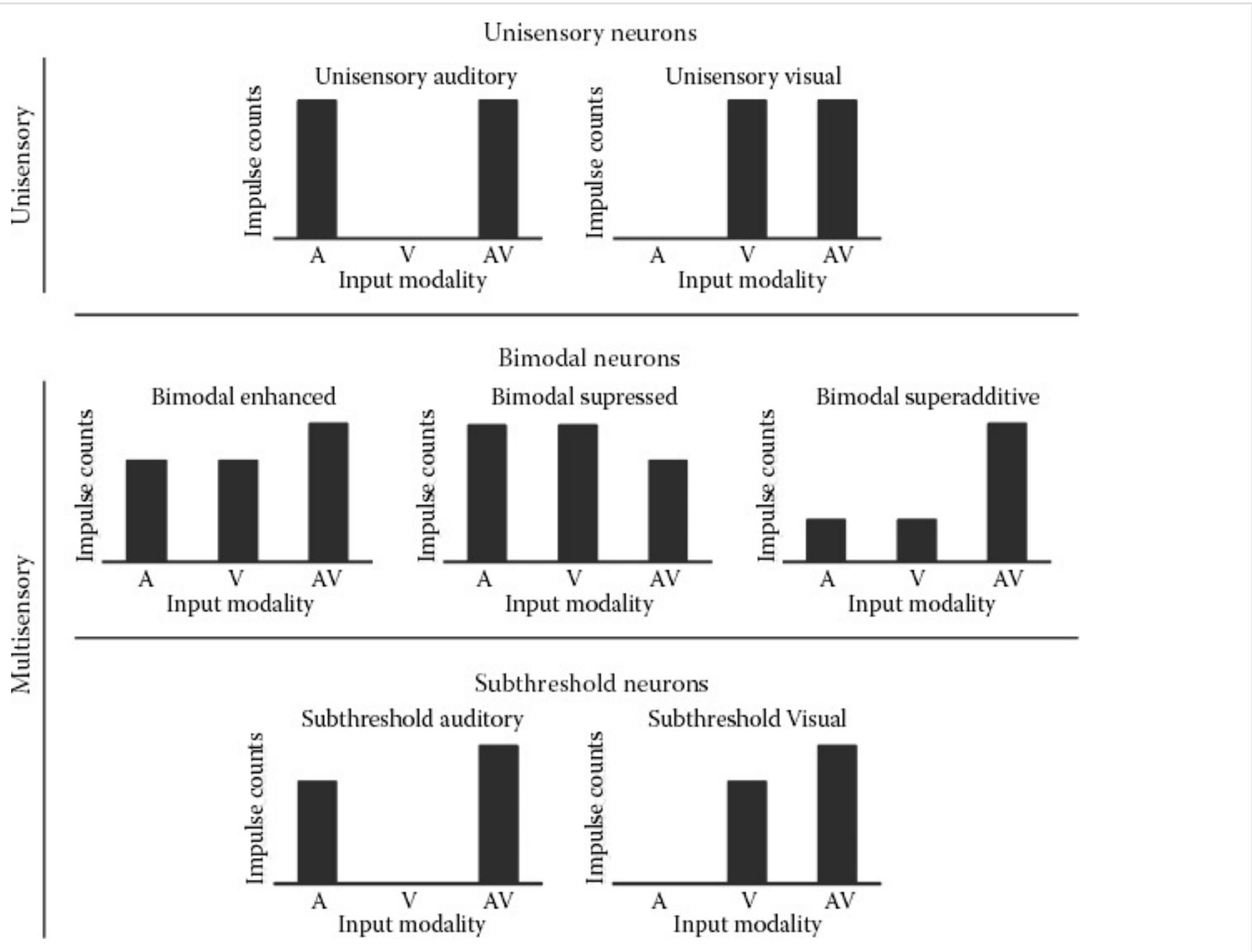