

The Integration of Object Levels and Their Content: A Theory of Global/Local Processing and Related Hemispheric Differences

Ronald Hübner and Gregor Volberg
Universität Konstanz

This article presents and tests the authors' integration hypothesis of global/local processing, which proposes that at early stages of processing, the identities of global and local units of a hierarchical stimulus are represented separately from information about their respective levels and that, therefore, identity and level information have to be integrated at later stages. It further states that the cerebral hemispheres differ in their capacities for these binding processes. Three experiments are reported in which the integration hypothesis was tested. Participants had to identify a letter at a prespecified level with the viewing duration restricted by a mask. False reporting of the letter at the nontarget level was predicted to occur more often when the integration of identity and level could fail. This was the case. Moreover, visual-field effects occurred, as expected. Finally, a multinomial model was constructed and fitted to the data.

Object perception is a complex task for the human visual system. Accordingly, the investigation of object perception comprises various aspects. For instance, because most objects in our environment are organized hierarchically—that is, lower level objects (e.g., hands, arms, legs) are nested within higher level objects (e.g., bodies)—one aspect concerns the relation between the perception of the whole and its parts. Does the visual system process lower level units first and then construct higher level units, as the Structuralists (e.g., Wundt, 1874) assumed? Or is the global shape of an object perceived first and then parsed into its local components, as the Gestalt psychologists (e.g., Wertheimer, 1922, 1923) proposed? This question is still in dispute (e.g., Sanocki, 2001).

Another aspect concerns the fact that objects usually consist not only of wholes and parts but also of features such as form, color, movement, and so forth. We know from neurophysiology that some of these features are processed in specialized areas in the brain (e.g., Livingstone & Hubel, 1988; Zeki, 1978, 1993). This implies that at some stage of processing, the features are represented separately in the mental system. Consequently, they have to be integrated at some stage for construction of a mental representation of the whole object. This is the so-called *binding problem* (Treisman & Gelade, 1980).

In some sense, the two aspects of object perception just mentioned are related. Often, the low-level and high-level units of an object are also regarded as local and global features of that object, respectively, and it has been investigated, for instance, whether

these features are processed in succession or in parallel. Moreover, as for other features, it has been suggested that global and local features or units are processed in different channels or pathways in the brain. It has even been proposed that they are processed in different cerebral hemispheres (for overviews, see Van Kleeck, 1989; Yovel, Yovel, & Levy, 2001). In this case also, a distributed processing would require the integration of global and local units in order to obtain a complete object representation. However, in the global/local research considered here, not much attention has been paid to this integration. Rather, the processing of individual features and their order or relative strength have been of interest, whereas the question of how whole objects are represented has largely been neglected.

However, as is shown in the present article, distributed processing and binding play an important role even if consideration is restricted to the individual processing of global and local units. In a series of experiments and by means of multinomial models we demonstrate in this article that, at least under certain circumstances, the units are processed and represented independently of their levels at an early stage of processing. Consequently, a binding process is required in order to link the units to their respective levels. Furthermore, we hypothesize that the cerebral hemispheres differ with respect to this integration process. We contrast this binding hypothesis with what we call the *standard view*. In this view, global and local features are processed in different streams or channels and, therefore, are uniquely related to their respective levels. In other words, the standard view assumes that level information and its content are coded in combination. Although this assumption has been explicitly formulated by only some researchers (e.g., Lamb & Yund, 1996; Robertson, 1996), it seems to represent one of the core assumptions in almost all theories of global/local processing.

Before we present our alternative model in detail and its supporting data, we provide an overview of the relevant literature. We have tried to integrate the different aspects of the research into a coherent theoretical framework in order to illustrate the standard

Ronald Hübner and Gregor Volberg, Department of Psychology, Universität Konstanz, Konstanz, Germany.

This research was supported by Grant Hu 432/7–4 from the Deutsche Forschungsgemeinschaft (DFG) to Ronald Hübner.

We thank Lawrence Ward, William Prinzmetal, Dirk Vorberg, and Marco Steinhauser for their helpful comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Ronald Hübner, Universität Konstanz, Fachbereich Psychologie, Fach D29, D-78457, Konstanz, Germany. E-mail: ronald.huebner@uni-konstanz.de

view and its shortcomings. First, however, we describe the stimuli and tasks usually applied in global/local studies.

Stimuli and Tasks in Global/Local Studies

An example of the stimuli usually used in global/local studies is shown on the left in Figure 1. It consists of a compound letter whose global letter *H* is constructed from identical local letters *E*. Although any symbols can be used as forms, letters were chosen in most cases. Such stimuli were introduced by Kinchla (1974) but were first systematically applied in the investigation of global/local processing by Navon (1977).

Because the aim of most global/local studies was to investigate perceptual processes, the effects of various stimulus parameters were examined. However, these effects cannot be observed directly. Therefore, observers have to perform a task and to produce overt responses, which can then be used to test specific hypotheses about the perceptual processes. Consequently, any conclusion drawn with respect to the investigated perceptual hypotheses is based on the whole task (see also Kimchi, 1992; Ward, 1983). Despite this fact, little attention has been paid to the effects of task parameters.

To consider a typical global/local procedure, assume that letters are used as units. A frequently applied task is then the speeded classification of the letter at a prespecified target level. Categories are usually defined by a mapping of the letters onto specific responses. For instance, if there are two letters, then one letter might require a left-button press, whereas the other requires a right-button press. With a single letter in the display, classification is usually an easy task. After some training, each letter automatically triggers its associated response. However, for compound letters, in which two letters are presented simultaneously, the situation is more complex. In this case, the two letters of a stimulus can be mapped onto different responses. Therefore, to select the correct response for such *incongruent* stimuli, the target level must be taken into account. This is not necessarily the case for congruent stimuli, that is, stimuli in which both letters are mapped onto the same response.

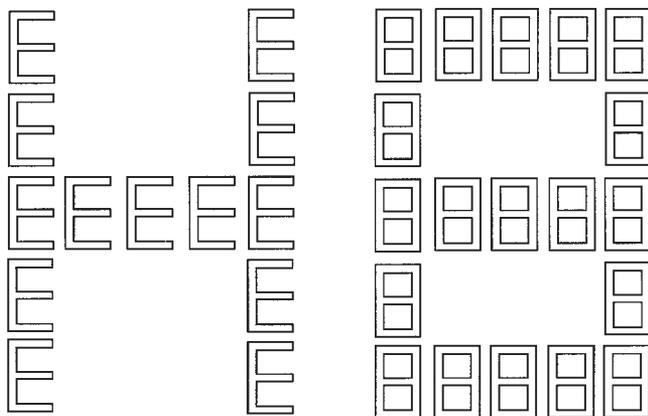


Figure 1. The pattern on the left depicts an example of the hierarchical letters used as stimuli. The pattern on the right shows the mask used for disrupting stimulus processing.

If there are only two letters in the letter set, then congruency implies that the letters at both levels are identical. To avoid this confound, one can use more than two letters. Moreover, in order to obtain a baseline for the congruency effects, one can also use neutral symbols as units for the nontarget level.

Relative Level Strength

By applying such a procedure, Navon (1977) observed that global letters were identified faster than local ones. Moreover, the congruency effect, that is, the latency difference between responses to incongruent and congruent stimuli, was larger for responses to local units than for those to global units. From this *global advantage*, Navon concluded that global features are processed first by early perceptual mechanisms. However, later studies showed that the global advantage is not a general phenomenon but depends on various stimulus factors such as spatial uncertainty (Grice, Canham, & Boroughs, 1983; Lamb & Robertson, 1988), stimulus distortion (Hoffman, 1980), size ratio (Kinchla & Wolfe, 1979), local density (Martin, 1979b), retinal position (Lamb & Robertson, 1988), and exposure duration (Paquet & Merikle, 1988). Accordingly, one can also easily produce a local advantage (e.g., Hoffman, 1980; Lamb & Robertson, 1988).

These results suggest that the advantage of a certain level does not depend on a single mechanism. Accordingly, relatively soon after Navon's (1977) seminal article appeared, several researchers proposed that attention also plays a role (e.g., Boer & Keuss, 1982; Hoffman, 1980; Miller, 1981; Ward, 1982). For instance, Miller (1981) assumed that response selection proceeds by a continuous evidence-accumulation process and that each response is represented by a corresponding evidence counter. When one of the counters exceeds a certain threshold, the corresponding response is triggered. He further assumed that attentional processes bias the input from the different sources into the counters. Because Miller's model is rather general, we have adopted its basic assumptions as a theoretical framework in order to interpret the different empirical results and to introduce our own view.

For performance in global/local tasks, it can be assumed that the units at both levels feed simultaneously into the evidence-accumulation process (Boer & Keuss, 1982; Hoffman, 1980; Miller, 1981), even though at different rates. The difference between the rates depends on various factors. As we have seen, stimulus parameters affect evidence accumulation in favor of one or the other level. However, because performance has to be goal directed, response selection cannot rely on pure stimulus-driven evidence accumulation. If this were the case, dominant units would almost always determine the response. Therefore, in order to respond to the units at a certain level, the respondent has to process the information of this level preferentially. Indeed, that participants can deliberately extract information from one level at the cost of the other level has been shown by so-called *attention operating characteristic* curves (Kinchla, Solis-Macias, & Hoffman, 1983; Ward, 1985). The observed curves are compatible with a simple binary-mixture model, in which it is assumed that the participants choose between two attentional strategies optimal for selecting information from the global or local level, respectively. Further evidence for attentional strategies has been provided, for instance, by cuing effects (Lamb, Pond, & Zahir, 2000; Robertson, Egly, Lamb, & Kerth, 1993; Stoffer, 1993).

These attentional strategies might be realized by biasing the different input sources in favor of the target level. A specific mechanism for controlling the input is filtering. For instance, absolute physical stimulus attributes can be used to select certain information (Broadbent, 1958). In global/local research, two such attributes have mainly been considered: size and spatial frequency.

The Use of Physical Stimulus Attributes

Because in many cases local and global units differ in absolute physical size, a possible filtering mechanism might be spatial attention (Lamb & Robertson, 1988; Stoffer, 1993) or regional selection (Robertson et al., 1993). On the basis of covert spatial orienting by means of an attentional spotlight (Posner & Cohen, 1984), it has been assumed that spatial attention can also operate like a zoom lens (Eriksen & St. James, 1986). This hypothesis has been supported by electrophysiological studies (e.g., Heinze, Hinrichs, Scholz, Burchert, & Mangun, 1998; Müller & Hübner, 2002) as well as by imaging studies (e.g., Sasaki et al., 2001; Weber, Schwarz, Kneifel, Treyer, & Buck, 2000). Thus, it seems that by adjusting the diameter of the attentional zoom lens, one can bias the evidence-accumulation processing in favor of one or the other level. However, like the position of the spotlight, the diameter of the zoom lens is not always under voluntary control. Rather, abrupt stimulus onsets determine the diameter automatically in favor of the global unit. This view is supported by a study in which the global advantage was eliminated by applying offset stimuli (Stoffer, 1994). On the other hand, the salience of the local units can be increased, for instance, by coloring some of their exemplars (Han, He, Yund, & Woods, 2001; Hübner & Malinowski, 2002).

Another physical and absolute attribute that often differs between global and local units is their spatial frequency content (Broadbent, 1977; Shulman, Sullivan, Gish, & Sakoda, 1986). Because it is known that low and high spatial frequencies are transmitted with different speeds in the visual system (Breitmeyer, 1975; Michimata, Okubo, & Mugishima, 1999; Tolhurst, 1975) and that there is also an asymmetric inhibition between the corresponding channels (Hughes, 1986), it has been proposed that differences in spatial frequency processing are responsible for the global advantage (cf. Hughes, Nozawa, & Kitterle, 1996). Furthermore, spatial frequencies have also been related to attention. On the one hand, low spatial frequencies seem to be processed preattentively and to serve as an early warning system (May, Gutierrez, & Harsin, 1995). On the other hand, attention can be voluntarily focused to specific spatial frequency channels (e.g., Hübner, 1996a, 1996b). Therefore, it seems that spatial frequencies are suitable for selecting units at a certain level (Ivry & Robertson, 1998; Robertson, 1996). However, in several studies, Lamb and Yund (1993, 1996, 2000) showed that such a selection does not depend on spatial frequency differences. Altogether, these results suggest that spatial frequencies might be used for the selection of local or global units. However, if this is not possible, because, for instance, the spatial frequency content is not sufficiently different between the units, then size or other discriminating features might be used.

Thus, a possible model that could account for the reported results is relatively simple. Information extracted from either stimulus level feeds simultaneously into an evidence-accumulation process. The relative contributions of the levels depend on

stimulus-driven as well as intentional factors. Goal-directed behavior is accomplished by means of an attentional mechanism that adjusts early filters in such a way that the units at the target level primarily feed the counters and, consequently, determine the response.

Limits of Physical Stimulus Attributes

Even though many data can be explained by such a simple account, it is not sufficient to explain all the data. This is because early filtering is not effective in all cases. Presumably, it is effective only if the stimulus position and the target level are fixed. In this case, it is easy to adjust the zoom lens or to bias other early filtering mechanisms in such a way that only the information of the target units feeds into the evidence-accumulation process. However, in other conditions, filtering by means of absolute stimulus attributes is not sufficient or at least is suboptimal. The main problem is that the attentional adjustments or task sets are not fully under voluntary control. For instance, if the target level is randomized across trials and, consequently, observers frequently have to shift from one task set to the other, the attentional adjustments are suboptimal, which is reflected in increased congruency effects (Hübner, 1997).

