Probing cognitive processes through the structure of event-related potentials during learning: an experimental and theoretical analysis

Jean-Paul Banquet and Stephen Grossberg

Data reporting correlated changes, due to learning, in the amplitudes and chronometry of several event-related potentials (ERPs) are compared to neural explanations and predictions of the adaptive resonance theory. The ERP components processing negativity (PN), early positive wave (P120), N200, and P300 covary with model processes of attentional priming and top-down expectancy learning, matching of bottom-up input patterns with learned top-down expectations, mismatch-mediated activation of the orienting subsystem, reset by the orienting subsystem of recognition codes in short-term memory, and direct activation of recognition codes via a bottom-up adaptive filter. These model mechanisms enable a recognition code to be learned in a self-stabilizing fashion in response to an input environment of arbitrary complexity. Thus spatiotemporal correlations among several ERPs during learning provide important evidence in support of postulated neural mechanisms for self-stabilizing self-organization of cognitive recognition codes.

I. Introduction

This paper describes a convergence between theoretical predictions and recent data concerning event-related potentials (ERP). The theory in question is adaptive resonance theory (ART), which was introduced by Grossberg[1,2] and has since undergone extensive empirical and formal development. This theory arose through an analysis of neural mechanisms that are capable of self-organizing and self-stabilizing the learning of cognitive recognition codes in response to arbitrarily complex input environments. Such mechanisms can buffer their learning against inappropriate recoding by the blooming buzzing confusion of a continuous stream of irrelevant experiences, yet can also learn quickly from novel environments which are important to behavioral survival.

This dynamic balance between memory stability and adaptive plasticity is controlled by the action of learned top-down expectations and pattern matching processes. An ideal experimental paradigm for testing such a theory is thus one in which expectations are learned, matching processes are parametrically manipulated, and the experimental measures are sensitive to state-dependent patterning of neuronal activities across large ensembles of cells. Banquet et al.,[3-7] Johnson and Donchin,[8] and Squires et al.,[9] have collected such data from event-related potential (ERP) experiments designed to determine the influence of probabilistic contextual information on processing strategies in a choice reaction time (RT) task.

The goal herein is to compare theoretical predictions of the adaptive resonance theory to data concerning the transformations which occur among four ERP components—processing negativity (PN), early positive wave (P120), N200, and P300—as learning proceeds (Table I). The theory anticipated the discovery of two of these components—processing negativity[10] or Nd[11] and early positive wave.[12,13] The data of Banquet et al.[3,4,6,7] provide detailed information which supports the theoretically postulated correlations between these waves. These data were selected herein for particular comparison with ART mechanisms because they analyze how the waveforms of several ERP components covary across experimental trials during which recognition learning occurs. Such spatiotemporal correlations provide a much stronger test of theoretical predictions than do ERP data which describe only the existence of an individual ERP component or the behavior of an individual component during performance trials which do not include a learning manipulation. Relevant ERP data from a number of other laboratories will also be discussed.

Jean-Paul Banquet is with Hopital de la Salpetriere, Laboratoire d'Electrophysiologie et de Neurophysiologie Appliquee-CNRS, 47 Boulevard de l'Hopital, 75651 Paris, France, and Stephen Grossberg is with Boston University, Center for Adaptive Systems, 111 Cummington Street, Boston, Massachusetts 02215.

ERP	ART mechanism
	Read-out learned top-down expectation from processing level F_2 to processing level F_1 and amplification of total activity at F_1 in a match situation.
P120	Reduction of total activity at F_1 in a mismatch situation.
N200	Disinhibition of the orienting subsystem A in a mismatch situation.
P300	Reset of short-term memory at F_2.

A brief review is first given of the main operations postulated by the adaptive resonance theory. More detailed recent expositions are given by Grossberg[14,15] and by Carpenter and Grossberg.[16-19] The Carpenter and Grossberg articles describe, moreover, mathematical analyses and computer simulations which characterize the trial-by-trial course of category learning and recognition in a number of specific examples. The theoretical introduction in Secs. II–VII is followed by a summary of the meaning of the different ERPs in Sec. VIII and a presentation of the experimental data in Sec. IX. The remainder of the paper compares theory with data and calls attention to theoretical predictions which have not yet been supported or disconfirmed by ERP experiments.

II. Attentional Subsystem and Orienting Subsystem

Within the adaptive resonance theory, interactions between two functionally complementary subsystems are often used to process familiar and unfamiliar events. An attentional subsystem learns evermore precise internal representations of familiar events. It also builds up the learned top-down expectations that help to stabilize the learned bottom-up recognition codes of familiar events. By itself, however, the attentional subsystem is unable simultaneously to maintain stable representations of familiar recognition codes and to create new recognition codes for unfamiliar patterns in response to certain input environments. An isolated attentional subsystem may be either too rigid to create new categories for unfamiliar patterns, or so unstable that it can ceaselessly recode the categories for familiar patterns as it learns about unfamiliar patterns.[1,16,18] The latter difficulty is typical of many learning systems whose plasticity is not switched off through time.

The second subsystem is an orienting subsystem that overcomes the rigidity of the attentional subsystem when unfamiliar events occur and enables the attentional subsystem to learn from these novel experiences without destabilizing its established learning. Interactions between the attentional subsystem and the orienting subsystem are essential for expressing whether a pattern is familiar and well represented by an existing recognition code, or unfamiliar and in need of a new recognition code.

All input events start to be processed by the attentional subsystem. A familiar event can activate a recognition code which reads out a top-down template, or expectation, which is matched against the input within the attentional subsystem (Fig. 1). A successful ap-

proximate match can deform, amplify, and sustain in short-term memory (STM) the activity pattern that was initially activated by the input within the attentional subsystem. Amplified, or resonant, STM activities throughout the attentional subsystem constitute the fully elaborated recognition event. They inhibit the orienting subsystem and engage the learning, or long-term memory (LTM), process. A familiar event can maintain or modify its prior learning as its recognition takes place.

An unfamiliar event also starts to be processed by the attentional subsystem. Such an event may also activate a recognition code which thereupon reads-out a top-down template. If the unfamiliar event can approximately match this template, it can be recognized as an exemplar of the recognition code on its first presentation. If the unfamiliar event is too different from familiar exemplars of the sampled code, it cannot approximately match this template. A sufficiently large mismatch within the attentional subsystem activates the orienting subsystem. Activation of the orienting subsystem functionally expresses the novelty, or unexpectedness, of the unfamiliar event. The orienting subsystem, in turn, rapidly resets the active representation within the attentional subsystem as it simultaneously energizes an orienting response.

The reset of the attentional subsystem by the orienting subsystem leads to the selection of a new representation within the attentional subsystem. This new representation may cause yet another mismatch,
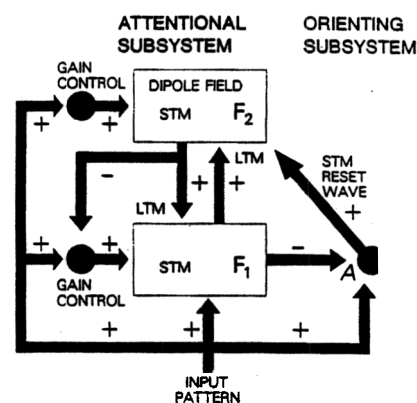

Fig. 1. Anatomy of the attentional-orienting system: Two successive stages, F_1 and F_2, of the attentional subsystem encode patterns of activation in short-term memory (STM). Bottom-up and top-down pathways between F_1 and F_2 contain adaptive long-term memory (LTM) traces which multiply the signals in these pathways. The remainder of the circuit modulates these STM and LTM processes. Modulation by gain control enables F_1 to distinguish between bottom-up input patterns and top-down priming, or template, patterns, as well as to match these bottom-up and top-down patterns. Gain control signals also enable F_2 to react supraliminally to signals from F_1 while an input pattern is on. The orienting subsystem generates a reset wave to F_2 when sufficiently large mismatches between bottom-up and top-down patterns occur at F_1. This reset wave selectively and enduringly inhibits active F_2 cells until the input is shut off. (Reprinted with permission from Carpenter and Grossberg.[18])
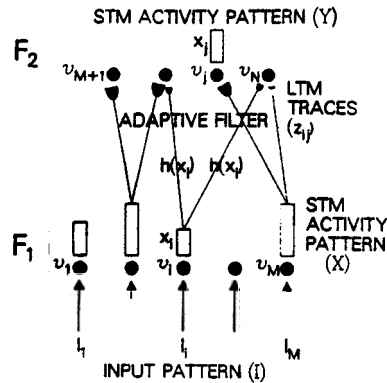
Fig. 2. Stages of bottom-up activation: The input pattern I generates a pattern of STM activation $X = (x_1, x_2, \ldots, x_m)$ across the nodes v_i of F_1. Sufficiently active F_1 nodes emit bottom-up signals $h(x_i)$ to F_2. This signal pattern, which is denoted by S in Fig. 3, is multiplied, or gated, by long-term memory (LTM) traces z_{ij} within the $F_1 \rightarrow F_2$ pathways. The LTM gated signals are summed before activating their target nodes in F_2. This LTM gated and summed signal pattern, which is denoted by T in Fig. 3, generates a pattern of activation $Y = (x_{M+1}, x_{M+2}, \ldots, x_N)$ across the nodes v_j of F_2. (Reprinted with permission from Carpenter and Grossberg.[18])

hence another STM reset event and the selection of yet another representation. In this way, the orienting subsystem mediates a rapid search which continues until a representation is found that does not cause a large mismatch. Then the search ends, an STM resonance develops, and the LTM learning process can encode the active representation to which the search led. The system's recognition codes are hereby altered in either of two ways. If the search leads to an established code, learning may refine the criteria, namely, the LTM bottom-up code and top-down expectation, for accessing that code. If the search leads to uncommitted cells, learning can add a new learned representation to the total recognition code.