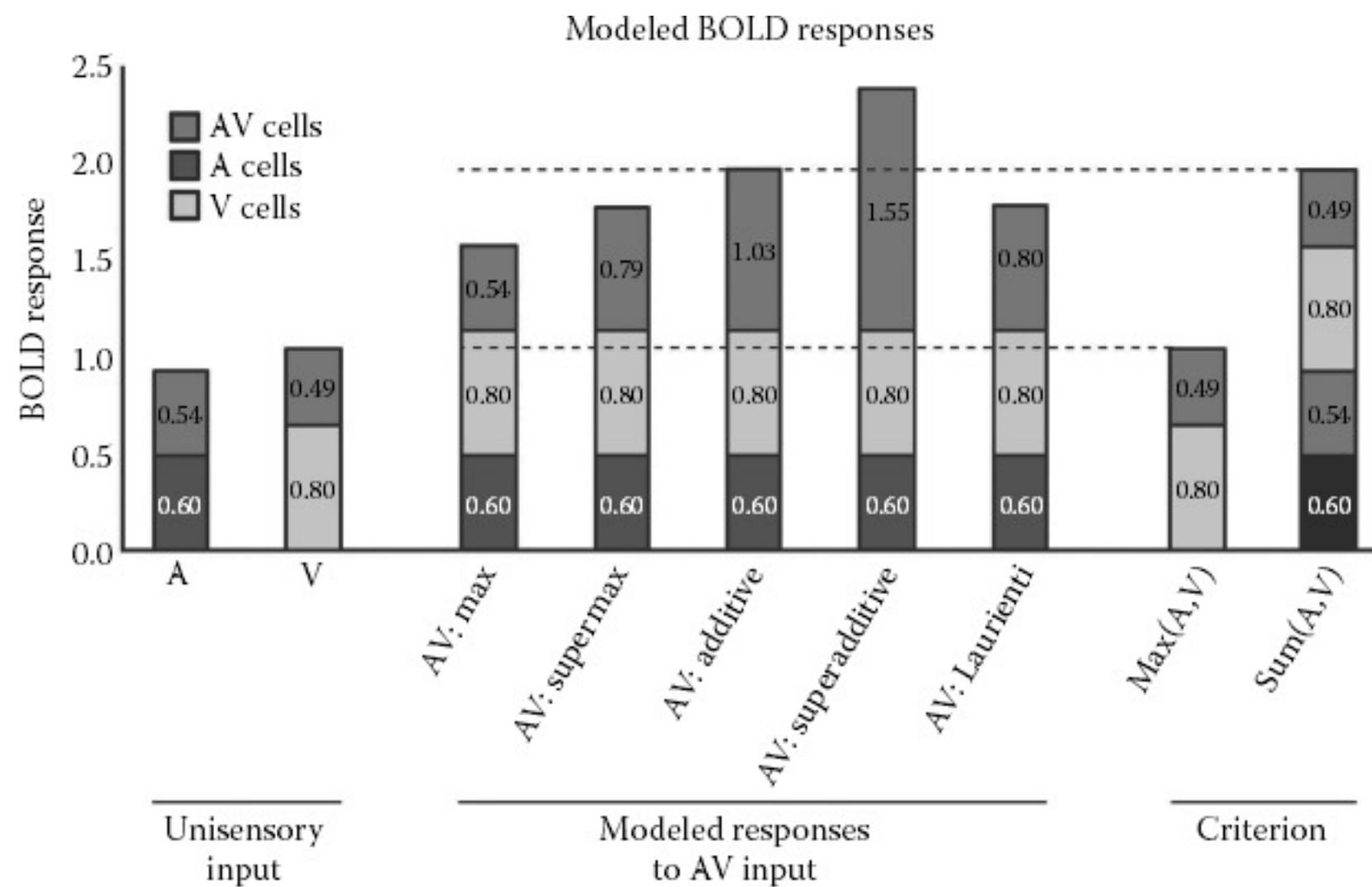
Figures