It can be assumed that the reduced performance under a variable target level is due to the inertia of the attentional adjustments. This assumption is supported by sequential effects, which were first observed by Ward (1982). He found an improved performance on trials in which the target level was the same as that on the previous trial. Moreover, Ward showed that this *level-readiness effect* was symmetric and was unaffected by spatial uncertainty, size of stimuli, necessity to use name codes for stimuli, divided or unitary attention to both levels or a particular level, and task (identification, same-different, or visual search). Later studies added further task (e.g., cue-stimulus interval, predictability, preparation) and stimulus (e.g., location, identity, spatial frequency) parameters that did not affect the level-readiness effect (Hübner, 2000; Lamb, London, Pond, & Whitt, 1998; Lamb & Yund, 2000). To account for his results, Ward (1982) proposed an attention switching/sharing model in which the relative speed of information processing at a given level depends on, besides the saliency of the features at that level, the relative amount of attentional resources allocated to that level. Moreover, he assumed that the processing at the target level uses the allocated attentional resources but also shifts the system to a state in which more resources are allocated to that level for future processing (see also Ward, 1985).

Robertson (1996) proposed another idea to explain level-readiness effects by assuming that the task set for responding to a given target level—what she called *attentional print*—persists over time and that, therefore, shifting from one task set to another produces costs in response time. She further assumed that the task sets consist of attentional weights for biasing certain spatial frequency channels. However, this latter assumption could not be confirmed (Hübner, 2000; Lamb & Yund, 1996). Therefore, Lamb and his coworkers (Lamb et al., 1998, 2000; Lamb & Yund, 1996) generalized the attentional print idea and proposed that sequential effects are due to the persisting activation of level-specific neural mechanisms. In essence, all ideas in this area are rather similar to the more general models considered in research on task-set shifting

(e.g., Allport, Styles, & Hsieh, 1994; Hübner, Futterer, & Steinhauer, 2001).

The Use of Relative Stimulus Attributes

A variable target level is only one example of how the efficiency of early filtering can be reduced. There are other even more severe conditions. For instance, if the stimulus is positioned randomly in different areas of the screen, then it is difficult to adjust the zoom lens optimally in advance (cf. Lamb & Robertson, 1988). Consequently, it is likely that the onset of the stimulus automatically attracts attention in favor of the global unit. In this case, early filtering might not be sufficient for achieving the usually required low error rate, at least for responding to local units. The case is even more difficult when the absolute size of the units varies. Such a variation is rather common in natural environments, where the observer and/or the objects can move; the global units in one moment can be of the same size as the local units in another moment. Thus, the relation between size or spatial frequency and level of an object is usually not constant.

However, even in cases in which early filtering is ineffective, observers are nevertheless capable of performing global/local tasks. This fact indicates that there are also other mechanisms that allow observers to select the correct response. Rather than relying on absolute physical features, these mechanisms presumably rely on derived stimulus features. For instance, even if the absolute features of the stimulus change, the relative hierarchical relations between the units remain invariant. Therefore, such relative stimulus attributes might also be used when the filtering of absolute features alone does not work. Some evidence for this conjecture has been provided by a study in which the absolute stimulus size varied across trials (Kim, Ivry, & Robertson, 1999). It turned out that level shift as well as size shift produced costs. This suggests that absolute as well as relative stimulus attributes are used for response selection.

These results show that the simple filter model outlined above is inappropriate. However, how would it have to be modified in order to cope with conditions in which early filtering alone does not work? Ineffective filtering means that both local and global units feed into the evidence-accumulation processes without much restriction. Consequently, dominant units almost always determine the response, and reliable goal-directed behavior is impossible. Therefore, to guarantee goal-directed performance even without efficient filtering, further mechanisms are required. We assume that these mechanisms utilize higher order representations. Up to now, mental representations have not played an important role for our simple filter model. The responses were merely represented by counters. However, to develop a more sophisticated model, it seems necessary to consider further representations. As in accounts for the Stroop phenomenon (cf. MacLeod, 1991), we assume that units such as letters automatically activate corresponding mental representations and that these representations serve as a link between stimulus and response. For instance, a compound letter activates corresponding letter representations, which in turn activate associated response counters. To adapt our simple model to this extension, we assume that early filtering now affects the activation of the letter representations. If it is effective, then the most active letter representation can be allowed to determine the response. Obviously, it is not necessary in this case to represent the

levels explicitly. The system trusts its filtering mechanism and simply assumes that the most active letter stems from the target level. However, the target level is represented implicitly by the selection of the task set that controls the filtering in favor of that level.

Now consider the case in which early filtering is not sufficient for reliable performance. In this case, a useful representation of the target level is required. The crucial question with respect to the objective of the present article is whether this information is already part of the letter representation, as the standard view assumes, or not. Advocates of the standard view could simply assume a further process that uses level information to select the letter whose level feature corresponds to the target level. This is possible because the letter representations also contain level information. However, in contrast to the standard view, our integration model assumes that early representations of the letters are relatively abstract and do not contain level information. As argued above, such abstract representations are sufficient if early filtering works. However, if this is not the case, the active letter representations have to be integrated with the representations of their levels. This results in representations that contain information about levels and their content.

What evidence do we have for our integration model? The model was originally derived from results collected in our studies concerned with hemispheric asymmetries for global/local processing (e.g., Hübner & Malinowski, 2002). For this reason, the relevant results in this area and how they support the integration theory are considered next.

Hemispheric Differences

It is widely assumed that the right and left hemispheres are specialized for processing the global and local levels, respectively (e.g., Ivry & Robertson, 1998). Evidence for this hypothesis comes from different sources, such as studies with neuropsychological patients (see Delis, Robertson, & Efron, 1986; Robertson & Lamb, 1991), response time studies (e.g., Hübner, 1997; Martin, 1979a), electrophysiological (event-related potential, or ERP) studies (e.g., Heinze & Münte, 1993; Malinowski, Hübner, Keil, & Gruber, 2002; Proverbio, Minniti, & Zani, 1998; Volberg & Hübner, 2004), and imaging studies (e.g., Fink et al., 1996; Heinze et al., 1998; Martinez et al., 1997).

However, there are also studies that did not find the expected hemispheric asymmetries. This is particularly true for response-time studies (e.g., Blanca-Mena, 1992; Boles & Karner, 1996; Van Kleeck, 1989). In their overview, Yovel et al. (2001) even came to the conclusion that the majority of studies with healthy participants revealed no hemispheric asymmetries for global/local processing. However, negative results can also be found in studies with patients (e.g., Lamb, Robertson, & Knight, 1989), in ERP studies (e.g., Han, Fan, Chen, & Zhuo, 1997; Han, He, & Woods, 2000; Johannes, Wieringa, Matzke, & Münte, 1996), and in imaging studies (e.g., Fink, Marshall, Halligan, Frith, & Frackowiak, 1997; Sasaki et al., 2001).

If we consider the electrophysiological and imaging studies, even those studies that found asymmetries differed with respect to the brain areas involved and the processing stages assumed. This difference led to a dispute as to whether asymmetries occur in areas reflecting early or late stages of processing. Some imaging

studies support the idea that hemispheric asymmetries occur at early stages (e.g., Fink et al., 1996), whereas others found evidence for hemispheric asymmetries at later stages (e.g., Heinze et al., 1998).

As for the global advantage, spatial frequencies have also been used to explain hemispheric asymmetries (for an overview, see Ivry & Robertson, 1998). For instance, Sergent (1982) proposed that the cerebral hemispheres differ in their capacity to process low and high spatial frequencies. However, at least for early perceptual processes such as contrast sensitivity, there seem to be no perceptual differences between the hemispheres (Peterzell, Harvey, & Hardyck, 1989). On the other hand, under conflicting stimulus conditions, hemispheric asymmetries have also been observed with compound gratings (Kitterle, Christman, & Conesa, 1993).

Taken together, the results with respect to hemispheric differences for global/local processing are inconclusive, which suggests that the phenomena are not as simple and robust as assumed. The results seem not to be invariant with respect to the different stimuli and procedures that have been applied in this area. In the best case, this indicates that asymmetries are real but difficult to observe. They might show up only under certain but still unknown conditions. Thus, an important endeavor is to determine the factors that reliably produce hemispheric asymmetries. In the search for such factors, various stimulus (e.g., Evert & Kmen, 2003) and task (e.g., Yovel et al., 2001) parameters have been manipulated with more or less success. However, with respect to the present objective, an interesting observation has already been made by Van Kleeck (1989) in his meta-analysis. He found that hemispheric asymmetries occurred more often for incongruent than for congruent stimuli. The positive effect of response conflicts for hemispheric asymmetries was later confirmed by behavioral experiments (e.g., Hübner, 1998) as well as by two ERP studies (Malinowski et al., 2002; Volberg & Hübner, 2004).

But how can congruency affect hemispheric asymmetries? Van Kleeck (1989) assumed that each hemisphere is capable of processing both local and global units but with different efficiencies. Within our framework, one would say that evidence accumulation proceeds at different rates depending on the visual field of stimulus presentation and on the target level. There could be various reasons for the different rates. Possible candidates are hemispheric differences with respect to feature extraction, early filtering, or feature processing. To consider an example, assume that a stimulus is presented in the left visual field and that the target level is local. In this case, a suboptimal amount of information feeds from the target level into the accumulation process. Moreover, it competes with the optimally extracted and processed information from the nontarget (global) level. Both factors increase the duration of response selection. However, if the target level is global, evidence accumulation is optimal for the target level and competes little with the suboptimal processed information at the nontarget (local) level. Analogous reasoning can be applied to the other conditions. In general, Van Kleeck assumed that the competition at the response selection stage caused by incongruent stimuli leads to an amplification of the hemispheric asymmetries.

A prediction of Van Kleeck's (1989) amplifier hypothesis is that the size of hemispheric asymmetries should covary, at least to some extent, with the size of the congruency effect. However, in a study by Hübner and Malinowski (2002), this prediction could not be confirmed. Although hemispheric asymmetries showed up only

when there was a congruency effect, the degree of congruency was relatively unimportant.

To explain this qualitative dependency of hemispheric asymmetries on congruency, Hübner and Malinowski (2002) speculated that at an early stage of processing, letter identity was already available but without level information, and that the hemispheres do not differ in this respect. As argued above, for congruent stimuli, the representation of letter identities is sufficient for a reliable response selection. This explains why there are no hemispheric asymmetries for this stimulus type. However, for responding to incongruent stimuli, the levels of the letters also have to be taken into account. This requires the integration of letter and level information at some stage, and it is at this stage that the hemispheres differ. Consequently, hemispheric asymmetries can be observed only when the binding of level and its content is necessary. Thus, the integration hypothesis nicely explains why incongruent stimuli are favorable in a qualitative way for hemispheric asymmetries.

A Masking Procedure

Even if one agrees that our integration hypothesis makes sense, up to now there has not been much supporting empirical evidence. Although our results with respect to hemispheric asymmetries support the hypothesis, the evidence is rather indirect. Furthermore, the interpretation of the results in favor of the integration theory is based on various assumptions that might be questioned. Therefore, we wanted to have a more direct test that was independent of hemispheric asymmetries. Fortunately, researchers interested in feature integration have developed a useful method for testing whether features are coded separately or in combination. The idea is to disturb or interrupt the integration process by masking. This produces different error types, which allow for the testing of various hypotheses. For instance, if features are coded separately at early stages and integration is prevented, then it might happen that different features combine accidentally. As a result, participants report so-called "illusory conjunctions" (Treisman & Schmidt, 1982), that is, the perception of feature combinations or objects that were not present in the display.

To test our integration hypothesis, we adopted the masking procedure and adapted it to the global/local paradigm. If letters are coded without level information, then observers should know the letters' identity before they know their level. Thus, if the letter at a certain level has to be reported but the integration processes are interrupted, then the letter at the other level should also be frequently reported. Such conjunction errors should at least occur more frequently than expected by the standard view, that is, when level and letter identity are represented in combination right from the beginning. The details become clear later when we introduce multinomial models. In any case, although derived from response-time data, our integration hypothesis is used to predict specific patterns of errors collected in an unspeeded identification task.

However, even though we hypothesize a separate coding of levels and letters, we do not assume illusory conjunctions in the sense that observers *see* a letter at one level although it appeared at the other level. For us it is sufficient to assume that binding might fail and that the observers then have to guess.

Next, three experiments are reported in which the masking paradigm was applied. In Experiment 1, the general method was

established and the results were used to test predictions derived from a simple multinomial model that reflects the standard view of global/local processing. Experiment 2 served to exclude an alternative account. Finally, a more complex data set was collected in Experiment 3 in order to fit a multinomial version of our integration hypothesis to the data.

Experiment 1

In this experiment, the participants had to identify a letter at a prespecified level of a hierarchical stimulus. Each stimulus was constructed from two of four different letters. In order to restrict the processing of the stimulus and to produce a sufficient number of errors, a mask (see right side of Figure 1) was presented shortly after stimulus onset. The possible responses with this procedure can be divided into three categories. First, responses are considered correct when the letter at the target level is reported, and these are coded as T. Second, when the letter at the nontarget level is reported, the responses are called “conjunction errors” and are coded as N. Third, a report of one of the two letters not present in the display is, following the tradition in illusory-conjunction research, called a “feature error” and is coded as O or \emptyset , respectively.

When a feature error occurs, then one can be relatively sure that no letter was seen at the target level. However, given a correct response, one cannot conclude that the letter at the target level was seen. This is because correct responses can also result from guessing. Thus, the actually interesting probabilities, such as the probability that the letter at the target level is seen, cannot be observed directly. Fortunately, there exists a method that allows one to estimate these probabilities. This method is based on multinomial models (cf. Batchelder & Riefer, 1990) and has already been successfully applied in illusory-conjunction research (Ashby, Prinzmetal, Ivry, & Maddox, 1996; Prinzmetal, Ivry, Beck, & Shimizu, 2002). However, parameter values were estimated only in Experiment 3. For the present experiment, testable predictions were derived and tested. This was possible even without knowing the exact parameter values.