III. Bottom-Up Adaptive Filtering and Contrast-Enhancement in Short-Term Memory

The main mechanisms of the theory are now introduced in a qualitative way by considering the typical network reactions to a single input pattern I within a temporal stream of input patterns. Each input pattern may be the output pattern of a preprocessing stage. The input pattern I is received at the stage F_1 of the attentional subsystem. Pattern I is transformed into a pattern X of activation across the nodes of F_1 (Fig. 2). The transformed pattern X represents a pattern in short-term memory (STM). In F_1 each node whose activity is sufficiently large generates excitatory signals along pathways to target nodes at the next processing stage F_2. A pattern X of STM activates across F_1 hereby elicits a pattern S of output signals from F_1. When a signal from a node in F_1 is carried along a pathway to F_2, the signal is multiplied, or gated, by the pathway's long-term memory (LTM) trace. The LTM gated signal (i.e., signal times LTM trace), not the signal alone, reaches the target node.

Each target node sums up all its LTM gated signals. In this way, pattern S generates a pattern T of LTM gated and summed input signals to F_2 [Fig. 3(a)]. The transformation from S to T is called an adaptive filter.

The input pattern T to F_2 is quickly transformed by interactions among the nodes of F_2. These interactions contrast-enhance the input pattern T. The resulting pattern of activation across F_2 is a new pattern Y. The contrast-enhanced pattern Y, rather than the input pattern T, begins to be stored in STM by F_2.

Only those nodes of F_2 which maintain stored activity in STM can elicit new learning at contiguous LTM traces. Whereas all the LTM traces in the adaptive filter, and thus all learned past experiences of the network, are used to determine recognition via the transformation $I \rightarrow X \rightarrow S \rightarrow T \rightarrow Y$, only those LTM traces in the pathways $S \rightarrow T$ whose STM activities Y in F_2 survive the contrast-enhancement process can learn in response to the activity pattern X.

IV. Top-Down Template Matching and Stabilization of Code Learning

We now summarize how top-down template matching can stabilize code learning. In order to do so, top-down template matching at F_1 must be able to prevent learning at bottom-up LTM traces whose contiguous F_2 nodes are only momentarily activated in STM. This ability depends on the different rates at which STM activities and LTM traces can change. The STM transformation $I \rightarrow X \rightarrow S \rightarrow T \rightarrow Y$ takes place very quickly; that is, much more quickly than the rate at which the LTM traces in the adaptive filter $S \rightarrow T$ can change. As soon as the bottom-up STM transformation $X \rightarrow Y$ takes place, the STM activities Y in F_2 elicit a top-down excitatory signal pattern U back to F_1. Only sufficiently large STM activities in Y elicit signals in U along the feedback pathways $F_2 \rightarrow F_1$.

As in the bottom-up adaptive filter, the top-down signals U are also gated by LTM traces before the LTM gated signals are summed at F_1 nodes. The pattern U of output signals from F_2 hereby generates a pattern V of LTM gated and summed input signals to F_1. The transformation from U to V is thus also an adaptive filter. The pattern V is called a top-down template or learned expectation.

Two sources of input now perturb F_1: the bottom-up input pattern I which gave rise to the original activity pattern X, and the top-down template pattern V that resulted from activating X. The activity pattern X^* across F_1 that is induced by I and V taken together is typically different from the activity pattern X that was previously induced by I alone. In particular, F_1 acts to match V against I. The result of this matching process determines the future course of learning and recognition by the network.

The entire activation sequence

$$I \rightarrow X \rightarrow S \rightarrow T \rightarrow Y \rightarrow U \rightarrow V \rightarrow X^* \tag{1}$$

takes place very quickly relative to the rate with which the LTM traces in either the bottom-up adaptive filter $S \rightarrow T$ or the top-down adaptive filter $U \rightarrow V$ can
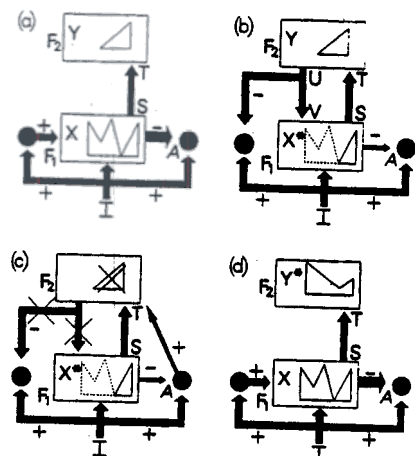
Fig. 3. Search for a correct F_2 code: (a) The input pattern I generates the specific STM activity pattern X at F_1 as it nonspecifically activates A. Pattern X both inhibits A and generates the output signal pattern S. Signal pattern S is transformed into the input pattern T, which activates the STM pattern Y across F_2. (b) Pattern Y generates the top-down signal pattern U which is transformed into the template pattern V. If V mismatches I at F_1, a new STM activity pattern X^* is generated at F_1. The reduction in total STM activity which occurs when X is transformed into X^* causes a decrease in the total inhibition from F_1 to A. (c) Then the input-driven activation of A can release a nonspecific arousal wave to F_2, which resets the STM pattern Y at F_2. (d) After Y is inhibited, its top-down template is eliminated, and X can be reinstated at F_1. Now X once again generates input pattern T to F_2, but since Y remains inhibited, T can activate a different STM pattern Y^* at F_2. If the top-down template due to Y^* also mismatches I at F_1, the rapid search for an appropriate F_2 code continues. (Reprinted with permission from Carpenter and Grossberg.[18])

change. Even though none of the LTM traces changes during such a short time, their prior learning strongly influences the STM patterns Y and X^* that evolve within the network. We now review how a match or mismatch of I and V at F_1 regulates the course of learning in response to the pattern I.

V. STM Reset and Search

Level F_1 can compute a match or mismatch between a bottom-up input pattern I and a top-down template pattern V, but it cannot compute which STM pattern Y across F_2 generated the template pattern V. Thus the outcome of matching at F_1 must have a nonspecific effect on F_2 that can potentially influence all the F_2 nodes, any one of which may have read-out V. The internal organization of F_2 must be the agent whereby this nonspecific event, which is called an arousal burst or a reset wave, selectively alters the stored STM activity pattern Y. A mismatch of I and V within F_1 generates a nonspecific arousal burst that inhibits the active populations in F_2 which read-out V. In this way, an erroneous STM representation Y at F_2 is quickly eliminated before any LTM traces can encode this error.

The attentional subsystem and the orienting subsystem work together to carry out these interactions. All learning takes place within the attentional subsystem.

All matches and mismatches are computed within the attentional subsystem. The orienting subsystem is the source of the nonspecific arousal bursts that reset STM within level F_2 of the attentional subsystem. The outcome of matching within F_1 determines whether such an arousal burst will be generated by the orienting subsystem. Thus the orienting system mediates reset of F_2 due to mismatches within F_1.

Figure 3 depicts a typical interaction between the attentional subsystem and the orienting subsystem. In Figure 3(a), an input pattern I instates an STM activity pattern X across F_1. The input pattern I also excites the orienting population A, but pattern X at F_1 inhibits A before it can generate an output signal.

Activity pattern X also generates an output pattern S which, via the bottom-up adaptive filter, instates an STM activity pattern Y across F_2. In Fig. 3(b), pattern Y reads a top-down template pattern V into F_1. Template V mismatches input I, thereby significantly inhibiting STM activity across F_1. The amount by which activity in X is attenuated to generate X^* depends on how much of the input pattern I is encoded within the template pattern V.