FIGURE 8.1

Activity profiles of neurons found in multisensory brain regions.

FIGURE 8.2

Criteria for assessing multisensory interactions in neuronal populations.

Modeled BOLD responses

Legend:
- AV cells
- A cells
- V cells

Unisensory input | Modeled responses to AV input | Criterion

Neural contributions by class

	A	V	Max	Supermax	Additive	Superadditive	Laurienti	Max(A,V)	Sum(A,V)
A cells	0.60	0.00	0.60	0.60	0.60	0.60	0.60	0.00	0.60
V cells	0.00	0.80	0.80	0.80	0.80	0.80	0.80	0.80	0.80
AV cells	0.54	0.48	0.54	0.79	1.03	1.88	0.80	0.49	1.03
BOLD	1.14	1.29	1.94	2.19	2.43	2.95	2.20	1.29	2.43

FIGURE 8.3

Models of BOLD activation with multisensory stimulation.
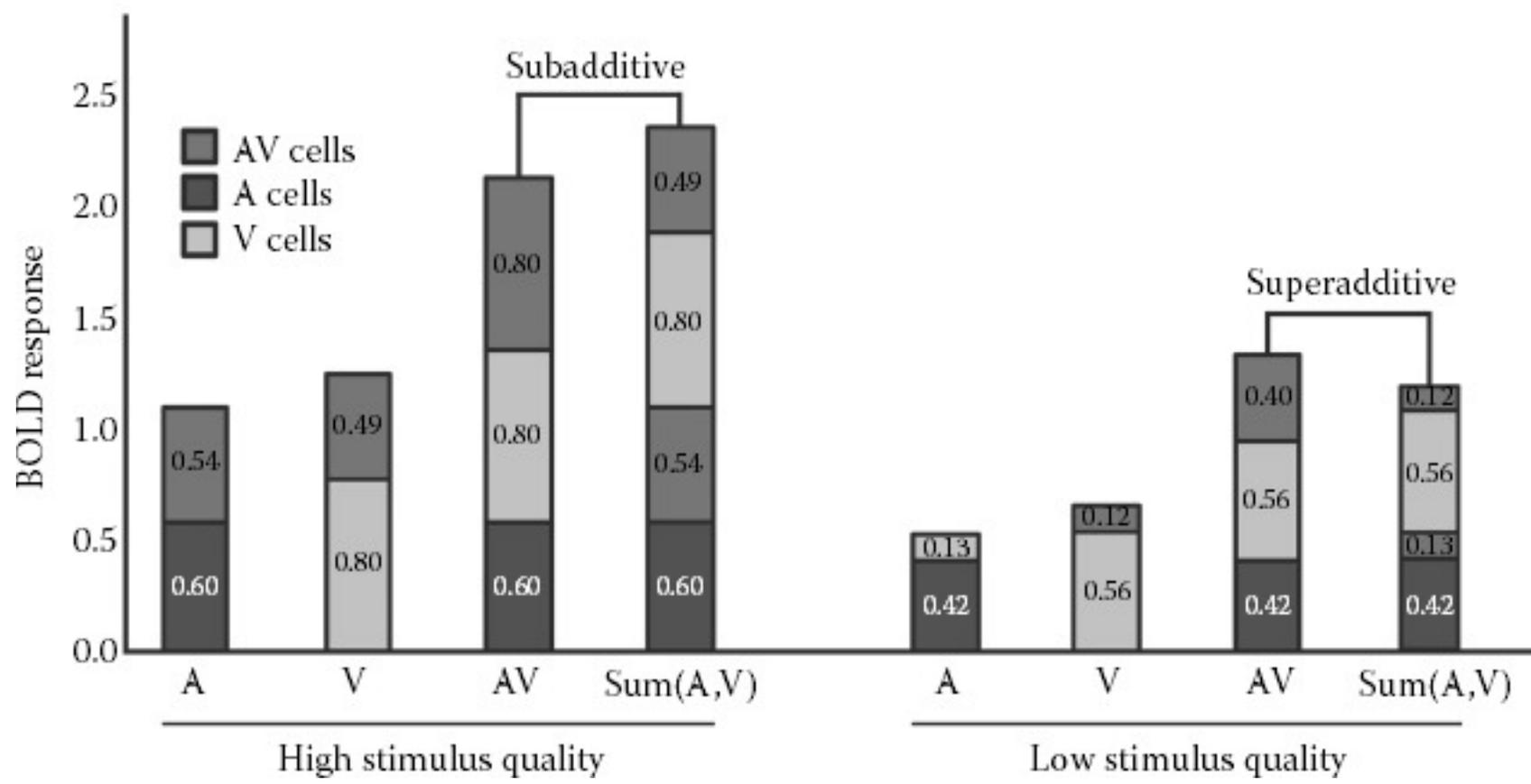
FIGURE 8.4