Consider the multinomial processing tree diagram of the so-called “null model” (see Figure 2), which illustrates how the different response categories arise according to the standard view,

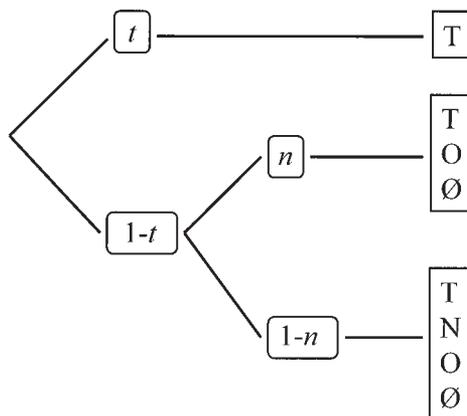


Figure 2. The multinomial processing tree diagram of the null model.

that is, under the assumption that levels and their contents are coded in combination. It predicts that whenever a letter is identified, its level is also known. Each branch in the diagram symbolizes a specific event or a specific set of events, and the parameters represent the probability of certain binary states. Here, the parameters t and n represent the probability that the letter at the target level and the letter at the nontarget level are seen, respectively. Thus, the event that the letter at the target level is seen has probability t . If the observer is in this state, a correct response (T) will occur. However, if one considers the other branches, it is obvious that a correct response might also occur when the letter is not seen. For instance, the probability that the letter at the nontarget level is seen but not the letter at the target level is given by the product of the probabilities of these corresponding two events, that is, by $(1 - t)n$. In this state, the observer has to guess the response by randomly choosing a letter from the set {T, O, \emptyset }. Thus, with probability $(1 - t)n(1/3)$, a correct response is given in this specific case. Finally, the total probability of a correct response is simply the sum of all of the probabilities of the pathways leading to the response T.

As can be seen, to utilize the multinomial methodology for our ideas, we have to translate the continuous evidence-accumulation models into models with discrete states. In the present case, it seems plausible to assume that the parameter t reflects a counter for the identity of the letter at the target level. If it exceeds a certain threshold value, the letter is “seen.” The probability that this state occurs depends, among other factors, on the efficiency of early filtering.

In order to calculate the proportion of conjunction errors (N), let us first assume that n (i.e., the probability that the letter at the nontarget level is seen) is zero, which means that filtering is perfect. When, for some reason, the letter at the target level cannot be identified, the observers have to guess. In this case, each of the four letters is chosen with probability 1/4. The choice of the letter at the target level leads to a correct response, whereas the choice of one of the other three letters produces an error. Consequently, on error trials, each letter not at the target level will be reported with a probability of 1/3. Because only one of these letters leads to a conjunction error, this simple version of the null model predicts that one third of the errors are conjunction errors.

This example shows that even without knowing the exact parameter values, it is possible to derive testable predictions from a multinomial model. If we now make the reasonable assumption that n is greater than zero, then even fewer conjunction errors are expected. This results from the fact that when the letter at the nontarget level is seen, it can be excluded from guessing, because according to the standard view the observer knows that it did not occur at the target level. However, because we do not know the parameter values, we will simply test whether the proportion of conjunction errors is greater than 1/3, which is a conservative test.

In contrast, our integration model assumes that letters and levels are coded separately at early stages and, therefore, must be integrated for the identification of the letter at the prespecified level. When this integration process is disrupted, the letters in the display might be seen without knowing at which level they occurred. Therefore, it can be expected that the letter at the nontarget level is falsely reported more frequently than either of the absent letters. Consequently, if our hypothesis is correct, the proportion of conjunction errors should be greater than 1/3.

However, our prediction holds only when the representations of both stimulus letters are activated. Therefore, we have to ensure that early filtering does not work very efficiently. Otherwise, the observers would simply have to report the mostly activated letter. Two measures were taken to make filtering difficult. One was the introduction of spatial uncertainty. Because we also wanted to investigate hemispheric asymmetries, stimulus position had to vary anyway. In addition, we randomized the target level.

With respect to hemispheric asymmetries, let us first consider the predictions of the standard view. According to the null model, hemispheric asymmetries with respect to the identification of the global and local letters are reflected by the individual parameters t and n for each hemisphere. Let us first assume that $n = 0$. In this case, only the parameter t differs between the hemispheres. As can be seen by considering Figure 2, any variation of t affects conjunction and feature errors to the same extent. To consider an example, assume that the target level is global and that in this case $t = .8$ for stimuli presented to the left visual field (LVF) and $t = .6$ for stimuli presented to the right visual field (RVF). With these parameters, we predict for conjunction errors corresponding rates of $[.25(1 - .8)] = .05$ and of $[.25(1 - .6)] = .1$, respectively. The rates for feature errors are .1 and .2. Thus, we expect twice as many errors for RVF stimuli as for LVF stimuli. Now, assume that n is greater than zero and that this parameter also differs between the hemispheres. In this case, the null model predicts smaller error rates and smaller visual-field (VF) effects. However, the reduction of VF effects is greater for conjunction errors than for feature errors. This is because n counteracts conjunction errors and differs in the opposite direction between the hemispheres relative to t . If we continue our example, but this time with $n = .2$ and $n = .4$ for LVF and RVF stimuli, respectively, then the corresponding conjunction error rates are now .04 and .06, whereas those for feature errors are .093 and .173. This shows that the reduction of VF effects is greater for conjunction errors. Thus, the null model predicts the same relative VF effects for both error types, or even smaller relative effects for conjunction errors.

According to our integration hypothesis, the hemispheres differ with respect to their capacity for integrating level and letter identity rather than in their capacity for filtering or letter identification. Therefore, we expect that VF effects are larger for conjunction than for feature errors.

Method

Participants

Sixteen students (9 women and 7 men; mean age = 22.3 years) from the Universität Konstanz, Germany, participated in the experiment. All had normal or corrected-to-normal vision and were paid for their participation.

Stimuli and Procedure

Stimuli were hierarchical letters (see Figure 1). Each stimulus was constructed from two out of four different letters (A, S, H, E). The size of the global letters was 4.28° of visual angle horizontally and 5.89° vertically. The respective size of the local letters was $0.71^\circ \times 0.93^\circ$. After the presentation of a cue (the letter l or g to indicate a local or global target level, respectively) for 300 ms and a subsequent blank interval of 300 ms, the stimulus was presented for 24 ms in white on a black background either to the LVF or to the RVF at an eccentricity of 3.03° (from the midline of

the screen to the center of the stimulus). VF and target level were randomized across trials. The stimulus-mask interval (SMI) was set to 96 ms at the beginning of the experiment and then was individually decreased as soon as a participant produced more than 70% correct responses in a block of trials. The mask remained present on the screen until the response occurred.

The task was to identify the letter at the cued level. Participants responded without time pressure by pushing one of four response buttons, each of which corresponded to a certain letter. The button-to-letter mapping was varied across participants to counteract any response biases. Participants were informed that each stimulus contained two different letters and that when they were sure that they had only seen the letter at the nontarget level, they should not report this letter but should randomly choose one of the remaining letters, because reporting the seen letter would produce an error. In addition, the participants received an auditory error feedback. Altogether, there were 16 blocks of 48 trials for each participant, which were run in a 1-hr session.

Results

The mean error rate was 37%. As can be seen in Figure 3, feature errors occurred in about 16% of the trials, whereas conjunction errors occurred in 22% of the trials. The rate of conjunction errors is significantly different, $F(1, 15) = 73.4, p < .01$, from the prediction by the null model (the gray bars in Figure 3).

As is obvious if one considers Figure 3, neither the rate of feature errors (local, 16.5%; global, 16.3%) nor that of conjunction errors (local, 22.0%; global, 22.9%) differed between the levels. However, there was a significant interaction in the expected direction between level and VF, $F(1, 15) = 16.2, p < .01$, for conjunction errors. When the letter at the local level had to be reported, it was more likely for LVF stimuli than for RVF stimuli that the letter at the global level was reported instead, whereas the opposite held for the global target level (see Figure 4, left graph). Although there was a similar trend for feature errors (see Figure 4, right graph), the interaction did not reach significance.

Discussion

The aim of Experiment 1 was to test whether conjunction errors (i.e., reporting the letter at the nontarget level instead of the letter

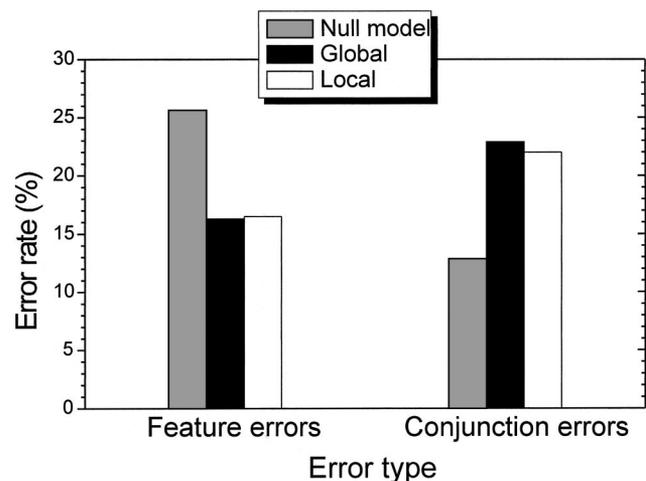


Figure 3. The percentage of conjunction and feature errors for global and local target letters in Experiment 1. The gray bars indicate the predictions by the null model.

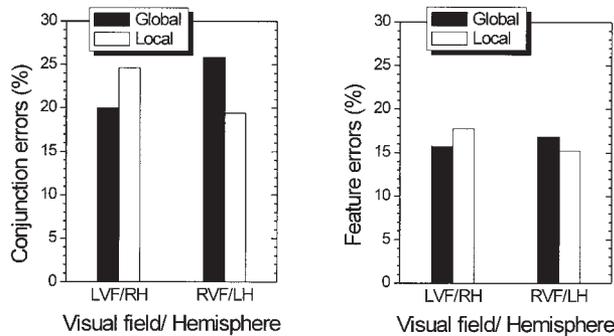


Figure 4. The interaction between visual field and target level for the conjunction errors (significant) and feature errors (not significant). LVF = left visual field; RVF = right visual field; RH = right hemisphere; LH = left hemisphere.

at the target level) occurred more often than predicted by the null model. This model represents the standard view that letter and levels are coded in combination. Thus, if a letter is perceived, its level is also known. The results clearly show that this was not the case. The number of conjunction errors largely exceeded the number predicted by the null model. Whereas the null model predicts 13% conjunction errors and 26% feature errors, we observed 22% conjunction errors and 16% feature errors. This supports our hypothesis that letters and levels are coded separately at early stages and that their integration takes place at later stages. If the integration is disturbed or prevented by a mask, as in the present experiment, then it is likely that on some trials the identities of both letters in the display are known but not the respective levels. Consequently, the letter at the nontarget level might falsely be reported, which produces a conjunction error. It also seems plausible to assume that the same error type occurs when only the letter at the nontarget level is seen but its level remains unknown. Because the participants focus their attention at a certain stimulus level, they expect that a single letter seen belongs to that level. This produces the tendency to report that letter.

Also, with respect to hemispheric asymmetries, the null model can be rejected. Contrary to its prediction, significant VF effects were present only for conjunction errors. This, however, is in accord with the integration hypothesis.

Thus, taken together, it seems that the present data strongly support the integration hypothesis. However, before we can be sure, we have to exclude a simple but compelling alternative account. One might argue that the standard view is still valid but that the participants confused the cues on some trials. Because the target level was randomized, a cue was presented at the beginning of each trial to indicate the relevant level. Thus, the possibility that the participants confused the cue on some trials and, therefore, falsely attended to the wrong level cannot be excluded. Consequently, participants may have reported the letter at the nontarget level, which would explain the increased number of conjunction errors.

How many cue confusions can be expected under a randomized-level condition? To obtain an estimate, we looked at the data of Hübner and Malinowski (2002). Although the participants had to produce speeded responses to unmasked stimuli in that study, the conditions in the critical interval from cue presentation to stimulus

presentation were similar to those in the present experiment. That is, there was a similar cue duration, there was a similar length of the cue–stimulus interval, and there were even similar stimuli. What is most important for the present objective, however, is that in Hübner and Malinowski's first experiment, two conditions were realized. In one condition the target level was blocked, whereas it was randomized in the other condition. For the blocked condition, it is reasonable to assume that the number of cue confusions is negligible (the error rate was 2.9%). Thus, if randomizing the levels produces cue confusions, then the number of cue confusions should be reflected by the difference in error rates between the two conditions. Therefore, we compared the error rates for incongruent stimuli (i.e., stimuli that contained two response-incompatible letters) because only for these stimuli would a confusion produce an error. It turned out that the error rate increased by 3.7%. Even though this increase is certainly not exclusively due to cue confusions, it still seems to be far too small to explain the relatively large number of conjunction errors in the present experiment. This raises the question of how many cue confusions would be necessary to account for our result.

To answer this question, we had to extend the null model. For this model, the processing tree shown in Figure 2 reflects the state in which the correct level is attended. Let us assume that this state occurs with probability c . Thus, with probability $1 - c$ the participants are in the state in which they attend to the wrong level. In this case a similar tree applies, but now T and N exchange their positions in the tree. For simplicity, let us again assume that $n = 0$. The probabilities $p(N)$ and $p(O)$ of conjunction and feature errors, respectively, can now be calculated by

$$p(N) = c(1 - t)(1/4) + (1 - c)t + (1 - c)(1 - t)(1/4)$$

and

$$p(O) = c(1 - t)(1/2) + (1 - c)(1 - t)(1/2).$$

The probability $p(T)$ for a correct response simply is $1 - [p(N) + p(O)]$. With these formulas we can search for the parameters c and t , given the probability $p(T)$, or calculate the probabilities $p(N)$ and $p(O)$ for certain parameters. For instance, if we assume, according to our estimate above, that $c = .96$ (i.e., that there are about 4% cue confusions) and search for the value of the parameter t (the probability of identifying the letter at the attended level) that produces the observed 61% of correct responses [i.e., $p(T) = .61$], then we find $t = .51$. However, with these parameters we predict 14% conjunction errors and 25% feature errors. Thus, this number of cue confusions hardly changes the prediction by the simple null model. If we now search for the value of the parameter c that predicts our data, we obtain $c = .79$. This value would produce 22% conjunction errors and 17% feature errors (t is .67 in this case). Thus, the participants would have had to confuse the cues in 21% of the trials. Compared with the estimated proportion of 3.7% cue-confusion trials in Hübner and Malinowski (2002), this is a large number. Even if we concede that speeded-response conditions might be favorable for reducing the number of cue confusions, it is hard to see how this could account for such a large difference. However, because we cannot definitively exclude the cue-confusion account, we tested a prediction of the extended null model in our second experiment.