When a mismatch causes a sufficient attenuation of STM activity across F_1, this activity no longer prevents the arousal source A from firing. Typically, if the total activity in X^* is less than a fixed fraction of the total activity in X, then A is activated. This fraction is called the vigilance parameter of the network.[18,19] The vigilance parameter can be altered by environmental feedback, notably punishment. Higher vigilance enables the network to make finer discriminations between pattern exemplars and to learn more selective recognition codes. Lower vigilance has the opposite effect.

Net activation of the orienting subsystem A is due to reduction in the total inhibition from F_1 to A when the total activity X^* decreases due to a pattern mismatch. Thus A is activated due to a disinhibitory process. Figure 3(c) depicts how disinhibition of A releases a nonspecific arousal burst to F_2. This arousal burst, in turn, selectively inhibits the most active populations in F_2. This inhibition is long-lasting. One physiological design for F_2 processing which has these reset properties is a gated dipole field.[15,20,21] A gated dipole field consists of opponent processing channels whose signals are multiplied, or gated, by habituating chemical transmitters. A nonspecific arousal burst induces selective and enduring inhibition within a gated dipole field.

In Fig. 3(c), inhibition of Y leads to inhibition of the top-down template V, and thereby terminates the mismatch between I and V. Input pattern I can thus reinstate the activity pattern X across F_1, which again generates the output pattern S from F_1 and the input pattern T to F_2. Due to the enduring, arousal-initiated, selective inhibition at F_2, the input pattern T can no longer activate the same pattern Y at F_2. A new pattern Y^* is thus generated at F_2 by I [Fig. 3(d)]. Despite the fact that some F_2 nodes may remain inhibited by the STM reset property, the new pattern Y^*

may encode large STM activities. This is because level F_2 is designed so that its total suprathreshold activity remains approximately constant, or normalized, despite the fact that some of its nodes may remain inhibited by the STM reset mechanism. This property is related to the limited capacity of STM. A physiological process capable of achieving the STM normalization property, based on recurrent on-center off-surround interactions among cells obeying membrane equations, is described by Grossberg.[15,20,21]

The new activity pattern Y^* reads-out a new top-down template pattern V^*. If a mismatch again occurs at F_1, the orienting subsystem is again engaged, thereby leading to another arousal-mediated reset of STM at F_2. In this way, a rapid series of STM matching and reset events may occur. Such an STM matching and reset series controls a search of LTM that sequentially engages the novelty-sensitive orienting subsystem. The mismatch-mediated search of LTM ends when an STM pattern across F_2 reads-out a top-down template which either matches I, to the degree of accuracy tolerated by the orienting subsystem due to the setting of the vigilance parameter, or which has not yet undergone any prior learning. In the former case, the bottom-up code and top-down template of the selected representation may be refined by learning any new information that is in the input exemplar I. In the latter case, a new recognition code is established as a bottom-up code and top-down template are learned for the first time by the selected representation in response to I.

The mismatch-mediated search of STM at F_2 may profitably be thought of as a sequential test of hypotheses. Each reset wave from A to F_2 inhibits an incorrect hypothesis. The next input wave from F_1 to F_2 is evaluated conditional on the hypothesis that the previous interpretations by F_2 of the input at F_1 were incorrect. Thus an adaptive resonance theory architecture is a cognitive system capable of discovering, testing, and learning hypotheses in a stable fashion in response to input environments whose statistical properties may change unpredictably or may be arbitrarily complex.

VI. Attentional Gain Control and Attentional Priming

The same top-down template matching process which stabilizes learning is also a mechanism of attentional priming. Consider, for example, a situation in which F_2 is activated by a level other than F_1 before F_1 is itself activated. In such a situation, F_2 can generate a top-down template V to F_1. The level F_1 is then primed, or ready, to receive a bottom-up input that may or may not match the active expectancy. Level F_1 can be primed to receive a bottom-up input without necessarily eliciting suprathreshold output signals in response to the priming expectancy. If this were not possible, every priming event would lead to suprathreshold consequences. Such a property would prevent subliminal anticipation of a future event.

On the other hand, an input pattern I must be able to generate a suprathreshold activity pattern X even if no top-down expectancy is active across F_1 (Fig. 3). How does F_1 know that it should generate a suprathreshold reaction to a bottom-up input pattern but not to a top-down input pattern? In both cases, an input pattern stimulates F_1 cells. Some auxiliary mechanism must exist to distinguish between bottom-up and top-down inputs. Such considerations led Grossberg[15,22] to distinguish this auxiliary mechanism, called attentional gain control, from attentional priming by the top-down template itself. Carpenter and Grossberg[16,18] have developed this qualitative distinction into a quantitative computational mechanism, and Grossberg and Stone[23] have used the distinction to help explain data from word recognition experiments.

In particular, the attentional priming mechanism delivers specific template patterns to F_1. The attentional gain control mechanism has a nonspecific effect on the sensitivity with which F_1 responds to the template pattern, as well as to other patterns received by F_1. With the addition of attentional gain control, a qualitative explanation can be given of how F_1 can tell the difference between bottom-up and top-down signal patterns.

The need to dissociate attentional priming from attentional gain control can also be seen from the fact that top-down priming events do not lead necessarily to subliminal reactions at F_1. Under certain circumstances, top-down expectancies can lead to suprathreshold consequences. Internal conversations or images can, for example, be experienced at will. Thus a difference exists between the read-out of a top-down template, which is a mechanism of attentional priming, and the translation of this operation into suprathreshold signals due to attentional gain control. An act of will can amplify attentional gain control signals to elicit a suprathreshold reaction at F_1 in response to an attentional priming pattern from F_2.

Figures 4(a)–(c) depict a scheme whereby subliminal reactions to top-down signals, supraliminal reactions to bottom-up signals, and supraliminal reactions to matched bottom-up and top-down signals can be achieved. Figure 4(d) shows how competitive interactions between the attentional gain control mechanisms of different modalities can prevent F_1 from automatically generating a supraliminal reaction to bottom-up signals when attention shifts from that modality to another.

VII. Matching via the 2/3 Rule

A rule for matching bottom-up input patterns with top-down templates, called the 2/3 Rule,[16,18] follows naturally from the distinction between attentional gain control and attentional priming. It says that two out of three signal sources must activate an F_1 node for that node to generate suprathreshold output signals. In Fig. 4(b), for example, during bottom-up processing, a suprathreshold node in F_1 is one which receives a specific input from the input pattern I and a nonspecific attentional gain control signal. All other nodes in F_1 receive only the nonspecific gain control signal. Since these cells receive inputs from only one pathway, they do not fire.
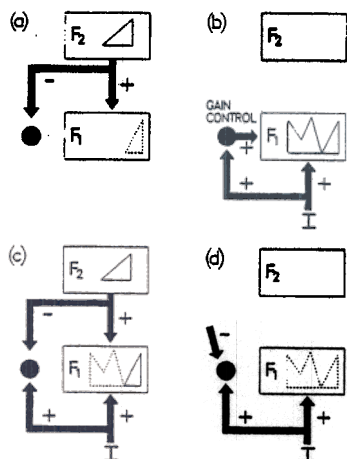
Fig. 4. Matching by 2/3 Rule: (a) A top-down template from F_2 inhibits the attentional gain control source as it subliminally primes target F_1 cells. (b) A bottom-up input activates both the (nonspecific) attentional gain control channel and certain F_1 cells. Only F_1 cells that receive bottom-up inputs and gain control signals can become supraliminally active. (c) When a bottom-up input pattern and a top-down template are simultaneously active, only those F_1 cells that receive inputs from both sources can become supraliminally active, since the gain control source is inhibited. (d) Intermodal competition can shut off the attentional gain control source and thereby prevent a bottom-up input from supraliminally activating F_1 when attention is directed to a different modality.

In Fig. 4(a), during top-down processing, or priming, some nodes in F_1 receive a template signal from F_2, whereas other nodes receive no signal whatsoever. All the nodes of F_1 receive inputs from at most one of their three possible input sources. Hence no cells in F_1 are supraliminally activated by a top-down template.

During simultaneous bottom-up and top-down signaling, the attentional gain control signal is inhibited by the top-down channel [Fig. 4(c)]. Despite this fact, some nodes of F_1 may receive sufficiently large inputs from both the bottom-up and the top-down signal patterns to generate suprathreshold outputs. Other nodes may receive inputs from the top-down template pattern or the bottom-up input pattern, but not both. These nodes receive signals from only one of their possible sources, hence do not fire. Cells which receive no inputs do not fire either. Thus only cells that are conjointly activated by the bottom-up input and the top-down template can fire when a top-down template is active. The 2/3 Rule clarifies the apparent paradox that the addition of top-down excitatory inputs to F_1 can lead to an overall decrease, or collapse, in the F_1 STM activity [Figs. 3(a) and (b)]. Carpenter and Grossberg[18] have shown that learning may become unstable if the 2/3 Rule is violated, but is absolutely stable when the 2/3 Rule is reinstated.