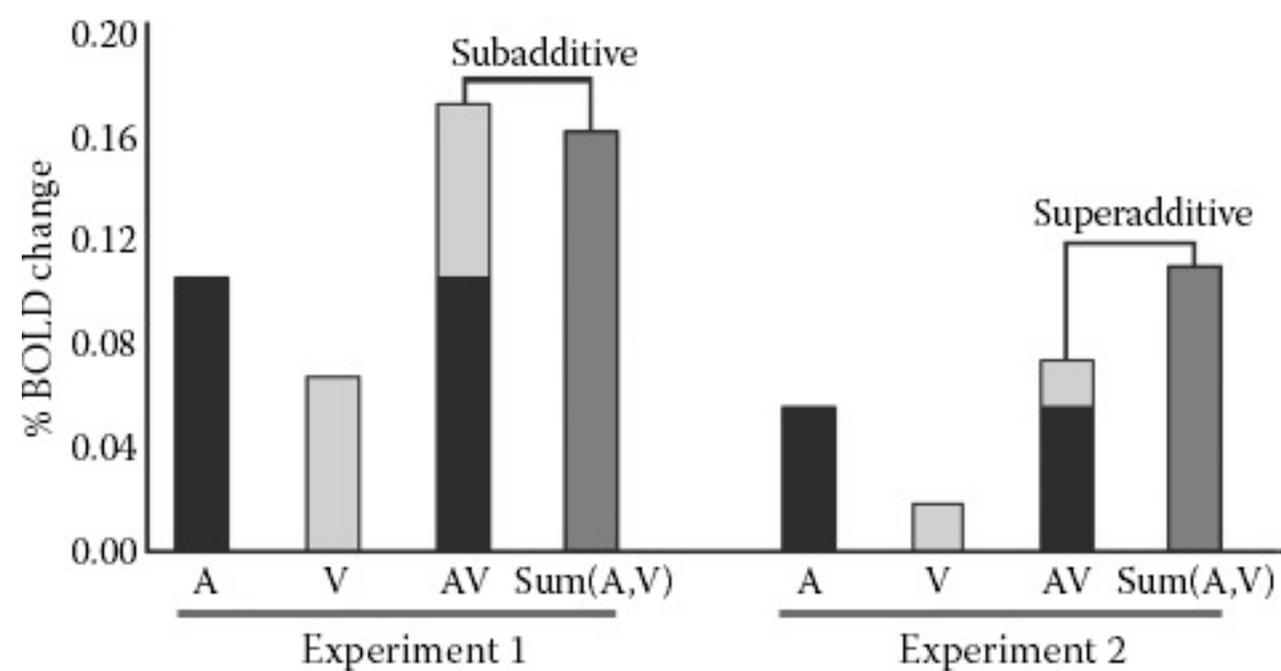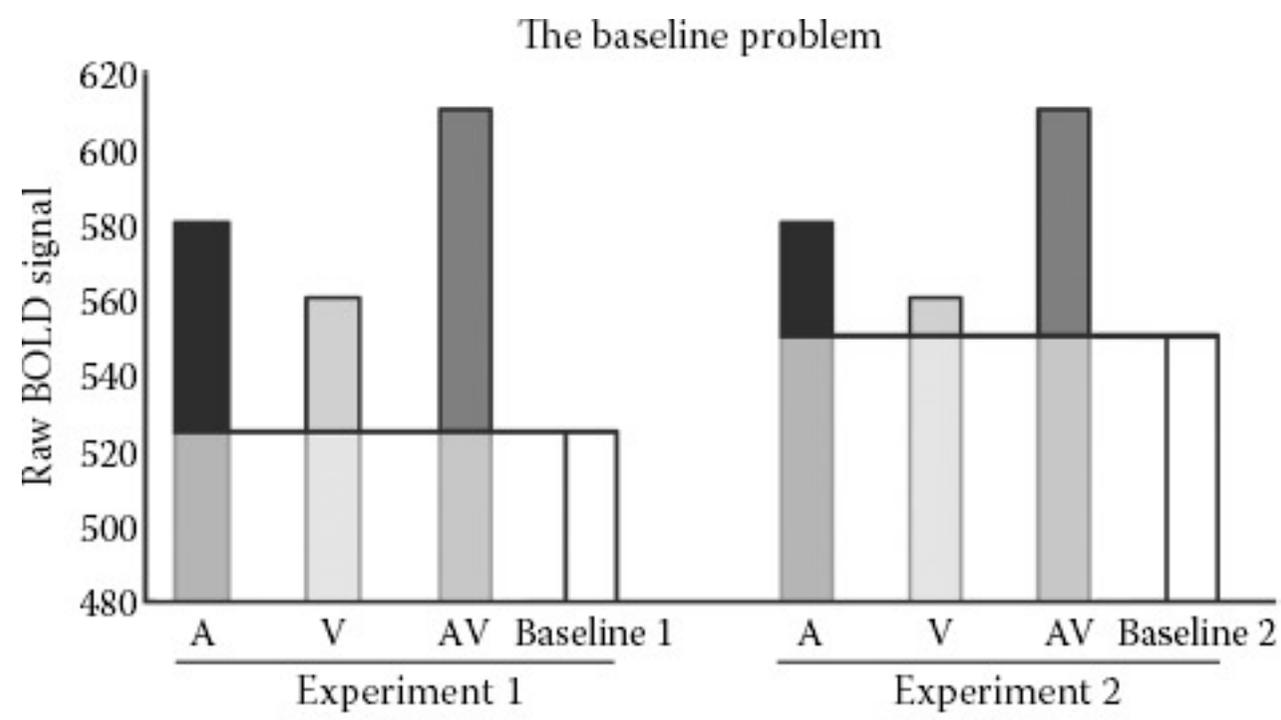Influence of inverse