Experiment 2

In our first experiment, we saw that the simple null model, which represents the standard view of global/local processing, cannot account for the data. However, if one additionally assumes that the participants attended to the wrong level on a certain number of trials, because they confused the cue, then the data could be explained. Even though it is highly unlikely that these trials occurred so frequently as would be necessary to account for the large number of observed conjunction errors, this model cannot be excluded yet. Therefore, our aim in the present experiment was to test a specific prediction of the extended null model.

This test should be achieved by varying the SMI across trials. It can be expected that the probability of identifying the letter at the attended level (the parameter t in the model) increases with the duration of stimulus availability. However, because this also applies to those trials on which the wrong level was attended, it increases the number of responses that would falsely be interpreted as conjunction errors. Given a sufficient number of cue-confusion trials, this can even lead to an increasing number of conjunction errors with increasing SMIs. On the other hand, according to our integration hypothesis, the probability that the letters are bound to their levels increases with the SMI. Consequently, we expected a steeply decreasing rate of conjunction errors with increasing SMIs. With respect to hemispheric differences, we expected the same VF effects as in Experiment 1.

Method

Sixteen students (11 women and 5 men; mean age = 24.2 years) from the Universität Konstanz participated in this experiment. All had normal or corrected-to-normal vision and were paid for their participation.

The method was identical to that of Experiment 1 except that four fixed SMIs (12, 24, 48, and 96 ms) were randomly chosen across trials. After some training, 12 blocks of 96 trials were run in a single 1.5-hr session. This resulted in 72 trials per condition.

Results

Conjunction Errors

The mean rate for conjunction errors was 21.6%. As can be seen in Figure 5a, the rate of conjunction errors decreased with an increasing SMI, $F(3, 45) = 59.0$, $p < .01$. Furthermore, as in Experiment 1, there was a significant interaction between level and VF, $F(1, 15) = 12.0$, $p < .01$ (see Figure 6). For the global target level, the mean difference between the visual fields was 4.60%, whereas it was 2.35% for the local target level. Also, the interaction between SMI and VF was reliable, $F(3, 45) = 2.83$, $p < .05$. For long SMIs, LVF stimuli produced fewer errors than RVF stimuli, whereas there was no difference for short SMIs.

Feature Errors

The mean rate of feature errors was 21.4%. Also, for feature errors the main effect of SMI was significant, $F(3, 45) = 48.2$, $p < .01$ (see Figure 5a). In contrast to the results of Experiment 1, this time the interaction between level and VF was reliable, $F(1, 15) = 6.04$, $p < .05$ (see Figure 6). For the global target level, the mean difference between the visual fields was 2.21%, whereas it was 1.63% for the local target level.

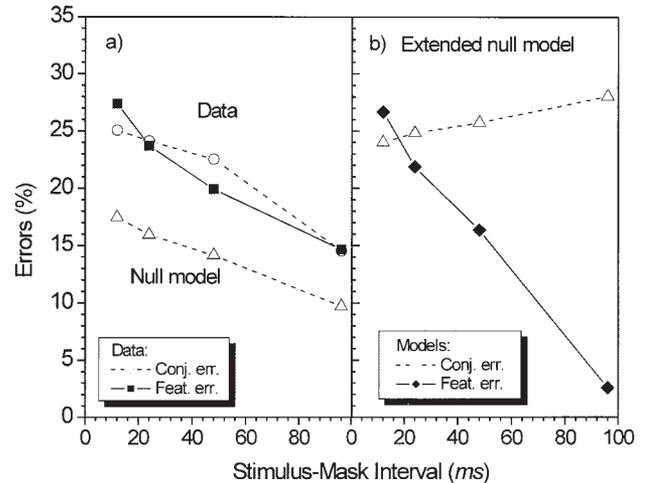


Figure 5. The results of Experiment 2 and the predictions made by different models: (a) The upper two curves show the conjunction errors (Conj. err.) and feature errors (Feat. err.). The lower dashed curve depicts the rate of conjunction errors predicted by the simple null model. (b) Conjunction and feature errors as predicted by the extended null model.

Discussion

This experiment was conducted to test whether the standard view of global/local processing can explain the relatively large number of conjunction errors if it is additionally assumed that on some trials the participants falsely attended to the wrong level because of cue confusion. On these trials they reported the letter at the nontarget level, which was falsely interpreted as a conjunction error. The test was attained by varying the length of the SMI across trials. Given that the probability of correctly identifying the letter at the attended level increases with the SMI, the number of errors that would falsely be interpreted as conjunction errors should also increase relative to the prediction of the simple null model. Depending on the number of trials on which the wrong level was attended, this can even lead to an increasing rate of conjunction errors with increasing SMIs. On the other hand, our integration hypothesis assumes that a longer period of time available for letter and level integration should largely reduce the number of conjunction errors. The data clearly show that the rate of conjunction errors decreased with increasing SMIs, which is in line with the integration hypothesis.

To further examine whether the data could also be explained by the extended null model, we fitted it to the data. We again set $n = 0$ (recall that this is to the advantage of the model) and estimated the value of parameter c (the probability of attending to the cued level) and the values of parameter t (the probability of identifying the letter at the attended level) for each of the SMIs, respectively. The estimation was accomplished by minimizing the deviation of the model from the averaged percentage of correct responses across the SMIs. As a result, we obtained $c = .73$. That is, cue confusions would have had to occur in 27% of the trials. The estimated values of the parameter t were .47, .56, .67, and .95 for the individual SMIs, respectively. The corresponding predicted rates of conjunction and feature errors are shown in Figure 5b. As can be seen, the large number of necessary cue confusions, to-

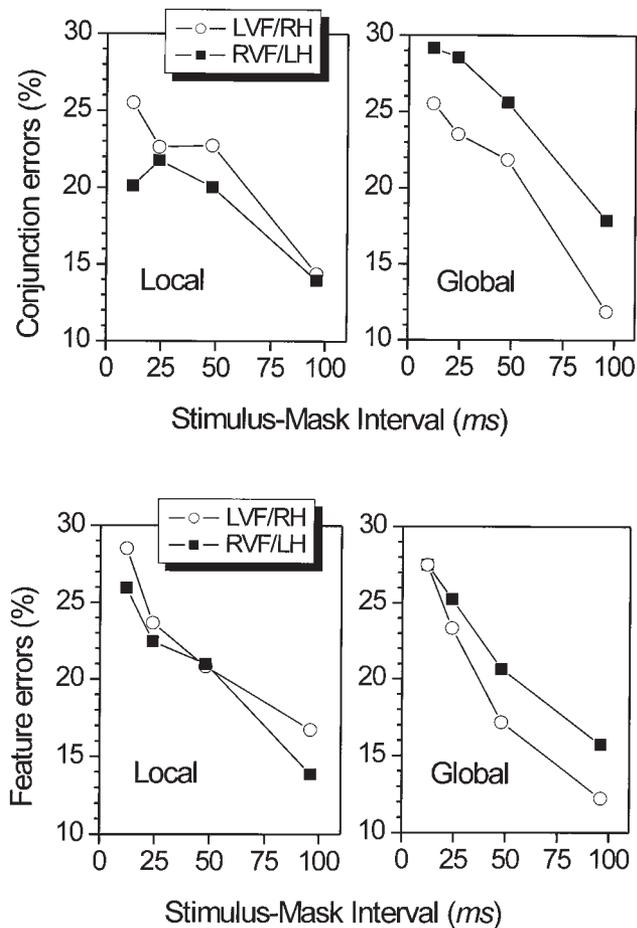


Figure 6. The interaction between level and visual field for the feature and conjunction errors in Experiment 2. LVF = left visual field; RVF = right visual field; RH = right hemisphere; LH = left hemisphere.

gether with the increasing probability of letter identification, leads to an increase of conjunction errors by 4% across the SMIs. In contrast to this prediction, these errors decreased by 10% in the data.

If we would reduce the number of assumed cue confusions so that their rate decreases with the SMI, then their number would largely be underestimated, especially at short SMIs. At the extreme, with $c = 1$, we would end up at the simple null model. For comparison, we included its predicted number of conjunction errors in Figure 5a. As can be seen, the simple null model largely underestimates these errors. Thus, as these analyses show, there is no reasonable way for the null model or for the extended null model to explain our data.

With respect to hemispheric differences, VF effects occurred in the expected direction for the conjunction errors. If we consider the feature errors, then their rate was increased, compared with the previous experiment. This might have contributed to the result that this time reliable VF effects also occurred for this error type. However, the rates were still smaller than those for the conjunction errors. Interestingly enough, the VF effects for both error types did not vary significantly with the SMI (see Figure 6).

Altogether, the two reported experiments clearly show that the null model cannot explain the data. This suggests that on some trials the identities of the stimulus letters were already known before the information about their respective levels was available and that, therefore, letters and levels have to be integrated at a later stage. Because the assumption of an integration stage leads to a more complex model of global/local processing than the standard view, it would also be desirable to have a multinomial model of this account. Such a model would not only help to interpret our pattern of results, it would also offer the possibility of fitting the model to the data and of estimating the parameters.

A multinomial model of the integration hypothesis needs, in addition to the parameters t and n for representing the probability that the letters at the target and nontarget levels, respectively, are identified, two further parameters. One parameter, α , is required for representing the probability that the target letter is correctly bound to its level. The other parameter, β , represents the corresponding probability for the letter at the nontarget level. However, adding two parameters leads to a problem for fitting the model to our data. Because we have only three response categories for a given SMI and visual field, the relation between the number of data points and the number of parameters is inappropriate. Therefore, we had to conduct a further experiment in which the number of response categories was increased.

Experiment 3

The participants in this experiment had, as before, to identify the letter at a prespecified target level. In addition, however, a second response was required in which they also had to identify the letter at the nontarget level. Because we could now observe two responses on each trial, the number of response categories was increased. We symbolize the two responses on a given trial by presenting the corresponding pair of single response categories. For instance, TN denotes the case in which the letter at the target level is reported first and the letter at the nontarget level is reported second; that is, both responses are correct. Accordingly, NT represents a double conjunction error. Actually, all 12 possible pairs of the four different letters could occur. Letter repetitions were not considered, because the letters within a given stimulus were always different and the participants knew that.

Thus, we have response pairs such as TN, NT, TO, T \emptyset , and so forth. However, because TO and T \emptyset have the same a priori probability, they were collapsed into one response category. By applying this principle to all possible response pairs, we obtained seven different response categories: {TN}, {TO, T \emptyset }, {NT}, {NO, N \emptyset }, {OT, \emptyset T}, {ON, \emptyset N}, and {O \emptyset , \emptyset O}. Each category is denoted by its first member. For instance, the last category is represented by O \emptyset . For these different response categories, a multinomial model was constructed that represents our integration hypothesis. The result can be seen in Figure 7.

The relative frequencies for the different response categories were registered not only for each VF and target level but also for the different SMIs. This allowed us to examine how the parameters change with the duration of stimulus processing. Because in the previous experiment conjunction errors were relatively numerous even at an SMI of 96 ms, here we included an additional SMI with a duration of 192 ms.

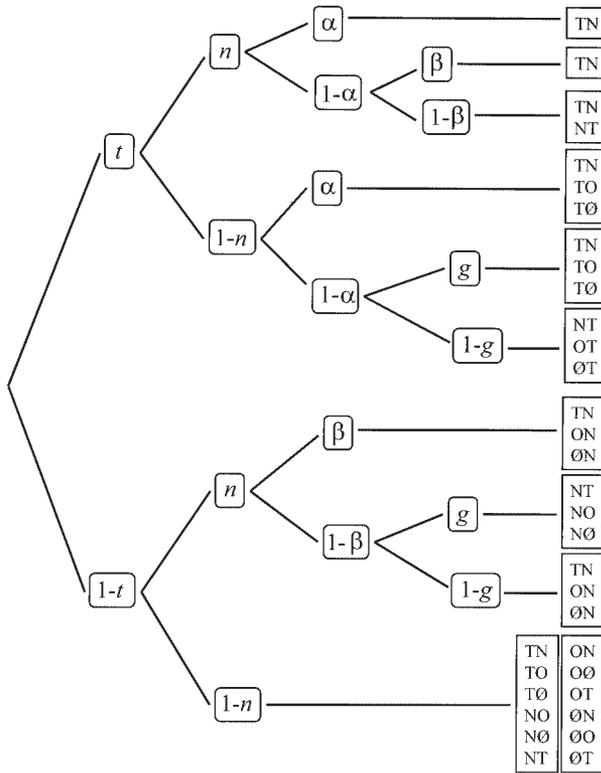


Figure 7. The multinomial tree diagram of the integration model.

Although we do not discuss our model in detail before the experiment and its results are reported, we consider one aspect with regard to hemispheric asymmetries at this point. As can be seen in Figure 7, the identification and the binding parameters are related in a multiplicative way. Consequently, we cannot expect to answer the question of whether identification or binding capacity differs between the cerebral hemispheres simply by fitting the model to our data. Either difference can be expected to explain the data. Therefore, we needed independent evidence for our hypothesis that the hemispheres do not differ with respect to their identification capacity. This should be achieved by including neutral stimuli, that is, stimuli that have a letter only at the target level and that have a neutral symbol at the nontarget level. If only one letter is present, binding should not be necessary for producing a correct first response. Therefore, if the hemispheres differ only with respect to their binding capacity, no VF effects should occur for neutral stimuli. On the other hand, if the hemispheres differ also or exclusively in their identification capacity for global and local letters, then VF effects can be expected.

Method

Sixteen students (9 women and 7 men; mean age = 22.5 years) from the Universität Konstanz participated in this experiment. All had normal or corrected-to-normal vision and were paid for their participation.