VIII. ERP Components

The understanding of the experimental results and of the discussion requires a brief presentation of the main ERP components. The term wave or deflection refers to the different possible morphologies of an ERP. The term component is attributed to an independent source of variability or a generator of the ERP waveform.

An early component (100 ms) labeled Nd (negative displacement) was first demonstrated by Hillyard et al.[24] in a dichotic listening experiment with short (100–800-ms) random interstimulus intervals (ISI). The attended stimuli elicited an enhanced negativity at ~100 ms from stimulus onset ($N1$), compared to the $N1$ elicited by an identical but unattended stimulus to the other ear. The authors interpreted this effect as a selective increase of the activity in the $N1$ generators and related it to the stimulus set mechanism of early stimulus filtering on the basis of physical features, which was proposed by Broadbent.[25]

With a longer and constant ISI (800 ms), Naatanen et al.[10] observed a selective attention effect going beyond the time window of the $N1$ component and extending for several hundred milliseconds. These authors proposed that this negative shift expresses a different ERP component of endogenous origin (the $N1$ being exogenous) which they labeled the processing negativity (PN). PN can be taken as an early sign of selective attention in a dichotic listening program. Our paradigm, however, is not a situation of dichotic listening; two identical stimuli in the two ears have to be attended. Nevertheless, a selection has to be made between what is a stimulus and what is not, and also between target and nontarget. Since the experimental situation and the method of subtraction by which our negativity has been revealed differs from the previous ones, the first endogenous negative component will be called early negativity (EN).

An early centroparietal positivity was first described as a $P165$ by Goodin et al.[13] and as a $P100$ by Desmedt et al.[12] in response to auditory and somesthesic stimuli, respectively. In both cases, the early positivity (EP) was elicited by attending rare relevant target stimuli. A $P165$–$N200$–$P300$ complex was observed in attended sequences of deviant tones compared with standard tones.[26] The early positivity was interpreted either as an early manifestation of decision processes related to the later $N200$–$P300$ potentials[13] or as an expression of the process of sorting out and identifying input signals against target templates.[12]

The $N200$ wave of the ERPs has been one of the less easy to interpret, largely because it reflects the existence of multiple components. We will mention here only the two most widely accepted components of the $N200$ deflection. The $N2a$ was named mismatch negativity (MMN) by Naatanen et al.[10] because of its occurrence in response to stimuli physically deviant from those in the immediate past, be they attended or unattended. This is in contrast to the PN which can be elicited only by attended stimuli. MMN is sensitive to dynamic changes in the stimuli presented, such as pitch or intensity, and to the magnitude of the change. In an attention condition, it precedes or overlaps the $P165$–$N200$–$P300$ complex, which is more centrally located. MMN seems therefore to represent an automatic process which is not influenced by selective attention. It could reflect short-duration memory

processes such as sensory registers or preattentive storage taking place in the sensory cortex.[26]

N2b (Refs. 27 and 28) is a negative component which precedes P300. The topography of N2b is distributed across modalities. It is elicited by temporally unexpected or rare stimuli. Its occurrence depends not only on the degree of stimulus change, but also on the orientation of focal attention to the stimulus source. N2b could also reflect transient activation of the subcortical centers releasing the orienting reflex.[29]

P300 has been one of the most explored ERP deflections since its discovery by Sutton et al.[30] First explained in terms of different psychological constructs (task relevance, expectancy, equivocation) or theories (information theory, signal detection theory), it was later explored for its specific functional role, as a scalp manifestation of information transactions in the brain.[31] P300 soon also appeared to be a nonunitary phenomenon. Squires et al.,[32] examining P300s in response to occasional shifts in ongoing trains of tones in conditions of attention and nonattention, found components of different latency and topography during the nonattended (P3a) and the attended (P3b) conditions. Courchesne et al.[33,34] further investigated various ERP components in situations of novelty.

The P3a wave, ever since its discovery,[32] appeared to reflect events distinct from the P3b complex. In some experiments, a P3a was elicited by an unpredictable shift in an ongoing repetitive series of auditory stimuli even though the sounds were task irrelevant or not attended.[35–37] Conversely, in dichotic listening tasks[10] or in distraction situations with slightly deviant auditory stimuli,[28,29] no P3a was elicited. Only an N200 mismatch negativity could be recorded. These apparently contradictory results may be explained by the fact that at present there is no reliable measure of the degree of subjects' awareness of the stimulus shift in the ignore condition. It could well turn out that the dichotic listening paradigm is a better guarantee of a true unattended situation than a simple ignore instruction. In any case, the N200–P3a complex has been interpreted by most authors as a reflection of a mismatch detector.[32,38] Courchesne[34] has shown that the P300 amplitude response to novel events shifts from a frontal to a parietal maximum with repeated presentations. More recently, it has been shown that this component is less sensitive to the prior probability of events than P3b.[3,4] Munson et al.[39] have also described a similar component, P300E, which does not react to prior probability. Therefore, the individuality of P3a from P3b seems clear.

The frontocentral P3a thus occurs, not only for attended task-relevant events, but also for unattended, task-irrelevant intermittent stimuli, its amplitude being related to the degree of physical contrast with the background and to immediately preceding probability rather than prior probability per se.[4] The similarity of these eliciting conditions with those of N200 led several authors to regard N2b and P3a as aspects of the same process and, in particular, to relate it to the orienting reaction.

On the contrary, P3b is a later component elicited by attended task-relevant target stimuli. Subjective probability, stimulus meaning, and information transmission are the three dimensions used in the model of Johnson[40] to explain variations in P3b amplitude. Since P3b covaries with so many different variables, it has also been suggested that it represents a general subroutine invoked in different cognitive operations, such as updating of the context or of models of the environment.[31,41] Grossberg[14,42,43] postulated the existence of two parallel output pathways from the orienting subsystem A whose effects on their target networks may be compared with data about the P3a and P3b, as in Sec. XI. One branch, from A to the attentional subsystem (Sec. II), causes reset of STM. The other branch of A activates processes associated with the orienting response, including processes which gate the release of orienting movements. Recent relation of P3b amplitude to the quality of subsequent recall seems to confirm the association of P3b to short-term memory processes.[40,44]

IX. Experimental Paradigm

A. Hypotheses

The original purpose of the experimental research was to determine the influence of probabilistic contextual information[45,46] on processing strategies in a choice RT task.[3,4,6,7] This influence of contextual probability processing on single trial processing, for events delivered in Bernoulli sequences, was suggested by previous results.[47] These results showed how chronometry and amplitude evolution of P300 during practice were correlated with performance. Subjects were divided into high performance and low performance groups.

In the high performance group, the P300 peak occurred after the reaction time (RT) and increased in amplitude with practice. Conversely, in the low performance group, P300 peaked before RT, and its amplitude decreased with practice. Such variable timing of RT and P300 was first demonstrated by Ritter et al.[48] A possible interpretation of these data was that more thorough processing and/or use of the probability information (indexed by P300 amplitude) was carried out by the better performing group.

B. Experimental Procedure

One of the simplest ways to manipulate the degree of contextual processing is a passive learning procedure in an oddball paradigm with a choice RT task. In an oddball paradigm, the subject receives Bernoulli series of two types of stimuli of complementary probability. One of the stimuli is frequent or standard; the other is rare and usually serves as the target. The subject has to perform a task such as counting target stimuli or releasing a motor response in response to each target stimulus.

In this experiment, Bernoulli series of high-pitched (2000-Hz) and low-pitched (500-Hz) tones of equal intensity and duration (10 ms) were delivered through headphones at fixed ISI (1500 ms). Target probabili-

ty was also manipulated: Five consecutive runs of unequally probable (0.2/0.8 or 0.8/0.2) stimuli were followed by two runs of equally probable (0.5/0.5) stimuli and then by five runs of the unequal complementary probability (0.8/0.2 or 0.2/0.8). Each unequal probability run ended after a total of fifteen rare stimuli had been delivered. One session consisted of these three consecutive probability conditions.

Subjects performed a Go/No Go task with a lever-press response to a single type of sound in both high and low probability conditions. They were not asked to monitor the probability changes. A second session was a replication of the first session one week later. Since perceptual discrimination and motor response tasks were easy to perform and identical in the different probability conditions, it was assumed that RT and ERP differences over consecutive runs and sessions would reflect probability learning.