effectiveness on simulated multisensory BOLD activation.
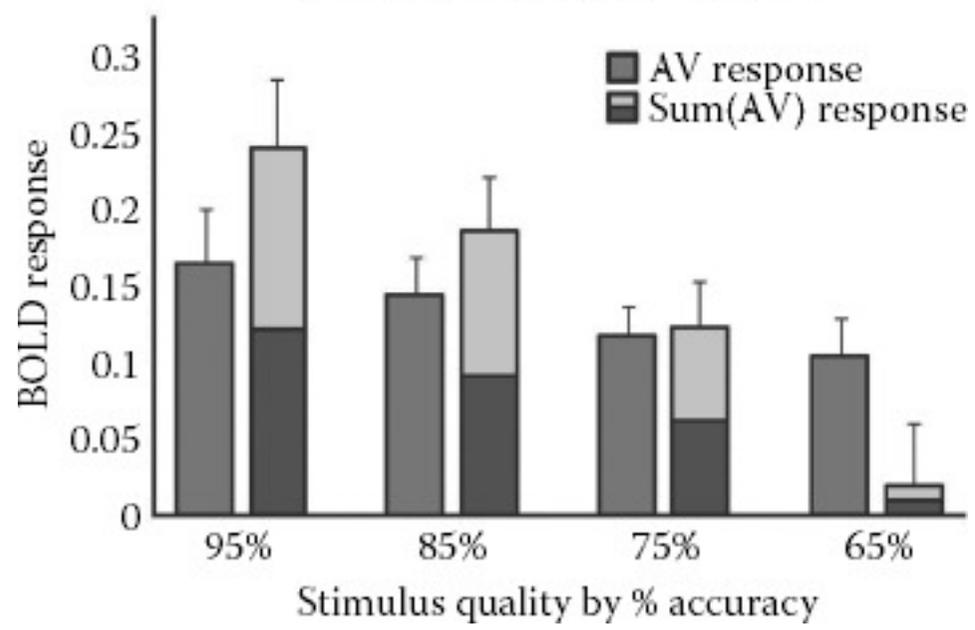
FIGURE 8.5