The method was similar to that in Experiment 2 except that an additional SMI length was used. Thus, there were five different SMIs (12, 24, 48, 96, and 192 ms). Furthermore, neutral stimuli were also presented. For these stimuli, either a *U* or an inverted *U* occurred at the nontarget level. The

main difference from the previous experiment, however, was that the participants, after reporting the letter at the target level (first response), also had to indicate the letter at the nontarget level (second response). Because this was possible only for nonneutral stimuli, an additional button with the meaning “I’m not sure” could be pressed for the second response on a trial. Moreover, auditory error feedback was given only for the first response on a trial.

Altogether, there were 1,800 trials for each participant, which were distributed across three 1-hr sessions. This resulted in 90 trials per condition per participant.

Results

Before performance with respect to the response categories is reported, a standard data analysis is provided for the first and second responses.

Target Identification (First Response)

Conjunction errors. The mean rate of conjunction errors was 11.3%. There was a significant main effect of SMI, $F(4, 60) = 54.3, p < .01$, reflecting that the rate of conjunction errors decreased from 16.1% to 6.19%. However, the two-way interaction between level and SMI was also significant, $F(4, 60) = 6.02, p < .01$, indicating that more errors occurred for the global than for the local target level at short SMIs, whereas the opposite held for long SMIs. The interaction between level and VF was also significant, $F(1, 15) = 7.93, p < .05$. However, there was also a significant three-way interaction among level, VF, and SMI, $F(4, 60) = 4.02, p < .01$. As can be seen in Figure 8, this interaction indicates that the two-way interaction between level and VF is absent at long SMIs.

Feature errors. The average rate of feature errors was 10.7%. As was the case for conjunction errors, there was a significant main effect of SMI, $F(4, 60) = 66.7, p < .01$. The feature error rate decreased from 18.3% to 5.03%. Further, the two-way interaction between level and SMI was reliable, $F(4, 60) = 3.45, p < .05$. As was true for conjunction errors, more feature errors occurred for the global than for the local target level, but only for short SMIs. Finally, there was a significant three-way interaction among level, VF, and SMI, $F(4, 60) = 1.92, p = .12$. As can be seen in Figure 8, the usual two-way interaction between level and VF was present only at an SMI of 96 ms.

Neutral. The average error rate for neutral stimuli was 20.7%. There were significant main effects for SMI, $F(4, 60) = 39.1, p < .01$, and for level, $F(1, 15) = 6.17, p < .05$. However, there was also a two-way interaction between these two factors, $F(4, 60) = 10.7, p < .01$. As was also true for conjunction and feature errors, this interaction reflects the fact that the local level has an advantage, but only at short SMIs. The error rate decreased from 38.4% to 9.61% for the global level, and from 22.8% to 13.1% for the local level. Most important, however, the interaction between level and VF was far from significant, $F(1, 15) = 0.032, p = .86$. This was also true for the three-way interaction among level, VF, and SMI, $F(4, 60) = 0.939, p = .45$.

Nontarget Identification (Second Response)

Conjunction errors. On average, the rate for conjunction errors was 17.0%. Only the main effect of SMI was significant, $F(4,$

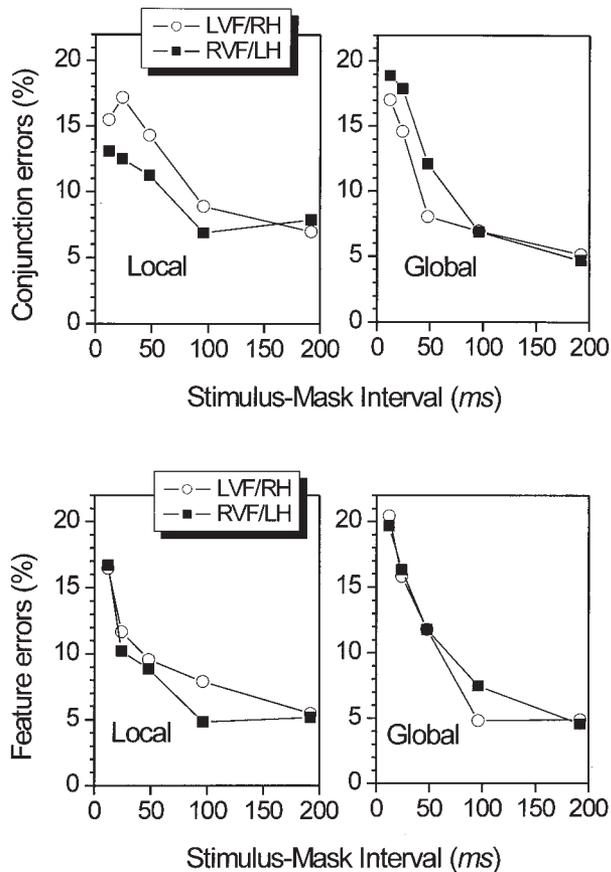


Figure 8. The identification performance for the letter at the target level in Experiment 3. LVF = left visual field; RVF = right visual field; RH = right hemisphere; LH = left hemisphere.

60) = 15.1, $p < .01$. The error rate decreased from 23.9% to 9.28% (see Figure 9).

Feature errors. On average, there were 32.1% feature errors. The main effect of SMI was significant, $F(4, 60) = 48.7$, $p < .01$ (a decrease from 41.9% to 19.3%), as was the main effect of level, $F(1, 15) = 4.58$, $p < .05$. However, there was a significant two-way interaction between level and SMI, $F(4, 60) = 3.04$, $p < .05$, indicating that there were more errors for the global than for the local target level, especially at short SMIs. Furthermore, the two-way interaction between level and VF was significant, $F(1, 15) = 11.7$, $p < .01$. However, there was also a three-way interaction among SMI, level and VF, $F(4, 60) = 2.73$, $p < .05$. As can be seen in Figure 9, the typical VF effects were largest for the medium range of SMIs.

Response Categories (Paired Responses)

The observed letter pairs were classified into the seven different response categories. Because there were only a few letter repetitions (1.42%), they were ignored. Presumably, they reflect error corrections.

The relative frequencies of the response categories for the different conditions are listed in Appendix A. Furthermore, the

data for four response categories are also shown in Figure 10. These data are discussed in connection with the model fit.

Discussion

This experiment was primarily conducted to collect an appropriate number of response categories in order to fit a multinomial model of our integration hypothesis to the data. This objective was achieved by requiring not only the identification of the letter at the target level but also the identification of the letter at the nontarget level. However, the registration of two responses on each trial also produced interesting results in itself. Although the task as a whole was different from and presumably required different mental processes than the task in the previous experiment, the results with respect to the identification of the letter at the target level largely replicated those of Experiment 2. There were again VF effects. For the conjunction errors, however, VF effects were absent at the longest SMI of 192 ms. For the feature errors, VF effects were present only at a SMI of 96 ms. On average, the VF effects were smaller for feature errors than for conjunction errors.

The performance with respect to reporting the letter at the nontarget level was different. First of all, performance was reduced compared with first-response performance. This finding indicates that attentional

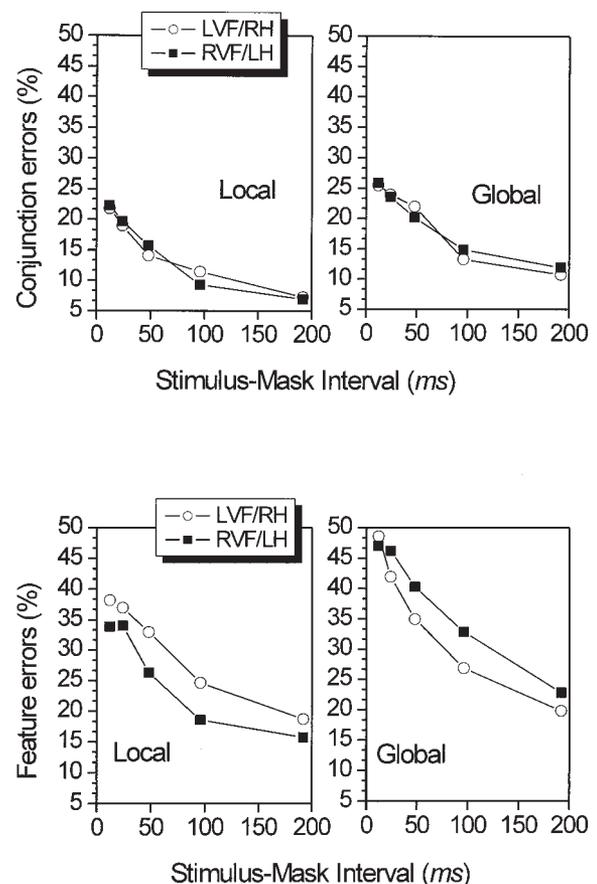


Figure 9. The identification performance for the letter at the nontarget level in Experiment 3. LVF = left visual field; RVF = right visual field; RH = right hemisphere; LH = left hemisphere.

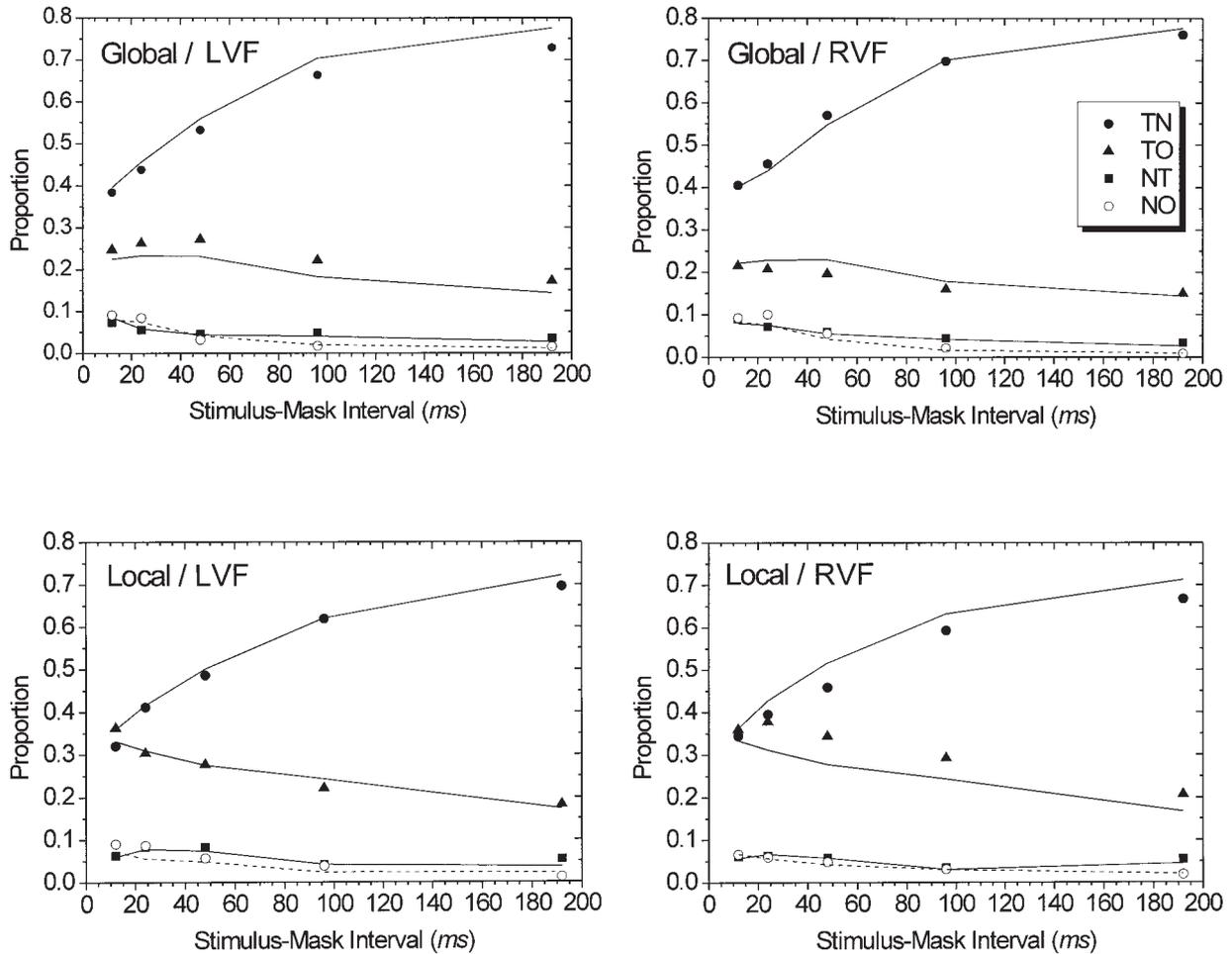


Figure 10. The proportion of four out of seven response categories (symbols) and the corresponding performance predicted by our integration model (lines). T = letter at the target level; N = letter at the nontarget level; O = letter not present in the display. The first letter of the code for the response categories indicates the response to the target level, which was required first. The second letter indicates the response to the nontarget level, which was required second. For instance, TN means that both responses are correct, whereas NT means both responses are conjunction errors. LVF = left visual field; RVF = right visual field.

filtering was involved and that response selection had a serial component. Obviously, the participants shifted their attention from the target level to the nontarget level. This explains why the identification of the letter at the nontarget level was less successful than identification of the letter at the target level, especially for short SMIs. Moreover, as can be seen in Figure 9, refocusing attention was more difficult going from the local to the global level than vice versa. In other words, zooming in was easier than zooming out, a phenomenon already observed by Vorberg (1992).

Furthermore, for the second response, there were no VF effects for conjunction errors, but there were substantial VF effects for feature errors. However, these effects are presumably due to the dependency between the responses. Thus, it seems that this complicated pattern of results can only be interpreted by means of a detailed model. As mentioned previously, our multinomial model cannot be used to decide whether the hemispheres differ in their capacity for letter and level binding or in their capacity for letter

identification. Therefore, neutral stimuli (with a letter only at the target level and a task-irrelevant neutral symbol at the nontarget level) were included in the experiment. Because identification, but not integration, is necessary for these stimuli to produce a correct response, we hoped that the corresponding results would help to answer the question of how the hemispheres differ. Fortunately, the results were clear-cut. There was not even a tendency of VF effects for neutral stimuli. This finding supports our hypothesis that the hemispheres differ with respect to their capacity for binding the letters to their respective levels. With respect to the identification performance for global and local letters, there seem to be no hemispheric differences.