Learning effects could be detected within the paradigm in two different ways. A within-condition analysis was performed of the unequal probability data for each session. RTs and ERPs averaged separately during the first two runs and the last two runs of each unequal probability condition (0.2/0.8 and 0.8/0.2) were compared. A between-session analysis compared the grand averages across the five runs of the unequal probability conditions.

The shift from unequal to equal probabilities after a block of five runs served two functions. It created a mismatch condition in which prior probability learning in one condition became inappropriate due to the unsignaled change in probability. In addition, the block of equal probability runs was chosen sufficiently long to damp previous learning effects due to unequal probability, and thereby to prepare the subject for the reversed unequal probability conditions.

In summary, three experimental factors were explored by the paradigm: (1) a prior probability factor due to the different probability conditions; (2) a practice factor resulting from the five-run blocks and two sessions for each condition; (3) a mismatch factor by the unwarned shift in probability.

C. Data Recording and Analysis

EEG was recorded from six electrodes referred to linked ears, spaced at intervals of 10% of the nasion–inion distance, starting from F_z and including C_z and P_z. Supraorbital and suborbital electrodes around the right eye monitored ocular potentials. RTs and ERPs were averaged separately for the different experimental situations. An analysis of these averaged waveforms was described by Banquet et al.[3] In the present article ERPs were measured both by subtracting ERPs to the frequent stimuli from ERPs to the infrequent stimuli[49,50] and by subtracting the auditory evoked potentials obtained in a passive situation with purely random stimuli (no task) from ERPs to both rare and frequent stimuli in a task situation.[3] This last procedure was proposed by Naatanen[51] as a better way to compute processing negativity (PN). The main purpose of both subtraction techniques is to neutralize the

overlapping of $N100$ and $P200$ exogenous evoked potentials on the early endogenous ERPs. The results were tested by a two way repeated measures analysis of variance. Only results at $p < 0.01$ were considered significant.[7] Factorial analysis[45,52,53] was also performed on these ERP data.

In a Go/No Go paradigm, a motor response is made only to the target stimulus. However, to selectively perform this response, both target and nontarget stimuli are relevant and thus must be actively attended and processed. This is confirmed by the approximately equal amplitudes of the ERP components to targets and nontargets of equal probability, thereby showing an absence of task effect.[3]

The absence of task effect on the amplitude of the ERPs in this experiment (Go and No Go ERP profiles are similar) indicates: (1) the preeminence of probability over other factors in determining the amplitude of $N200$ and $P300$; (2) the efficacy of the random washout session (0.5/0.5) between the two learning sequences of complementary probability (0.8/0.2 and 0.2/0.8), even if it cannot be excluded that the learning in the second sequence of runs was faster than in the first one. The absence of task effect enables the comparison of No Go ERPs with RTs in the same probability condition, since Go and No Go ERP amplitudes are similar. These nontarget data provide ERP measures of cognitive processing that are relatively uncontaminated by motor components and are therefore the primary focus of the experimental analysis.

The analysis of the first experimental session was broken up into four cases, two of which generate almost identical ERP profiles: (1) frequent stimuli in the early runs (Fe); (2) rare stimuli in the early runs (Re); (3) frequent stimuli in the late runs (Fl); (4) rare stimuli in the late runs (Rl). According to the law of the stimulus probability effect, faster reaction times were expected to frequent target stimuli than to rare target stimuli. In view of previous ERP results, larger ERP
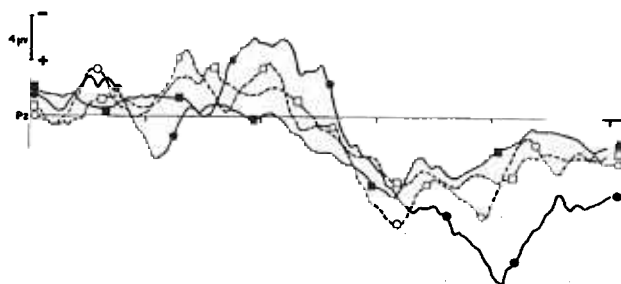

Fig. 5. Grand average chronograms of the ERPs recorded at P_z, obtained after subtraction of the auditory evoked potentials for each subject in a passive situation (no task) from the No Go condition of the task situation. The ERP amplitudes (in microvolts) are displayed as a function of time (a unit scale, 100 ms), stimulus probability (squares, high probability; circles, low probability), and runs (dotted lines, first two runs; full lines, last two runs). Cases Fe and Re (dotted lines with squares and circles, respectively) and cases Fl and Rl (full lines with squares and circles, respectively) are superimposed.

components were expected to rare stimuli than to frequent stimuli.

X. Experimental Results: ERP Profiles

The results will be reported in greater detail elsewhere.[7] There was good agreement between behavioral (RT) and ERP data. At both levels, the four cases combined to form three patterns or profiles of response (Fig. 5).

(1) and (2) *Cases Fe and Re:* During the early runs of the first session, whether with low or high probability nontargets (Fig. 5, dotted lines), there was practically no difference between the ERP profiles in response to rare or frequent stimuli. The only significant amplitude difference occurred for the $P300$ component at P_z, which is called $P3b$. In addition, RTs to frequent and rare stimuli were not significantly different.

(3) *Case Fl:* In contrast, during the late runs of the first session, a high probability nontarget caused a widespread flattening of both negative and positive components with only one alteration between negativity and then positivity (full-line square in Fig. 5). Simultaneously, RTs decreased dramatically.

(4) *Case Rl:* Greater amplitude peaks appeared at three points of the time axes: early positivity $P120$, $N200$ mismatch, and $P300$. In parallel, RT increased compared to cases *Fe* and *Re.* These trends were even more striking when the data are replotted as in Fig. 6, where the peaks of the ERP components are positioned at their mean latency.

A comparison of ERP amplitudes in cases (3) and (4) with cases (1) and (2) in light of the functional significance of the ERP components is consistent with the following conclusions. A high-frequency event leads to learning of an expectancy which tends to be matched during condition *Fl* and tends to be mismatched during condition *Rl.* This possibility is supported by a component-by-component parametric
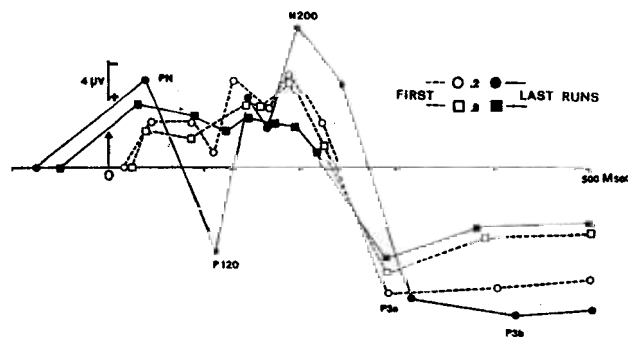

Fig. 6. Mean amplitudes of the ERP components over all subjects displayed with the same code as in Fig. 5. The peaks were measured on interpolated maps at their maximum amplitude value, in different locations. Therefore the diagram does not represent a unique electrode site. Cases *He, Le, Hl,* and *Ll* are superimposed as in the previous figure. The main advantage of this representation is to suppress the smoothing of the peaks due to intersubject latency jitter for the different components.

analysis of the chronometry and amplitude of the ERP profiles as a function of probability and learning. We now summarize the main features of this analysis.

A comparison of the square and circle curves in Fig. 5 showed that the $P3b$ was the first comparison to monitor the prior probability of the stimuli according to the classical law of the inverse relation between $P3b$ amplitude and stimulus probability. This result confirmed a multitude of experiments on the $P300$ probability effect (see Refs. 41 and 54–56 for a review). In particular, the inverse relation held both for cases *Fe* and *Re,* which correspond to the first practice runs of the unequal probability condition (dotted curves of Fig. 5) where little expectancy or random expectancy obtain. These relationships become more obvious when replotted as in Fig. 7(a). The inverse relation persisted during the two runs of the equal probability condition which create a condition of erroneous expectancy, as illustrated in Fig. 7(b).

These results show that $P3b$ monitors quite closely the actual probability of the stimuli and is relatively independent of prior expectancies. Independence is supported by the observation that the relationship of $P3b$ amplitude to prior probability[35] or probability shifts[8] has been shown to adapt after only a few trials. In the present experiment, the $P3b$ relationship to probability adapts during the first runs of practice and continues even after an erroneous expectancy is generated by a shift from the unequal probability condition to the equal probability condition. Yet, such independence of $P3b$ from expectancy is only relative because the inverse relationship between $P3b$ amplitude and probability is amplified by practice—that is, after the build-up of an expectancy—as can be seen by the comparison of the dotted and solid lines in Figs. 5 and 6.