Assessing inverse effectiveness empirically with BOLD activation. These are a subset of data reported elsewhere. (From Stevenson, R.A. and James, T.W., *NeuroImage*, 44, 1210–23, 2009. With permission.)

$$AV > A + V \qquad (8.1)$$

$$\frac{\text{AV-baseline}}{\text{baseline}} > \frac{\text{A-baseline}}{\text{baseline}} + \frac{\text{V-baseline}}{\text{baseline}} \qquad (8.2)$$

FIGURE 8.6

Influence of baseline activation on additive criterion.

$$\text{AV} - \text{baseline} > \text{A} + \text{V} - 2 \times \text{baseline}, \tag{8.3}$$
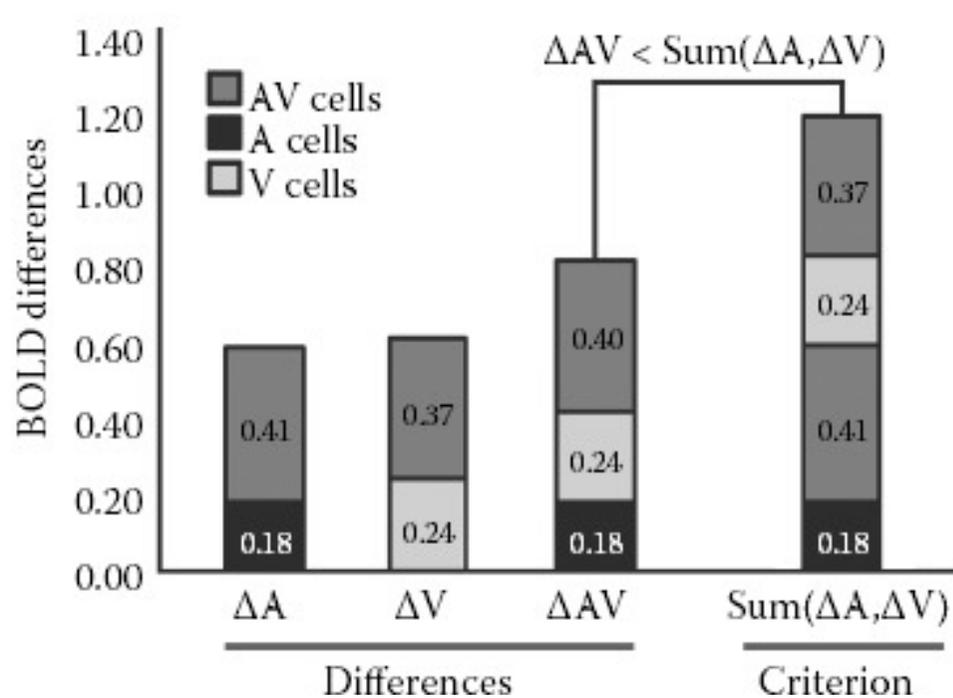
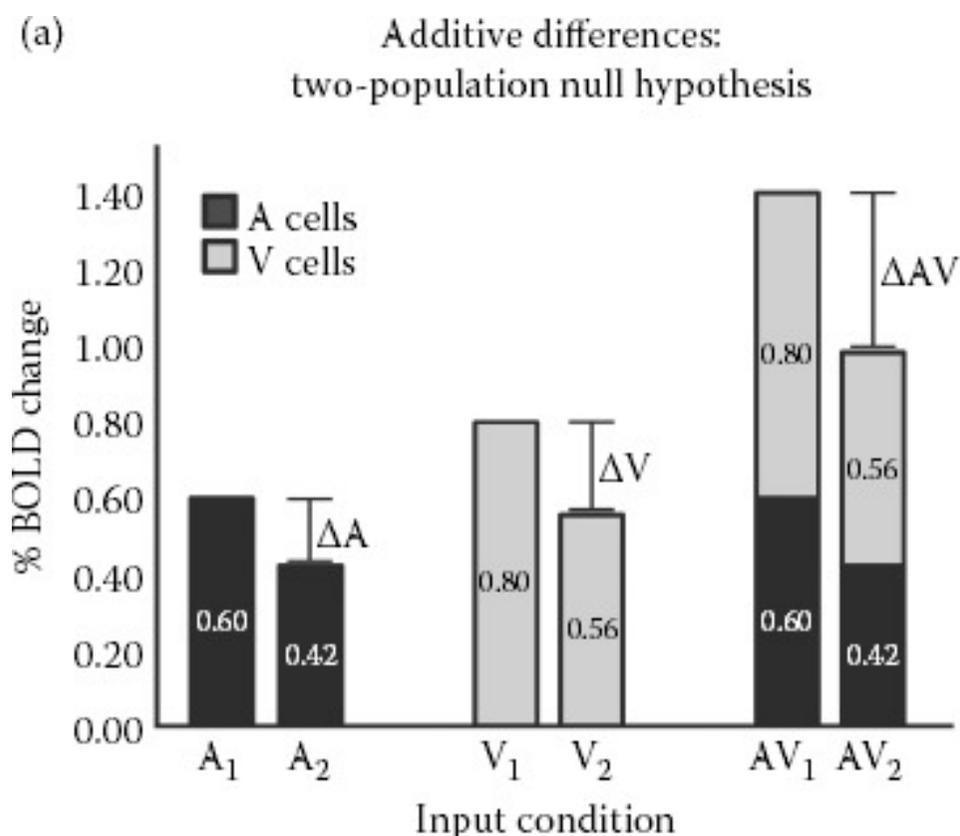$$AV > A + V - \text{baseline}. \tag{8.4}$$