A Multinomial Integration Model

The number of response categories now allows us to fit the multinomial version of our integration hypothesis, as shown in

Figure 7, to the data. The parameters t and n represent, respectively, the probability that the identity of the letter at the target level is seen and the probability that the identity of the letter at the nontarget level is seen. In this respect the model is similar to the former model of the standard view. Here, however, we have two additional parameters, α and β , representing the probability that the identity of the letter at the target and the nontarget levels, respectively, is correctly bound to its stimulus level. If we consider Figure 7, then, the top branch of the tree diagram shows the case in which both letters are perceived and integrated correctly. This event occurs with probability $t \cdot n \cdot \alpha \cdot \beta$ and leads to a correct report of both letters (TN). However, if both letters are seen, it is sufficient that only one is correctly bound to its level to produce a correct order. This is because the unbound letter can only belong to the other level. For instance, with probability $t \cdot n \cdot \alpha$, both letters are seen and the target letter is bound to its level. This produces a correct response pair irrespective of whether the letter at the nontarget level is bound. Analogously, a correct response pair is also reported when both letters are seen but only the nontarget letter is bound to its level. This event occurs with probability $t \cdot n \cdot (1 - \alpha) \cdot \beta$. If the two letters of the stimulus are seen but both remain unbound (the third branch from the top), then the participants have to guess the letter order, that is, they have to choose randomly between TN and NT. This situation occurs with probability $t \cdot n \cdot (1 - \alpha) \cdot (1 - \beta)$.

It might also happen that only one letter is perceived and that it is bound correctly. This letter is then reported correctly, whereas the letter for the other level has to be chosen from among the remaining three letters. If only one letter is perceived and remains unbound, one could assume that observers randomly choose one letter pair from the set of all possible pairs that contain this letter. However, it seems more reasonable to hypothesize that in this case the perceived letter is reported first and the second letter is guessed. In other words, if early filtering mechanisms are tuned to a certain level and a single letter is seen, then it makes sense to assume that it belongs to the attended level. In order to include this strategy, we added a guessing parameter g to the model (see Figure 7).

The tree diagram in Figure 7 shows all possible events. To obtain the expected proportion for a certain response category, one simply has to add the probabilities of the different events that contain the corresponding response pairs. To provide an example, consider the response category NO. Its predicted proportion is given by adding the probabilities of Branches 8 and 10, that is, $(1 - t) \cdot n \cdot (1 - \beta) \cdot g \cdot (2/3) + (1 - t) \cdot (1 - n) \cdot (2/12)$.

Before we fit the model to our data, we consider the null model for our double-identification paradigm. It is a special case of our integration model. One simply has to assume that $\alpha = \beta = 1$. According to this model, conjunction errors can occur only when none of the stimulus letters is seen. We can also derive testable predictions from the null model. For instance, it predicts that the two types of single conjunction errors, OT and NO, have the same probability: $(1 - t) \cdot (1 - n) \cdot (2/12)$. However, this does not correspond to our data, as can be seen in Appendix A. If we consider, for example, the proportions for the global level and the LVF, then it is obvious that NO responses were about twice as frequent as OT responses (.0481 vs. .0264), $t(79) = 3.82$, $p < .001$.

To provide a further example, consider double conjunction errors. The null model predicts that they occur with only half of the probability just considered, that is, with probability $(1 - t) \cdot (1 - n) \cdot (1/12)$. However, this also does not correspond to our data (see Appendix A). A test comparing NT with NO responses revealed that there was no main effect but that there was an interaction with SMI, $F(4, 60) = 5.59$, $p < .001$. At short SMIs, NT responses occurred less frequently than NO responses (7.21% vs. 8.53%). From the third SMI onward, however, NT responses were given more frequently. For the longest SMI, the proportions are 4.47% versus 1.38%. The time course for these response categories can also be seen in Figure 10.

As these examples show, there are again many reasons to reject the null model. Therefore, we now examine how successfully our integration model explains the data. The model shown in Figure 7 was fitted to the data in order to estimate the different parameters. If it is assumed that t , n , α , and β differ between the target levels, VFs, and SMIs, we have four parameters for each of the seven corresponding data points (response categories). However, because the data for the neutral stimuli and other results (Hübner & Malinowski, 2002) support our hypothesis that the hemispheres do not differ with respect to their identification capacities, the number of parameters can be reduced by using the common parameters t and n for a given level. Only the individual binding parameters α and β were allowed to vary between the VFs. Thus, we had to estimate two identification parameters (for a given target level) and four binding parameters (two for each VF) in a single step from the corresponding 14 data points for both VFs. The estimation procedure was then run individually for each target level, SMI, and participant.

Because it is reasonable to assume that the guessing parameter g represents a global strategy, it was fixed across all conditions and participants. After trying three different values for g , we set it to .9. The other parameters were estimated by minimizing the goodness-of-fit measure G^2 , which is defined as (cf. Prinzmetal et al., 2002)

$$G^2 = \sum_{i=1}^m \left[(2 \cdot f_i^{ob}) \ln \left(\frac{p_i^{ob}}{p_i^{pr}} \right) \right].$$

In this formula, m denotes the number of response categories, f_i^{ob} the observed frequency, p_i^{ob} the observed proportion, and p_i^{pr} the predicted proportion of response category i . The smaller the value of G^2 , the better the fit. The function “fminsearch” from MATLAB served as the algorithm for the minimization. To ensure that the fits did not represent local minima, we used four different starting values for each parameter.

The obtained mean parameters and corresponding goodness-of-fit values, G^2 and the sum of squared errors (SSE), are listed in Appendix B. The corresponding predicted values for all seven response categories are given in Appendix C. A sample of these values is shown in Figure 10 together with the corresponding empirical data. For the figure we chose the two response categories with the largest proportion and the two response categories that contained conjunction errors for the target level. As can be seen, the fits are close to the observed data. Only those for the local target level and RVF stimuli deviate to some extent from the data.

If we consider, as for the null model, the predicted values for NT and NO responses (see Figure 10), then it turns out that they are similar to the observed data. There is no significant main effect for

the response categories, but there is a significant interaction with the SMI, $F(4, 60) = 3.76$, $p < .01$. At short SMIs, fewer NT responses are predicted than NO responses (at the shortest SMI, 7.10% vs. 7.56%). At longer SMIs, however, NT responses are predicted to occur more often than NO responses (at the longest SMI, 3.41% vs. 1.63%). This demonstrates that even such relatively small effects are mimicked by the model.

To compare the goodness of fit of the integration model with that of the null model, the latter was also fitted to the data. With its four parameters (t and n for each VF), we obtained for each of the 14 data points an average G^2 of 125 and an average SSE of .0553. With six instead of four parameters (common t and n for both VFs, one α and β for each VF), the integration model reached an average G^2 of 60.3 and an average SSE of .0252. Thus, the two additional parameters produced not only a qualitative but also a substantial quantitative improvement of the fit.

The comparison between the models clearly favors the integration model. The null model cannot describe the data. This supports our hypothesis that letter identification and binding are two separate and serial stages.

By considering the obtained parameters (see Appendix B), it can be seen that the identification parameters t and n increase, as expected, monotonically with the SMI. The values of t also mirror the result that the local units had an advantage at short SMIs, whereas global units were superior at long SMIs (compare the size of t in the top half of Table B1 with that in the bottom half of Table B1). The values of n are generally smaller than those of t . This reflects early attentional filtering and suggests that the participants first focused their attention to the target level and then shifted it to the nontarget level. One could have expected the difference between t and n to decrease with the SMI because of the increasing amount of time available for selecting information from the nontarget level. However, this seems to have been the case only for the local target level. Furthermore, the difference between t and n is generally larger for the local target level. Thus, the relations between these parameters reflect our result that it was more difficult to shift attention from the local to the global level than vice versa.

On the basis of our results for the neutral stimuli, we assumed that the hemispheres do not differ with respect to letter identification. Thus, t and n were held constant for both VFs. Consequently, VF effects could be reflected only by the binding parameters. If we consider the values in the bottom half of Table B1, then it is obvious that, as expected, the binding parameter α for the global target level is generally greater for LVF stimuli than for RVF stimuli. The opposite should hold for the local target level, which was the case for the three shortest SMIs (see Appendix C, the top half of Table C1).

In contrast to the identification parameters, the values of the binding parameter for the target level were not generally larger than those for the nontarget level. This suggests that binding was not affected by attentional focusing. Furthermore, the binding parameters did not systematically increase with the SMI. Although α increased to some extent for the global target level (see the top half of Table B1), we had expected a more general increase with the SMI. This result could indicate that binding is a relatively slow and fragile process such that even 200 ms are not sufficient for reaching a stable stimulus representation. However, it could also indicate that the mask was crucial and destroyed the result of the

binding process to some extent irrespective of the SMIs. Moreover, the mask could even have a specific effect for each of the SMIs. That is, depending on the state and stage of processing, it could have a more or less negative effect. This could explain the variability of the binding parameters across the SMIs.

Thus, although the obtained parameters were not as expected in all respects, by and large they are compatible with our general ideas about global/local processing.

General Discussion

Our aim in the present article was to examine the mechanisms and structures involved in the processing of hierarchical stimuli. Although this is an intensively investigated area, many questions are still open. For instance, little is known about how the units at a given stimulus level are selected and represented. Another question concerns the involved brain areas. In what sense are the left and right cerebral hemispheres specialized for processing local and global units, respectively? If one considers the relevant literature with respect to these questions, then one learns that the results are inconclusive. One reason for the inconsistency might be the implicit assumption in many studies that the involved processes are relatively simple and robust. As a consequence, task parameters vary considerably across studies, and little attention has usually been paid to their possible effects.

Here, we have argued that even in relatively simple global/local tasks, the involved perceptual and cognitive processes are relatively complex and susceptible to minor experimental changes. Therefore, because one can usually measure only the end product of the whole processing, all components of the task and their possible mental representations should be taken into account. This could finally lead to more complete models for processing from the stimulus to the response. In the present article, we adopted the idea of evidence accumulation (Miller, 1981) as a theoretical framework for describing the proposed mental structures and processes. In a first approach, we assumed that there are mental representations for the identities of the different stimulus units and for the required responses, and that the relative activations of these representations serve for response selection. For ambiguous stimulus conditions, we hypothesized that attentional mechanisms adjust perceptual filters in such a way that the processing of the different units is biased in favor of those at the target level. If this filtering is effective, then the identity of the units is obviously sufficient for a reliable response selection.

The Necessity of Binding Level and Its Content

The important question with respect to the objective of the present study is how the system proceeds when early filtering is insufficient. It seems inevitable that in this case, level information has to be taken into account as well. However, how this information is represented in the mental system is an open question. There are two main alternative answers. One possibility is that level and identity information are coded together right from the beginning. This hypothesis was called the standard view, because we think that this view is more or less explicitly shared by most of the researchers in this field. The other possibility, proposed and favored by us, is that level and identity information are coded separately at early stages. It follows from the latter hypothesis that

identity and level information have to be integrated at some later stage in order to obtain a complete stimulus representation.

Many authors in this area have not been precise about the mechanisms underlying their theoretical views, and therefore it is unclear whether they would really endorse the standard view as presented here. Unquestionably, many researchers have remained rather vague about their assumed mechanisms. Usually, they have investigated specific problems such as the global advantage or the properties of early filtering mechanisms, which can be examined with minimal theoretical assumptions. However, researchers who went further and proposed more detailed models usually ended up proposing that the levels correspond to specific mechanisms in the brain that process their content or features. This is most obvious in theories that relate global/local processing to transient/sustained channels (e.g., May et al., 1995), to low/high spatial-frequency channels (e.g., Broadbent, 1977; Shulman et al., 1986), or to magnocellular/parvocellular pathways (e.g., Michimata et al., 1999). Corresponding ideas can also be found in the area of hemispheric asymmetries for global/local processing (cf. Ivry & Robertson, 1998).

In all of these theories, each level is processed by an individual mechanism, which implies that content is automatically related to its level. The predominance of this view can also be seen by considering one of the most detailed models in this area. It was proposed by Lynn Robertson (1996), a leading theorist in this field. In her flowchart (see Figure 4 in Robertson, 1996), global and local information are separated from the beginning and processed in independent spatial-frequency channels up to an identification stage.

Another leading researcher in this field, Marvin Lamb, proposed together with his co-worker E. Yund a similar idea. They merely substituted level-specific neural mechanisms (Lamb & Yund, 1996) for the spatial-frequency channels in Robertson's (1996) model. It is interesting that they also briefly considered the possibility that target identity is represented separately from target level. After mentioning that Biederman and Cooper (1992) provided evidence that the mechanisms for representing object shape are different from those for representing object size, Lamb and Yund (1996) hypothesized that "it is possible that target identity and target level are represented separately" (p. 372). However, they instantly rejected this idea and favored their level-specific mechanism hypothesis, because it "escapes this difficulty altogether" (p. 372).

These considerations show that the standard view is more than a straw man for the present article and in fact represents a core assumption of at least the majority, if not all, of the current theories of global/local processing. On the other hand, the idea that levels and their contents are processed independently at early stages and, consequently, have to be integrated at a later stage has never before been considered a serious alternative. Only when we had difficulties explaining our VF effects with the standard view did we come up with the hypothesis that the identity of the units might be available before the knowledge about their level (Hübner & Malinowski, 2002).

To provide evidence for our integration hypothesis, we conducted three experiments with compound letters for which we adapted an experimental paradigm from the research area of illusory conjunctions (cf. Prinzmetal et al., 2002). The task was to identify the letter at a prespecified target level. However, the

duration of stimulus processing was restricted by a mask. We hoped that on some trials this would provide sufficient time for letter identification but not for binding identity and level information. In these cases, letter identities and levels were expected to remain unbound and to lead to a relatively high rate of conjunction errors, that is, of false reports of the letter at the nontarget level.