By contrast with the $P3b$ component, the $N200$ was not an early index of stimulus probability. Indeed, during the first practice runs (cases *Fe* and *Re*), there was no amplitude difference in the $N200$ responses to rare and frequent stimuli at times where $P3b$ amplitude was already well differentiated [Fig. 7(a)]. Nonetheless, the $N200$ amplitude difference for high probability and low probability nontargets became as large, and in the same direction, as the $P3b$ difference by the last runs of the unequal probability conditions (cases *Fl* and *Rl*). Furthermore, $N200$ amplitude adapted slowly to stimulus probabilities, whether during the first runs of the unequal probability conditions or after the shift from unequal probabilities to equal probabilities. This lag of $N200$ suggests that this component may reflect a learning process, which develops slowly compared to the rate of $P3b$ adaptation to stimulus probabilities.

At an earlier stage of processing, a positivity ($P120$ ms) abruptly interrupts the early negativity (EN), but only in case Rl (Figs. 5 and 6). This result suggests that the $P120$ component reflects the mismatch of a learned expectancy with a rare (low probability) nontarget. A similar type of component has already been
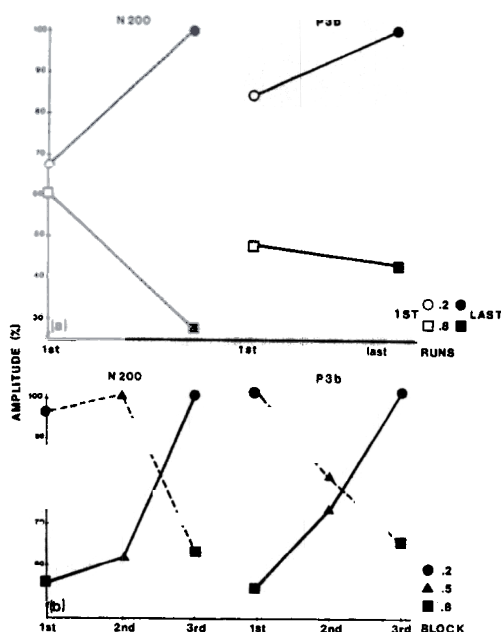
Fig. 7. Relative amplitudes of the N200 and P3b components are plotted as a function of stimulus probability and practice. (a) First runs, no practice; last runs, end of the session. (b) Relative amplitude of the N200 and P3b components are plotted as a function of stimulus probability and the relative ordering of the probability condition blocks. The transition from unequal to equal probability is accurately reflected by the P3b amplitude. The response of N200 to equal probability continues to reflect the previous unequal probability condition.

described by Desmedt et al.[12] and Goodin et al.[13] in analogous conditions.

Finally, a negative potential arises ~100 ms prior to stimulus delivery in cases Fl and Rl, but only 50 ms after stimulus delivery in cases Fe and Re (Fig. 6). We suggest that the learning of the expectancy which leads to the P120 in a mismatch situation (case Rl) is reflected by the shift in onset of this negative component with respect to the stimulus. Thus relationships among changes in components across conditions Fe, Re, Fl, and Rl provide additional constraints on the possible interpretations of each component.

XI. Comparison of ERP Profiles with Adaptive Resonance Theory Mechanisms

Adaptive resonance theory models how neural information is processed in a cyclic fashion in both the bottom-up and the top-down directions. Top-down expectancy read-out and consequent priming of an expected event can be induced in at least two ways: by instruction, as in a dichotic listening paradigm wherein the subject is asked to selectively pay attention to only one channel, and in this channel to a single target stimulus; or if there is no preselected channel and all stimuli must be attended, by practice of Bernoulli series in unequal probability conditions wherein the subject learns to expect the most frequent stimulus. Both cases result in a priming of the expected stimulus.

In the latter condition, prior to any practice there is no clearly defined expectancy about the forthcoming inputs. Thus top-down expectancy read-out can at first be triggered only by the initiation of bottom-up filtering and coding. As practice proceeds, however, top-down expectancies become progressively better learned, and hence stronger and more precise. In addition, internal representations which control these top-down expectancies may perseverate in short-term memory if they are frequently reactivated by their target stimuli. Thus, as learning proceeds, the balance between bottom-up and top-down processing may shift, such that top-down processes may become more dominant in the information processing chain. Then specific expectancies about the nature of future events may be read-out without prior bottom-up processing. Subliminal priming of F_1 by such a top-down template generates a faster supraliminal activation to a matched bottom-up input, but delays the response to a mismatched event which attenuates F_1 activation and drives the search for a better hypothesis unless the gain of F_2 is set too high to prevent easy reset. A more detailed correspondence between observed ERP components and adaptive resonance theory mechanisms will now be articulated.

We first consider the negativity which arises during cases Fl and Rl. In these cases, the negativity arises prior to the delivery of the stimulus (Fig. 6). Right after the stimulus delivery, the electrical activity profiles diverge according to whether an expectancy is either matched or mismatched. In the Fl case of expectancy match, the electrical profile exhibits only one alternation between negative and then positive activity. We suggest that the enduring poststimulus negativity includes processing negativity (PN)[10] or Nd,[24] which is typically recorded in a dichotic listening paradigm. Indeed, the plateau of activity occurs just before 100 ms when there is no P120 to interrupt it, and two peaks, prior to and after 100 ms, can be located when a P120 is generated. This early negativity (EN) therefore occurs in the time window of the PN. The PN component has properties of the match process between the subliminal template and the input code. In particular, the negativity is greater in the Fl case. According to the theory, this match process induces a supraliminal reverberant STM activity between the F_1 and F_2 levels. The subliminal priming of F_1 by a top-down template before a stimulus occurs enables F_1 to respond more rapidly to input patterns that match the primed template.[23] This property of F_1 matching is sufficient to explain the ~50-ms difference in RT between case Fl and case Rl in the target condition. It can also account for part of the smaller size of the early negative components in the no expectancy (Fe and Re) situation.

Although processing negativity of maximal amplitude is elicited by relevant target stimuli, recent results[57] also indicate that the more irrelevant stimuli resemble the relevant ones, the larger the PN they elicit, which is consistent with the 2/3 Rule (Sec. VII). Also consistent with the 2/3 Rule is the striking rever-

sal in the amplitude relations between the components EN and $N200$ across the experimental conditions Fl and Rl. After 150 ms, Rl shows a greater negativity than Fl. Indeed, 150–250 ms is the time window of the $N200$, which arises in a mismatch condition.

The long duration early negativity is followed by a minimal amplitude $N200–P300$ (Figs. 5 and 6). This is also expected from the theory. The rapid emergence of vigorous supraliminal activation at F_1 during expectancy match prevents activation of the orienting subsystem A (Sec. V) and thereby prevents STM reset of F_2. Because the $N200$ is interpreted to reflect the arousal burst, there is only a small $N200$ in this match situation. Since the $P300$ is interpreted to reflect the reset of STM at F_2, there is only a small $P300$ in the match situation.

The alternative case Rl is initiated by occurrence of a rare stimulus which leads to an expectancy mismatch. Here the prestimulus negativity and its poststimulus continuation are abruptly interrupted by an early (<50-ms onset) polarity reversal which peaks at ~120 ms. A possible interpretation for this pattern of positive activity (the $P120$) is the collapse of activity at F_1 due to input-template mismatch (Sec. V). Such an interpretation would provide strong support for the theory, because it would confirm that a mismatch between two input patterns, each of which is generated by excitatory signals, can cause less activation than the bottom-up input pattern alone due to the 2/3 Rule.

This interpretation must, however, be tested further before it is accepted without reservation. The positive polarity of $P120$ does not necessarily imply an inhibitory process. The scalp polarity of the evoked potentials depends not only on the nature of the underlying process (activation or inhibition), but also on the depth of the concerned neuronal populations. An alternative explanation is consistent with factorial analysis results which locate $P120$ and $P300$ on the same (unrotated) factor and therefore indicate a correlation between the two components: It may be that $P120$ conveys a similar function to that of $P300$. This issue is complicated further by the theory's suggestion that a causal link exists between $P120$, $N200$, and $P300$ in the situation of expectancy mismatch, as the following discussion indicates.

If the interpretation of $P120$ as a measure of mismatch at F_1 is confirmed, the next property provides even stronger support for the theory because it mirrors the theory's postulated causal link between mismatch within the attentional subsystem at F_1 and activation of the orienting subsystem at A. In the theory, mismatch causes a collapse of F_1 activation (interpret: $P120$), which thereby disinhibits the activation of A (interpret: $N200$). Thus within the situation of expectancy mismatch, one expects to find a maximal $N200$. Moreover, one expects to find that the two electrical indices $P120$ and $N200$ should covary in amplitude across experimental trials even though they are of opposite polarity. This relation is verified in Fig. 5. Therefore, the experimental results provide striking support for the hypothesis of a mismatch-mediated

burst of arousal. Further experimental tests of this correlation in other experimental paradigms which include a learning manipulation and a mismatch condition are much to be desired.