$$\Delta AV \neq \Delta A + \Delta V. \qquad (8.5)$$

$$\Delta A = \frac{\left(A_1 - \text{baseline}\right)}{\text{baseline}} - \frac{\left(A_2 - \text{baseline}\right)}{\text{baseline}}, \qquad\qquad (8.6)$$
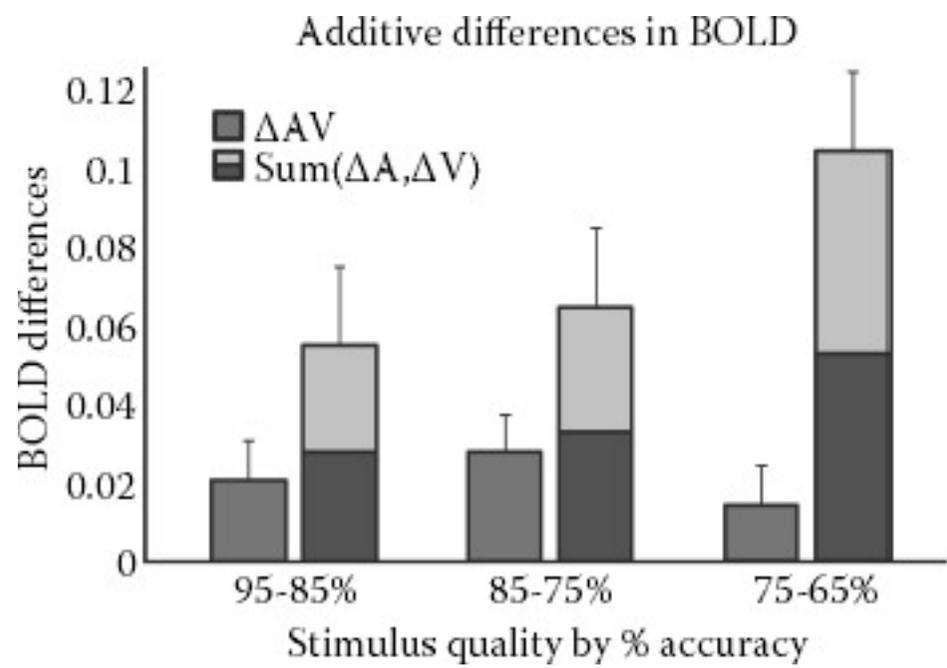
$$(AV_1 - AV_2) \neq (A_1 - A_2) - (V_1 - V_2). \qquad (8.7)$$
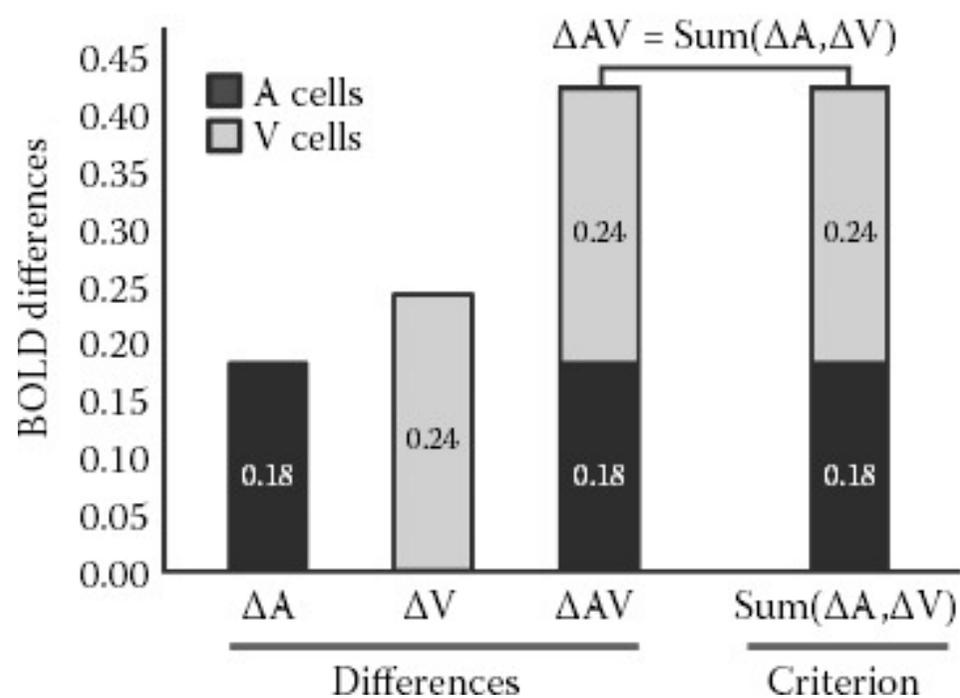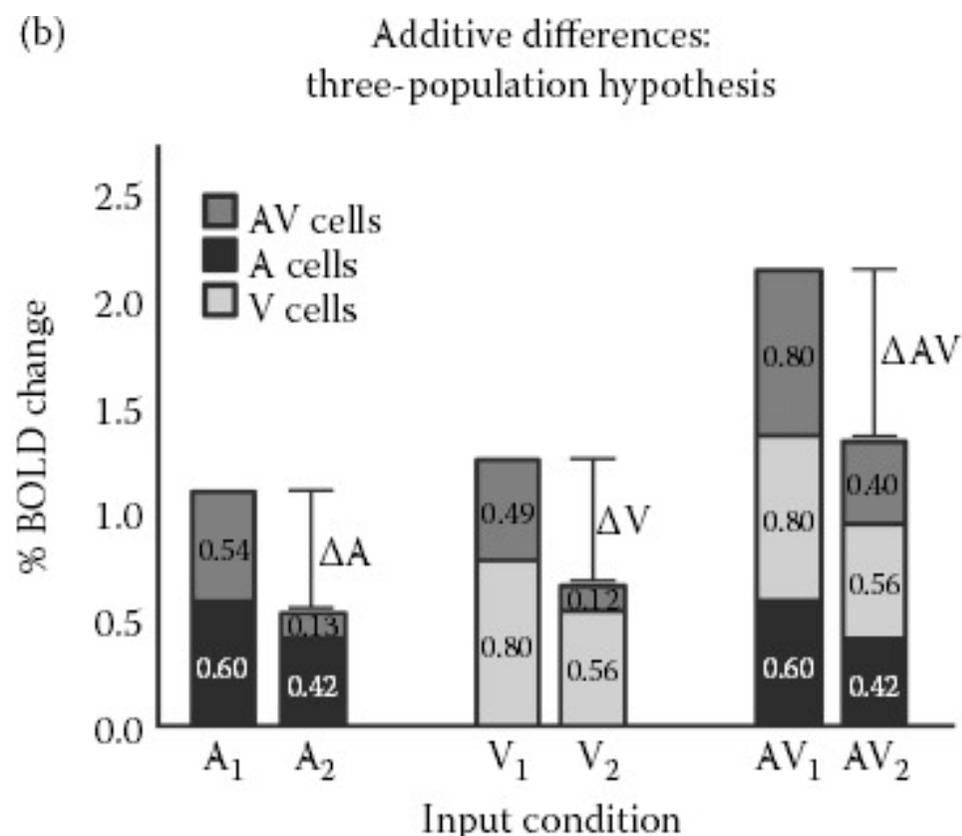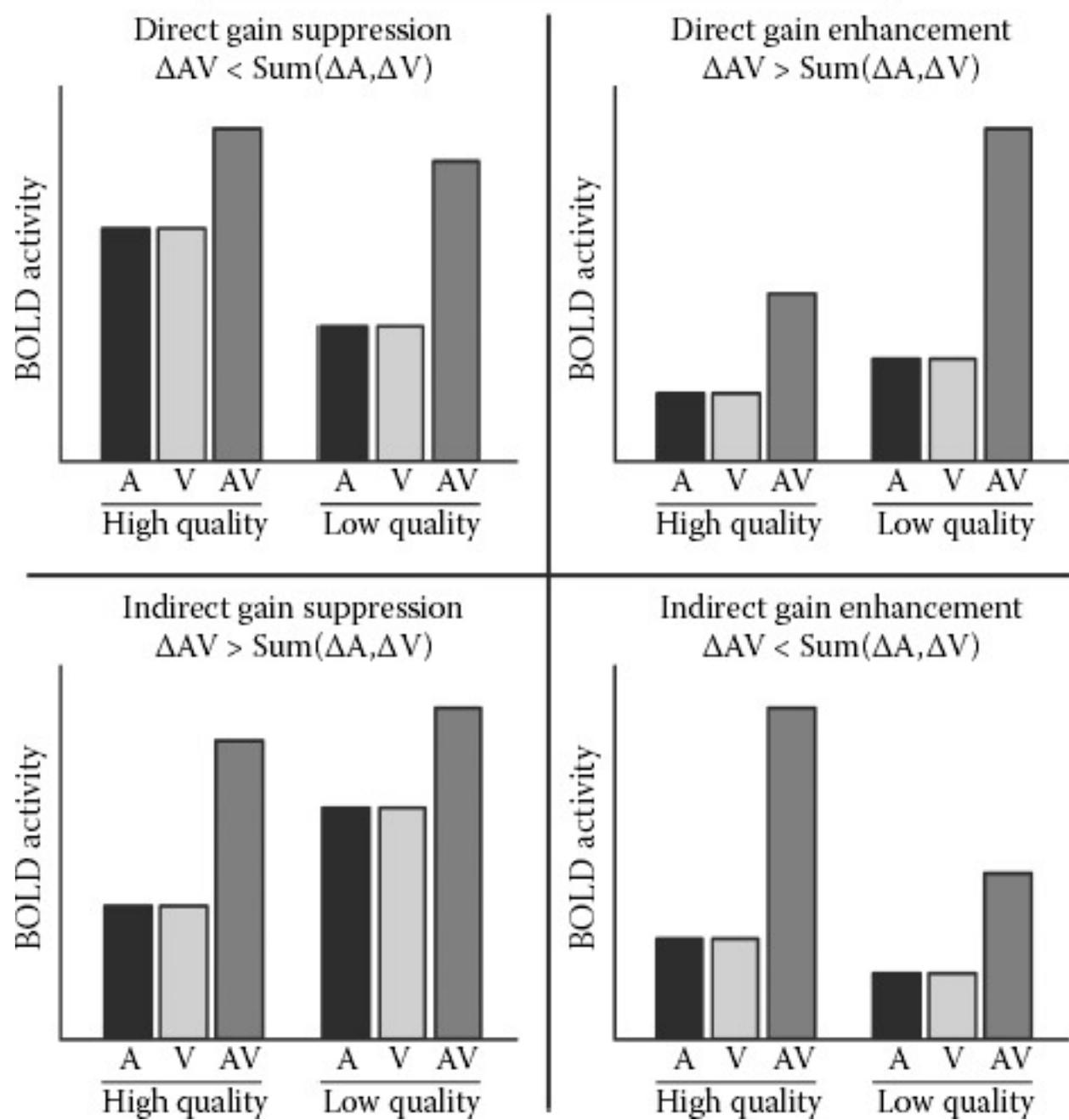
FIGURE 8.7

Additive differences criterion.

FIGURE 8.8

Assessing multisensory interactions empirically with additive differences.

Possible BOLD additive-difference interactions

Direct gain suppression
$\Delta AV < Sum(\Delta A, \Delta V)$

Direct gain enhancement
$\Delta AV > Sum(\Delta A, \Delta V)$

Indirect gain suppression
$\Delta AV > Sum(\Delta A, \Delta V)$

Indirect gain enhancement
$\Delta AV < Sum(\Delta A, \Delta V)$

(Each panel: x-axis labeled A V AV under High quality and A V AV under Low quality; y-axis labeled BOLD activity.)

FIGURE 8.9

A whole-brain statistical parametric map of regions demonstrating audiovisual neuronal convergence as assessed by additive differences criterion.