Our integration hypothesis was contrasted with the standard view of global/local processing, that is, with the assumption that units and levels are coded in combination. According to this view, knowing the identity of a stimulus letter implies that its level is also known. Therefore, only a few conjunction errors should occur. In order to obtain a quantitative prediction, we formalized the standard view as a multinomial null model and calculated the rate of expected conjunction errors. We then tested this prediction in our first experiment. The results showed that there were considerably more conjunction errors than predicted, which supported our integration model.

In a second experiment, we varied the SMI in order to exclude the alternative account that the conjunction errors in Experiment 1 were simply due to the participants confusing the cues and, consequently, attending to the wrong level. If this had been the case, an improved identification rate with increasing SMIs should also have increased the number of conjunction errors. However, this was not the case. Rather, the number of conjunction errors decreased with the length of the SMI. Thus, together, the results of the first two experiments supported our hypothesis that letters and levels are coded separately at early stages.

In order to fit the integration hypothesis to data, we translated it into a multinomial model that differed from the null model by two additional parameters. Thus, there were not only two parameters for representing the identity of the letter at the target level and the nontarget level, respectively, but also two parameters for representing their binding state. Because the number of data points collected in the first two experiments were not sufficient for applying the model, we conducted a third experiment in which the participants had to identify not only the letter at the target level but also the letter at the nontarget level. Relative to the first two experiments, this third experiment increased the number of observable response categories from three to seven.

The resulting parameter values from fitting the model to the data were in line with the observed results and largely reflected our hypotheses. For instance, the values of the identification parameters for the letter at the target level were generally larger than those for the letter at the nontarget level. This reflects the result that the performance for the target level was better than that for the nontarget level and indicates that attention was biased in favor of the target level at the beginning of a trial. In other words, the participants did not distribute their attention equally across the levels but first focused their attention on the target level and then shifted it to the nontarget level.

As expected, the identification parameters also increased in size with the duration of stimulus processing. However, unexpectedly, this was not the case for the binding parameters. Rather, their values varied relatively unsystematically across the SMIs. This finding could indicate that the negative effect of the mask for the binding process was different from that for the identification process. It seems that the mask could destroy the result of the binding process even at long SMIs. Moreover, the data suggest that the disrupting effect of the mask was particularly strong at some

intermediate intervals. Presumably, these intervals correspond to sensible stages of processing.

Thus, although not all of the obtained parameter values varied as expected, by and large the multinomial model reflects our main ideas rather well.

Abstract Representations of Level and Content

Altogether, the data and the model fit provide strong evidence for our integration hypothesis, that is, for the idea that the identities of the units of a hierarchical stimulus are coded separately from the information about their levels at early stages of processing. This demonstrates that questions about how stimuli are represented and how these representations are used for executing the required task are important for explaining performance. Our results indicate that the early stimulus representations are relatively abstract. This explains why it was obviously so difficult for our participants to decide whether a letter occurred at one or the other level even though global and local letters differed in various respects such as size and number. It can be assumed that at early stages, letter identity was represented as an isolated property. This assumption is compatible with the feature integration theory (e.g., Treisman, 1998), which states that all that is available before binding takes place is evidence about the presence or absence of features, not evidence about their individual instantiations.

If we formulate these ideas within the framework of evidence accumulation, we can say that there is one evidence counter for each relevant letter identity. The process that registers the evidence for the presence or absence of a given letter is not necessarily aware of the origin of this evidence, that is, of the sources that activated the respective evidence counter. Thus, at early stages, the system might have evidence about the presence of a certain letter in the display, but it knows neither its number and size nor the level at which it occurred. All of this additional knowledge depends on integration processes at later stages. The representation of levels can be conceptualized in a similar manner. They might also be represented rather abstractly at early stages. As for letters, it can be assumed that evidence counters at early stages merely signal the presence or absence of a level for a given stimulus.

Hemispheric Differences

An important part of our integration hypothesis is the assumption that the hemispheres are equivalent with respect to letter identification but differ in their capacities for integrating identity and level information. This idea was originally developed in a former article (Hübner & Malinowski, 2002) and is in line with the older notion that functional hemispheric asymmetries reflect differences at higher stages rather than at lower ones (cf. Moscovitch, 1979). The results obtained in the present experiments further support this hypothesis. In all three experiments, there were VF effects in the expected direction for conjunction errors. Similar but smaller effects also showed up for feature errors, at least in the last two experiments. Because the different error types are not independent of each other, corresponding VF effects alone are not suited for deciding how the hemispheres differ. Unfortunately, the multinomial model did not resolve this issue either. Because the identity and binding parameters are combined in a multiplicative way, the model cannot be used to decide whether the hemispheres

differ with respect to identification, binding, or both. Therefore, in order to obtain other information for answering this question, we also included neutral stimuli in Experiment 3. Because these stimuli had a letter only at the target level and a neutral symbol at the nontarget level, the response could in principle be selected by using identity information alone, without any binding. Thus, according to the integration hypothesis, there should be no hemispheric differences for neutral stimuli. This was precisely the case. On the basis of this result, the number of parameters in the model was reduced by using identical identification parameters for both VF conditions. Only the binding parameters were allowed to differ. As a result, the parameter values obtained by fitting the model to the data differed between the two VF conditions in the expected direction.

Altogether, our results support the hypothesis that the hemispheres differ in their binding capacities rather than in their identification capacities. A recent patient study provides further evidence for this conjecture. Doricchi and Incoccia (1998) examined a patient with damage to the right hemisphere of the brain. She had no problems seeing the global shape as long as there were no relevant local shapes. Thus, her identification performance was still intact, but her binding capacity seemed to be reduced. However, one might argue that the hemispheres differ in their identification capacities but only if there is competition between the levels. Although we cannot definitively rule out this account, given our pattern of results it is obvious that the integration hypothesis is not only more precise and elegant but also more parsimonious.

Generality of the Results

An interesting question with respect to the present results is whether the stimuli used are special. The compound-letter stimuli used here and in most other global/local studies belong, according to Pomerantz (1981), to the class of *Type P* objects. For these objects, only the position of the local units matters for the identity of the global unit. Moreover, all of the local units are identical. In natural environments, hierarchical objects are usually *Type N* objects, that is, they consist of different local elements whose position as well as nature matter for the identity of the object (cf. Pomerantz, 1981). However, *Type P* stimuli have generally been preferred for experimental studies, because the identity of one unit can be changed without seriously affecting the identity of the other unit. Moreover, the same identity can be used for local as well as for global units.

It is reasonable to assume that the properties of *Type P* stimuli were helpful for obtaining our results. Also, the use of well-learned letters as stimuli was presumably advantageous. It is likely that the letters automatically activated associated mental representations irrespective of their level. This clearly favored the separate coding of units and levels. Thus, one might wonder whether the present results can be generalized to *Type N* or more natural stimuli. We think that the binding of levels and content is also necessary for other stimuli. However, it might be more difficult to observe, because semantic knowledge about objects can usually be used to resolve ambiguities. Assume, for instance, that a picture of a human body is presented very briefly and then masked. If one sees a hand, then, even if it remains unbound, one simply knows that it is a local part of the body. Thus, although *Type P* stimuli are

special, they allow one to uncover a common phenomenon, which can hardly be observed with natural stimuli.

Finally, given our method and the close analogies between our results and conclusions and those found for other binding phenomena, one might wonder whether there were also illusory conjunctions between levels and letters. Although we do not exclude this possibility, because our model fitted the data sufficiently well without assuming illusory conjunctions, we had no reason to include such a hypothesis at this point.

References

- Allport, A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance XV* (pp. 421–452). Cambridge, MA: MIT Press.
- Ashby, F. G., Prinzmetal, W., Ivry, R., & Maddox, W. T. (1996). A formal theory of feature binding in object perception. *Psychological Review*, *103*, 165–192.
- Batchelder, W. H., & Riefer, D. M. (1990). Multinomial processing models of source monitoring. *Psychological Review*, *97*, 548–564.
- Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 121–133.
- Blanca-Mena, M. J. (1992). Can certain stimulus characteristics influence the hemispheric differences in global and local processing? *Acta Psychologica*, *79*, 201–217.
- Boer, L. C., & Keuss, P. J. (1982). Global precedence as a postperceptual effect: An analysis of speed–accuracy tradeoff functions. *Perception & Psychophysics*, *31*, 358–366.
- Boles, D. B., & Karner, T. A. (1996). Hemispheric differences in global versus local processing: Still unclear. *Brain and Cognition*, *30*, 232–243.
- Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, *15*, 1411–1412.
- Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.
- Broadbent, D. E. (1977). The hidden preattentive processes. *American Psychologist*, *32*, 109–118.
- Delis, D. C., Robertson, L. C., & Efron, R. (1986). Hemispheric specialization of memory for visual hierarchical stimuli. *Neuropsychologia*, *24*, 205–214.
- Doricchi, F., & Incoccia, C. (1998). Seeing only the right half of the forest but cutting down all the trees? *Nature*, *394*, 75–78.
- Eriksen, C. W., & St. James, J. D. (1986). Visual attention within and around the field of focal attention: A zoom lens model. *Perception & Psychophysics*, *40*, 225–240.
- Evert, D. L., & Kmen, M. (2003). Hemispheric asymmetries for global and local processing as a function of stimulus exposure duration. *Brain and Cognition*, *51*, 115–142.
- Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J., & Dolan, R. J. (1996). Where in the brain does visual attention select the forest and the trees? *Nature*, *382*, 626–628.
- Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., & Frackowiak, R. (1997). Hemispheric specialization for global and local processing: The effect of stimulus category. *Proceedings of the Royal Society of London (B)*, *264*, 487–494.
- Grice, G. R., Canham, L., & Boroughs, J. M. (1983). Forest before trees? It depends where you look. *Perception & Psychophysics*, *33*, 121–128.
- Han, S., Fan, S., Chen, L., & Zhuo, Y. (1997). On the different processing of wholes and parts: A psychophysiological analysis. *Journal of Cognitive Neuroscience*, *9*, 687–698.
- Han, S., He, X., & Woods, D. L. (2000). Hierarchical processing and level-repetition effect as indexed by early brain potentials. *Psychophysiology*, *37*, 817–830.
- Han, S., He, X., Yund, E. W., & Woods, D. L. (2001). Attentional selection in the processing of hierarchical patterns: An ERP study. *Biological Psychology*, *56*, 113–130.
- Heinze, H. J., Hinrichs, H., Scholz, M., Burchert, W., & Mangun, G. R. (1998). Neural mechanisms of global and local processing: A combined PET and ERP study. *Journal of Cognitive Neuroscience*, *10*, 485–498.
- Heinze, H. J., & Münte, T. F. (1993). Electrophysiological correlates of hierarchical stimulus processing: Dissociation between onset and later stages of global and local target processing. *Neuropsychologia*, *31*, 841–852.
- Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 222–234.
- Hübner, R. (1996a). The efficiency of different cue types for reducing spatial-frequency uncertainty. *Vision Research*, *36*, 401–408.
- Hübner, R. (1996b). Specific effects of spatial-frequency uncertainty and different cue types on contrast detection: Data and models. *Vision Research*, *36*, 3429–3439.
- Hübner, R. (1997). The effect of spatial frequency on global precedence and hemispheric differences. *Perception & Psychophysics*, *59*, 187–201.
- Hübner, R. (1998). Hemispheric differences in global/local processing revealed by same–different judgements. *Visual Cognition*, *5*, 457–478.
- Hübner, R. (2000). Attention shifting between global and local target levels: The persistence of level-repetition effects. *Visual Cognition*, *7*, 465–484.
- Hübner, R., Futterer, T., & Steinhauser, M. (2001). On attentional control as a source of residual shift costs: Evidence from two-component task shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 640–653.
- Hübner, R., & Malinowski, P. (2002). The effect of response competition on functional hemispheric asymmetries for global/local processing. *Perception & Psychophysics*, *64*, 1290–1300.
- Hughes, H. C. (1986). Asymmetric interference between components of suprathreshold compound gratings. *Perception & Psychophysics*, *40*, 241–250.
- Hughes, H. C., Nozawa, G., & Kitterle, F. (1996). Global precedence, spatial frequency channels, and the statistics of natural images. *Journal of Cognitive Neuroscience*, *8*, 197–230.
- Ivry, R. B., & Robertson, L. C. (1998). *The two sides of perception*. Cambridge, MA: MIT Press.
- Johannes, S., Wieringa, B. M., Matzke, M., & Münte, T. F. (1996). Hierarchical visual stimuli: Electrophysiological evidence for separate left hemispheric global and local processing mechanisms in humans. *Neuroscience Letters*, *210*, 111–114.
- Kim, N., Ivry, R. B., & Robertson, L. C. (1999). Sequential priming in hierarchically organized figures: Effects of target level and target resolution. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 715–729.
- Kimchi, R. (1992). Primacy of wholistic processing and global/local paradigm: A critical review. *Psychological Bulletin*, *112*, 24–38.
- Kinchla, R. A. (1974). Detecting target elements in multielement arrays: A confusability model. *Perception & Psychophysics*, *15*, 149–158.
- Kinchla, R. A., Solis-Macias, V., & Hoffman, J. (1983). Attending to different levels of structure in a visual image. *Perception & Psychophysics*, *33*, 1–10.
- Kinchla, R. A., & Wolfe, J. M. (1979). The order of visual processing: “Top-down,” “bottom-up,” or “middle-out.” *Perception & Psychophysics*, *25*, 225–231.
- Kitterle, F. L., Christman, S., & Conesa, J. (1993). Hemispheric differences