The $N200$ has a major modality-specific component $N2a$ or MMN[10,28] whose location depends on the nature of the stimulus.[49] A nonspecific component, $N2b$, peaking in the frontocentral region is added when attention is directed to the stimulus.[27,28] Both are largely overlapping in time and space. Further experimentation is needed to determine which of the two represents the arousal burst. For the orienting subsystem to work well, each F_1–F_2 processing channel must be able to calibrate its own internal mismatches by evaluating the collapse in STM activity at F_1 against the total excitatory input to its orienting subsystem A [Fig. 3(b)]. After this comparison is carried out at A to determine whether an arousal burst should be released, A can, in principle, broadcast this arousal burst to a single F_2 level or to several F_2 levels in different modalities. This is possible because the STM reset signal is nonspecific; it need not encode detailed featural properties of any code or modality. The actual distribution field of each orienting subsystem A to its target modalities also requires further experimental study.

It remains to consider how the data compare to the theoretical concept that activation of A causes a reset of STM at F_2. We compare activation of A with data about $N200$ and reset of STM at F_2 with data about $P300$. In the model, such an STM reset can initiate a rapid series of mismatch-mediated STM reset events, which constitute a search for an appropriate F_2 code (Fig. 3). This hypothesis testing scheme of the theory utilizes both the attentional and the orienting subsystems disposed in parallel, even though the search which they generate operates sequentially in time. It is well known since Sternberg's seminal experiments that each single hypothesis testing cycle is very fast, not exceeding 40 ms.[58] Each rapid reset of the F_2 level causes an enduring inhibition of the previously most active F_2 cells in order to allow for the sequential testing of new hypotheses, as well as to avoid error perservation. It is hereby suggested that, in a task requiring iterative hypothesis testing, the search process could be reflected by a rapid succession of $N200$ negative components which may merge with an incremental build up of $P300$ positivity resulting from enduring reset-contingent inhibition.

In support of this conception, it is found in many complex tasks that the reaction time (RT) may occur hundreds of milliseconds after the positivity onset. Thus several search cycles may be needed before a code capable of meeting a behavioral criterion is accessed. The masking of individual search cycles by an enduring and cumulative inhibition may partially explain this delay. Kramer et al.[59] have reported further data which are consistent with this analysis. They have shown in a varied mapping paradigm that a larger $N200$ merges with a late frontal negativity when the memory set size is increased from one to four items. It

was concluded that the prolonged negativity may reflect the need for additional controlled processing of the stimuli in the mismatch situation. In the present experiments, one expects a brief search since the memory set has only two elements.

The final event to occur, the maximal late positive complex of $P300$, has two components: $P3a$ and $P3b$. Adaptive resonance theory links two functionally different types of process to the $P300$ complex: STM reset and orienting processes. Because the theory admits multiple coding levels in addition to the simplified two-level scheme summarized herein,[21,23,43,60] one can envisage within the theory STM reset processes going on at different levels of perceptual and cognitive organization. In addition, the STM reset wave elicited by a nonspecific arousal burst from an orienting generator A may be distributed either within a sensory modality or across a wide range of intermodality processing levels. Thus one cannot infer from differences in $P300$ topography alone qualitative differences in underlying mechanism or function. Finally, Grossberg[14,42,43] pointed out that an orienting generator A may give rise to two parallel output branches. One branch, such as the one posited from A to F_2, causes reset of STM. The other branch activates processes associated with the orienting response, including gain control signals which gate the release of orienting movements. The coordinated parallel action of these two branches can activate rapid movements oriented toward an unexpected source of information as they simultaneously prepare STM to efficiently process the unexpected data. Because the interplay of these orienting and STM reset factors across all relevant processing levels may *in vivo* be complex, we content ourselves herein with qualitative comparisons between theory and data and summarize some hypotheses which are in need of further experimental tests.

Donchin et al.[61] have related $P3a$ to a brain equivalent of the somatic orienting response, which is also related to attention shifts. Our primary attention will be focused on the theoretical conception of how STM reset at F_2 (and possibly higher levels) is triggered by an arousal burst from A that is contingent on a pattern mismatch at F_1.

The major effect of the arousal burst is to inhibit the most active F_2 cells. Limited capacity STM resources are hereby freed for reallocation to less active F_2 representations. Due to this inhibition of the most active sources of top-down template signals, the mismatch at F_1 is eliminated [Fig. 3(c)]. The F_1 reaction to the bottom-up input pattern is hereby unmasked and begins once again to activate the $F_1 \rightarrow F_2$ adaptive filter [Fig. 3(d)].

As this is happening, a complementary effect of the arousal burst begins to take effect. This effect is the STM enhancement (dishabituation, unblocking) of F_2 representations which previously were only weakly activated. This unblocking effect may be intuitively understood as follows. The previously attenuated STM representations may have been encoding important information which was erroneously unattended,

thereby leading to activation of the wrong hypothesis and read-out of the wrong expectation. Unblocking rectifies this error by endowing these STM representations with large activation levels. Unblocking begins to occur as the unmasked inputs from F_1 begin to influence F_2. The combination of unblocking at F_2 and inputs from F_1 generates a new pattern of activity across F_2; that is, a new code, or hypothesis, is instated in STM at F_2. Thus reset of STM includes two types of operation, those which are modulated by learned top-down expectancies and those which are directly induced by bottom-up processing of the input. These functional properties are consistent with the amplitude response of the $P3a$ and $P3b$ components [Fig. 7(a)], which shows a tendency to increase in response to rare events across learning trials (top-down effect), as well as base line differences between responses to frequent and rare events before learning develops (bottom-up effect).

As noted above, the inhibitory effect of the arousal burst at F_2 is enduring, so that when F_1 input again activates F_2, perseveration of the old erroneous hypothesis is prevented. This enduring inhibition persists and accumulates when a series of mismatch-mediated arousal bursts develops in a task that triggers several hypothesis testing cycles. Such hypothesis testing cycles may generate a sustained F_2 inhibition, or positivity, superimposed on more momentary activations, or negativities, thereby generating longer late positive complexes. A different but possibly related situation is created by increasing the difficulty of a discrimination task. In this case Ruchkin et al.[62-65] have shown an increased amplitude of the slow wave component. This last component is made of simultaneous positive and negative activity, each predominant at different locations. In summary, the STM reset properties within a cortical F_2 field, including its activity-dependent sustained inhibition mechanism, suggest a physiological model of how properties of $P3b$ may reflect contextual updating.

According to recent ERP results relating $P300$ and memory, during certain strategies of memorization the amplitude of the $P300$ component at the moment of the stimulus delivery is related to the strength of the consecutive LTM trace.[44] Such a result does not, however, imply that $P300$ is a direct reflection of the long-term memorization process. Grossberg[14,66] has, for example, noted that an unexpected event can be stored in STM with amplified activity by inheriting limited capacity STM resources from the F_2 representations which it has just reset. These larger STM values can generate larger learning signals. Larger learning signals support faster encoding into LTM. Thus the relationship between $P300$ amplitude and subsequent LTM strength may be mediated by properties of the STM reset process which is hypothesized to be a major cause of the $P300$ components.

Adaptive resonance theory has suggested several detailed neural network models which contain candidates for $P3a$ and/or $P3b$ generators. The theory admits several processes that are candidates for $P3a$

generators. Further experiments are needed to decide between them. One possibility is that a P300 (possibly a P3a) occurs when the orienting subsystem activates orienting responses, as it inhibits the midbrain reinforcement circuits which motivate consummatory motor commands.[42,43] A second possibility is that a P300 is generated when STM reset at F_2 indirectly causes an STM reset at midbrain reinforcement circuits.[14,66] Such a secondary reset can disconfirm the motivational bias that had been set by the erroneous cognitive representation. This type of STM reset was predicted to involve circuits which include the hippocampus.[20,42] In partial support of this prediction, P300 activity has been recorded in the hippocampus[67-69] or indirectly located in this structure.[70] To further clarify this situation, learning experiments capable of dissociating cortical and hippocampal generators have been suggested.[15,66]

It is also worth emphasizing why, even in the absence of an experimentally trained expectancy, P300 amplitude reflects in real time the difference in stimulus probability and N200 does not. One explanation of this derives from the nature of the direct activation of higher processing levels by lower processing levels in the theory. Bottom-up STM encoding in the theory is sensitive to the frequency and temporal ordering of individual events.[21,23,60,71,72] Therefore, even if a top-down template does not yet differentiate individual events, STM reset due to direct bottom-up activation can reflect the prior probability of events. In particular, less reset of a frequent event may be expected due to the stronger STM preseveration of that event as a result of its many previous occurrences.