- in the interference among components of compound gratings. *Perception & Psychophysics*, *54*, 785–793.
- Lamb, M. R., London, B., Pond, H. M., & Whitt, K. A. (1998). Automatic and controlled processes in the analysis of hierarchical structure. *Psychological Science*, *9*, 14–19.
- Lamb, M. R., Pond, H. M., & Zahir, G. (2000). Contributions of automatic and controlled processes to the analysis of hierarchical structure. *Journal of Experimental Psychology: Human Perception and Performance*, *26*, 234–245.
- Lamb, M. R., & Robertson, L. C. (1988). The processing of hierarchical stimuli: Effects of retinal locus, locational uncertainty and stimulus identity. *Perception & Psychophysics*, *44*, 172–181.
- Lamb, M. R., Robertson, L. C., & Knight, R. T. (1989). Attention and interference in the processing of global and local information: Effects of unilateral temporal-parietal junction lesions. *Neuropsychologia*, *4*, 471–483.
- Lamb, M. R., & Yund, E. W. (1993). The role of spatial frequency in the processing of hierarchically organized stimuli. *Perception & Psychophysics*, *54*, 773–784.
- Lamb, M. R., & Yund, E. W. (1996). Spatial frequency and attention: Effects of level-, target-, and location-repetition on the processing of global and local forms. *Perception & Psychophysics*, *58*, 363–373.
- Lamb, M. R., & Yund, E. W. (2000). The role of spatial frequency in cued shifts of attention between global and local forms. *Perception & Psychophysics*, *62*, 753–761.
- Livingstone, M., & Hubel, D. (1988, May). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, *240*, 740–749.
- MacLeod, C. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*, 163–203.
- Malinowski, P., Hübner, R., Keil, A., & Gruber, T. (2002). The influence of response competition on cerebral asymmetries for processing hierarchical stimuli revealed by ERP recordings. *Experimental Brain Research*, *144*, 136–139.
- Martin, M. (1979a). Hemispheric specialization for global and local processing. *Neuropsychologia*, *17*, 33–40.
- Martin, M. (1979b). Local and global processing: The role of sparsity. *Memory & Cognition*, *7*, 476–484.
- Martinez, A., Moses, P., Frank, L., Buxton, R., Wong, E., & Stiles, J. (1997). Hemispheric asymmetries in global and local processing: Evidence from fMRI. *NeuroReport*, *8*, 1685–1689.
- May, J. G., Gutierrez, C., & Harsin, C. A. (1995). The time-course of global precedence and consistency effects. *International Journal of Neuroscience*, *80*, 237–245.
- Michimata, C., Okubo, M., & Mugishima, Y. (1999). Effects of background color on the global and local processing of hierarchically organized stimuli. *Journal of Cognitive Neuroscience*, *11*, 1–8.
- Miller, J. (1981). Global precedence in attention and decision. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1161–1174.
- Moscovitch, M. (1979). Information processing in the cerebral hemispheres. In M. S. Gazzaniga (Ed.), *Handbook of behavioral neurobiology: Vol. 2. Neuropsychology* (pp. 379–446). New York: Plenum Press.
- Müller, M. M., & Hübner, R. (2002). Can the spotlight of attention be shaped like a donut? *Psychological Science*, *13*, 119–124.
- Navon, D. (1977). Forest before the trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*, 353–393.
- Paquet, L., & Merikle, P. M. (1988). Global precedence in attended and nonattended objects. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 89–100.
- Peterzell, D. H., Harvey, L. O., Jr., & Hardyck, C. D. (1989). Spatial frequencies and the cerebral hemispheres: Contrast sensitivity, visible persistence, and letter classification. *Perception & Psychophysics*, *46*, 443–455.
- Pomerantz, J. R. (1981). Perceptual organization in information processing. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 141–180). Hillsdale, NJ: Erlbaum.
- Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma & D. G. Bouwhuis (Eds.), *Attention and performance X: Control of language processes* (pp. 531–556). London: Erlbaum.
- Prinzmetal, W., Ivry, R. B., Beck, D., & Shimizu, N. (2002). A measurement theory of illusory conjunctions. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 251–269.
- Proverbio, A. M., Minniti, A., & Zani, A. (1998). Electrophysiological evidence of a perceptual precedence of global vs. local visual information. *Cognitive Brain Research*, *6*, 321–334.
- Robertson, L. C. (1996). Attentional persistence for features of hierarchical patterns. *Journal of Experimental Psychology: General*, *125*, 227–249.
- Robertson, L. C., Egly, R., Lamb, M. R., & Kerth, L. (1993). Spatial attention and cuing to global and local levels of hierarchical structure. *Journal of Experimental Psychology: Human Perception and Performance*, *19*, 471–487.
- Robertson, L. C., & Lamb, M. R. (1991). Neuropsychological contributions to theories of part/whole organization. *Cognitive Psychology*, *23*, 299–330.
- Sanocki, T. (2001). Interaction of scale and time during object identification. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 290–302.
- Sasaki, Y., Hadjikhani, N., Fischl, B., Liu, A. K., Marret, S., Dale, A. M., & Tootell, R. B. (2001). Local and global attention are mapped retinotopically in human occipital cortex. *Proceedings of the National Academy of Sciences, USA*, *98*, 2077–2082.
- Sergent, J. (1982). The cerebral balance of power: Confrontation or cooperation? *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 253–272.
- Shulman, G. L., Sullivan, M. A., Gish, K., & Sakoda, W. J. (1986). The role of spatial frequency channels in the perception of local and global structure. *Perception*, *15*, 259–273.
- Stoffer, T. H. (1993). The time course of attentional zooming: A comparison of voluntary and involuntary allocation of attention to the levels of compound stimuli. *Psychological Research*, *56*, 14–25.
- Stoffer, T. H. (1994). Attentional zooming and the global-dominance phenomenon: Effects of level-specific cueing and abrupt visual onset. *Psychological Research*, *56*, 83–98.
- Tolhurst, D. J. (1975). Sustained and transient channels in human vision. *Vision Research*, *15*, 1151–1155.
- Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society, Series B, Biological Sciences*, *353*, 1295–1306.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107–141.
- Van Kleeck, M. H. (1989). Hemispheric differences in global versus local processing of hierarchical visual stimuli by normal subjects: New data and a meta-analysis of previous studies. *Neuropsychologia*, *27*, 1165–1178.
- Volberg, G., & Hübner, R. (2004). On the role of response conflicts and stimulus position for hemispheric differences in global/local processing: An ERP study. *Neuropsychologia*, *42*, 1805–1813.
- Vorberg, D. (1992). Aufmerksamkeitswechsel zwischen globalen und lokalen Reiz-Ebenen [Attention shifts between global and local stimulus levels]. In L. Montada (Ed.), *Bericht über den 38. Kongreß der Deutschen Gesellschaft für Psychologie in Trier 1992. Band 1* (p. 761). Göttingen, Germany: Hogrefe.
- Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception and Performance*, *8*, 562–581.

- Ward, L. M. (1983). On processing dominance: Comment on Pomerantz. *Journal of Experimental Psychology: General*, 112, 541–546.
- Ward, L. M. (1985). Covert focussing of the attentional gaze. *Canadian Journal of Psychology*, 39, 546–563.
- Weber, B., Schwarz, U., Kneifel, S., Treyer, V., & Buck, A. (2000). Hierarchical visual processing is dependent on the oculomotor system. *NeuroReport*, 11, 241–247.
- Wertheimer, M. (1922). Untersuchungen zur Lehre von der Gestalt. I. Prinzipielle Bemerkungen [Investigations of the laws of organization in perceptual forms. I.]. *Psychologische Forschungen*, 1, 47–58.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II [Investigations of the laws of organization in perceptual forms. II.]. *Psychologische Forschungen*, 4, 301–350.
- Wundt, W. (1874). *Grundzüge der physiologischen Psychologie* [Principles of physiological psychology]. Leipzig, Germany: Engelmann.
- Yovel, G., Yovel, I., & Levy, J. (2001). Hemispheric asymmetries for global and local visual perception: Effects of stimulus and task factors. *Journal of Experimental Psychology: Human Perception and Performance*, 27, 1369–1385.
- Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature*, 274, 423–428.
- Zeki, S. (1993). *A vision of the brain*. Oxford, England: Blackwell.

Appendix A

Observed Data for Experiment 3

Table A1
Proportions for the Different Response Categories in Experiment 3 for the Global Target Level

Response category	Stimulus–mask interval (milliseconds)				
	12	24	48	96	192
LVF stimuli					
TN	.384	.438	.533	.664	.729
TO	.248	.263	.272	.222	.172
NT	.074	.055	.045	.049	.035
NO	.091	.084	.032	.018	.015
OT	.054	.042	.021	.009	.006
ON	.105	.090	.068	.031	.041
OØ	.045	.027	.029	.009	.002
RVF stimuli					
TN	.405	.456	.570	.699	.761
TO	.215	.207	.196	.160	.150
NT	.089	.073	.060	.045	.034
NO	.093	.100	.059	.022	.009
OT	.046	.038	.031	.010	.004
ON	.118	.092	.074	.060	.042
OØ	.034	.033	.013	.004	.000

Note. LVF = left visual field; RVF = right visual field; T = letter at the target level; N = letter at the nontarget level; O and Ø = letters not present in the display. The first letter of the code for the response categories indicates the response to the target level, which was required first. The second letter indicates the response to the nontarget level, which was required second. For instance, TN means that both responses are correct, whereas NT means both responses are conjunction errors.

Table A2
Proportions for the Different Response Categories in Experiment 3 for the Local Target Level

Response category	Stimulus–mask interval (milliseconds)				
	12	24	48	96	192
LVF stimuli					
TN	.320	.412	.487	.619	.696
TO	.362	.303	.277	.221	.182
NT	.063	.083	.083	.040	.054
NO	.090	.087	.057	.038	.012
OT	.060	.033	.024	.013	.004
ON	.067	.054	.051	.056	.049
OØ	.037	.028	.021	.012	.004
RVF stimuli					
TN	.345	.395	.459	.593	.667
TO	.360	.377	.344	.292	.206
NT	.062	.064	.059	.035	.055
NO	.067	.061	.049	.031	.019
OT	.057	.036	.021	.014	.007
ON	.062	.040	.053	.027	.042
OØ	.047	.027	.015	.008	.003

Note. LVF = left visual field; RVF = right visual field; T = letter at the target level; N = letter at the nontarget level; O and Ø = letters not present in the display. The first letter of the code for the response categories indicates the response to the target level, which was required first. The second letter indicates the response to the nontarget level, which was required second. For instance, TN means that both responses are correct, whereas NT means both responses are conjunction errors.

(Appendixes follow)

Appendix B

Parameters for the Integration Model

Table B1

Parameters and Goodness-of-Fit Values for the Integration Model for the Global and Local Target Levels

Parameter or goodness-of-fit measure	Stimulus-mask interval (milliseconds)				
	12	24	48	96	192
Global					
<i>t</i>	.573	.633	.762	.897	.926
<i>n</i>	.454	.480	.547	.690	.751
α , LVF	.406	.677	.640	.710	.790
α , RVF	.377	.468	.593	.547	.770
β , LVF	.616	.631	.633	.366	.499
β , RVF	.566	.517	.581	.586	.596
G^2	46.407	55.204	47.131	63.765	60.554
SSE	.025	.023	.017	.018	.021
Local					
<i>t</i>	.647	.728	.794	.869	.901
<i>n</i>	.288	.372	.481	.579	.711
α , LVF	.537	.444	.474	.691	.659
α , RVF	.617	.544	.559	.658	.453
β , LVF	.492	.558	.458	.642	.561
β , RVF	.635	.507	.583	.534	.655
G^2	52.526	68.778	64.448	65.630	78.794
SSE	.043	.032	.031	.018	.024

Note. LVF = left visual field; RVF = right visual field; SSE = sum of squared errors.

Appendix C

Predicted Proportions for Experiment 3

Table C1

Predicted Proportions for the Different Response Categories in Experiment 3 for the Global Target Level

Response category	Stimulus–mask interval (milliseconds)				
	12	24	48	96	192
LVF stimuli					
TN	.396	.456	.559	.704	.776
TO	.225	.233	.232	.183	.144
NT	.086	.059	.044	.041	.027
NO	.080	.074	.042	.020	.012
OT	.052	.042	.028	.010	.010
ON	.118	.010	.074	.035	.025
OØ	.044	.035	.021	.007	.006
RVF stimuli					
TN	.402	.439	.548	.702	.776
TO	.221	.229	.230	.178	.144
NT	.080	.076	.056	.043	.023
NO	.082	.076	.043	.017	.010
OT	.056	.047	.030	.015	.009
ON	.115	.098	.073	.039	.027
OØ	.044	.035	.021	.007	.006

Note. LVF = left visual field; RVF = right visual field; T = letter at the target level; N = letter at the nontarget level; O and Ø = letters not present in the display. The first letter of the code for the response categories indicates the response to the target level, which was required first. The second letter indicates the response to the nontarget level, which was required second. For instance, TN means that both responses are correct, whereas NT means both responses are conjunction errors.

Table C2

Predicted Proportions for the Different Response Categories in Experiment 3 for the Local Target Level

Response category	Stimulus–mask interval (milliseconds)				
	12	24	48	96	192
LVF stimuli					
TN	.360	.416	.502	.622	.722
TO	.332	.310	.276	.243	.173
NT	.061	.079	.074	.042	.037
NO	.072	.056	.049	.024	.023
OT	.056	.046	.031	.016	.008
ON	.076	.062	.049	.043	.032
OØ	.043	.032	.020	.011	.006
RVF stimuli					
TN	.363	.428	.517	.633	.713
TO	.334	.312	.278	.243	.167
NT	.058	.067	.059	.031	.045
NO	.068	.057	.043	.030	.021
OT	.054	.043	.029	.016	.014
ON	.081	.061	.055	.036	.034
OØ	.043	.032	.020	.011	.006

Note. LVF = left visual field; RVF = right visual field; T = letter at the target level; N = letter at the nontarget level; O and Ø = letters not present in the display. The first letter of the code for the response categories indicates the response to the target level, which was required first. The second letter indicates the response to the nontarget level, which was required second. For instance, TN means that both responses are correct, whereas NT means both responses are conjunction errors.

■