Although bottom-up activation of certain processing levels is sensitive to event frequency and even to temporal order, both bottom-up code learning and top-down template learning can influence the form of ERPs through time as events become more familiar. In particular, results reported elsewhere[4] demonstrate the influence of long-term memorization on the P3b amplitude response. The amplitude of this component increases after long-term (1-week interval) learning, reflecting the learning of new top-down templates. Simultaneously, the N200 amplitudes to rare and frequent stimuli become less different.[7] These results illustrate that a classification of ERPs in terms of their different topographies alone provides an insufficient measure of their functional independence, since direct bottom-up activation of STM and top-down read-out from LTM into STM can converge on the same cell targets and can vary in different ways through time as a function of the experimental task.

The learning-dependent N200 changes discovered by Banquet et al.[4] and Banquet and Guenther[73] can be explained using the following concepts. The top-down template of the frequent stimulus may be learned relatively rapidly, whereas the template corresponding to the rare stimulus may be learned at a slower rate. When the frequent stimulus template alone is active, sharp matches or mismatches would be expected with the frequent and rare events, respective-

ly. As the template of the rare stimulus is progressively but more slowly learned, the total top-down template becomes a composite of frequent and rare event templates. In this later phase of learning, pure matches or mismatches are replaced by partial matches or mismatches, due to the presence of both template components. This analysis is consistent with the existence of two learning phases, and with the fact that after long-term learning, the N200 to rare and frequent stimuli becomes less different in amplitude. In addition, the hypothesized difference in the rate of template learning in response to frequent and rare stimuli is supported by the behavioral results which show no significant difference in RT at the beginning of the first session, while a large difference occurs by the end of the first session. This difference results from a decreased RT to frequent events and an increased RT to rare events which is consistent with theoretical properties of matching both frequent and rare events against a learned template for the frequent event. Thus, at least during the first session, most of the learning occurs in response to the frequent event.

XII. Conclusion: the Relationship of Learning to ERPs

The concepts and mechanisms of adaptive resonance theory which are most important for the analysis of ERPs—learning of top-down templates, matching of bottom-up input patterns with learned top-down templates, frequency-sensitive bottom-up encoding of events in STM, activation of the orienting subsystem, and reset of event codes in STM—were all derived from an analysis of how a cognitive system can learn recognition codes in a self-stabilizing and globally self-consistent fashion. This theoretical framework has, by now, been useful for analyzing and predicting data in a number of fields, such as visual perception,[2,71,74,75] classical and instrumental conditioning,[15,21,66,76-78] speech processing,[21,22,60,71,72,79] word recognition and recall,[21-23,71] decision making under risk,[80] and self-organization of cognitive recognition codes,[15-19,21] and has suggested neural principles and mechanisms for interpreting and sharpening many concepts within the ERP literature. In particular, the relationship of adaptive resonance theory concepts to the concepts of Donchin about P300 are reviewed in Ref. 14 and reprinted in Ref. 15 and to those of Naatanen about PN and N200 are reviewed in Ref. 17.

Despite the critical role of learning constraints on the design of cognitive mechanisms, much of the ERP literature has utilized performance paradigms to analyze individual ERPs. It seems to us that an informative way to understand the role of learning constraints on the cognitive designs probed by ERPs is to investigate paradigms in which correlated changes in amplitudes and chrometric relationships among several ERP components as a function of learning are the units of the experimental and theoretical analysis. The present article contributes to this enterprise.

Due to the fact that only a small number of ERP experiments have explicitly tested how multiple ERPs covary as a function of learning manipulations, many

more experimental studies of such correlations will be needed before ART mechanisms are unequivocally supported in many behavioral situations. It is also to be expected that such data will provide useful guidelines for further theoretical development and refinement. On the other hand, the facts that known ART mechanisms have predicted both the existence of key ERPs and their main correlations in available data provide a hopeful beginning for such a systematic analysis, as well as a serious challenge to alternative cognitive theories in which these ERP data have no natural interpretation.

References

1. S. Grossberg, "Adaptive Pattern Classification and Universal Recoding I: Parallel Development and Coding of Neural Feature Detectors," Biol. Cybern. 23, 121 (1976).
2. S. Grossberg, "Adaptive Pattern Classification and Universal Recoding II: Feedback, Expectation, Olfaction, and Illusions," Biol. Cybern. 23, 187 (1976).
3. J.-P. Banquet, B. Renault, and N. Lesevre, "Effect of Task and Stimulus Probability on Evoked Potentials," Biol. Psychol. 13, 203 (1981).
4. J.-P. Banquet, J. Baribeau-Braun, and N. Lesevre, "Learning of "Single Trial" and "Contextual" Information Processing in an Odd-Ball Paradigm," in Brain and Information: Event Related Potentials, R. Karrer, J. Cohen, and P. Tueting, Eds. (New York Academy of Sciences, 1984), pp. 162–165.
5. J.-P. Banquet, F. El Massioui, and J. L. Godet, "ERP-RT Chronometry and Learning in Normal and Depressed Subjects," in Cerebral Psychophysiology: Studies in Event-Related Potentials, W. C. McCallum, R. Zappoli, and F. Denoth, Eds. (Elsevier, Amsterdam, 1986).
6. J.-P. Banquet, W. Guenther, and M. Smith, "Probability Processing in Depressed Patients," In Current Trends in Event-Related Potential Research, R. Johnson, Jr., R. Parasuraman, and J. W. Rohrbaugh, Eds. Current Electroencephalogr. Clin. Neurophysiol. Suppl. 40, (1987), in press.
7. J.-P. Banquet, W. Guenther, and M. Smith, "Probability Mapping, Task Performance and Learning: an ERP Model," (1987), submitted for publication.
8. R. Johnson, Jr., and E. Donchin, "Sequential Expectancies and Decision Making in a Changing Environment: an Electrophysiological Approach," Psychophysiology 19, 183 (1982).
9. K. C. Squires, C. Wickens, N. K. Squires, and E. Donchin, "The Effect of Stimulus Sequence on the Waveform of the Cortical Event-Related Potentials," Science 193, 1142 (1976).
10. R. Naatanen, A. W. Gaillard, and S. Mantysalo, "Early Selective Attention Effect on Evoked Potential Reinterpreted," Acta Psychol. 42, 313 (1978).
11. J. C. Hansen and S. A. Hillyard, "Endogenous Brain Potentials Associated with Selective Auditory Attention," Electroencephalogr. Clin. Neurophysiol. 49, 277 (1980).
12. J. E. Desmedt, N. Tran Huy, and M. Bourguet, "The Cognitive P40, N60, and P100 Components of Somato-Sensory Evoked Potentials and the Earliest Electrical Signs of Sensory Processing in Man," Electroencephalogr. Clin. Neurophysiol. 56, 272 (1983).
13. D. S. Goodin, K. C. Squires, B. H. Henderson, and A. Starr, "An Early Event-Related Cortical Potential," Psychophysiology 15, 360 (1978).
14. S. Grossberg, "Some Psychophysiological and Pharmacological Correlates of a Developmental, Cognitive, and Motivational Theory," in Brain and Information: Event Related Potentials, R. Karrer, J. Cohen, and P. Tueting, Eds. (New York Academy of Sciences, 1984).
15. S. Grossberg, Ed., The Adaptive Brain I: Cognition, Learning, Reinforcement, and Rhythm (Elsevier/North-Holland, Amsterdam, 1987).
16. G. A. Carpenter and S. Grossberg, "Neural Dynamics of Category Learning and Recognition: Attention, Memory Consolidation, and Amnesia," in Brain Structure, Learning, and Memory, J. Davis, R. Newburgh, and E. Wegman, Eds. (AAAS Symposium Series, 1987), in press.
17. G. A. Carpenter and S. Grossberg, "Neural Dynamics of Category Learning and Recognition: Structural Invariants, Reinforcement, and Evoked Potentials," in Pattern Recognition and Concepts in Animals, People, and Machines, M. L. Commons, S. M. Kosslyn, and R. J. Herrnstein, Eds. (Erlbaum, Hillsdale, NJ, 1987).
18. G. A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for a Self-Organizing Neural Pattern Recognition Machine," Comput. Vision Graphics Image Process. 37, 54 (1987).
19. G. A. Carpenter and S. Grossberg, "ART 2: Self-Organization of Stable Category Recognition Codes for Analog Input Patterns," Appl. Opt. 26, 4919 (1987).
20. S. Grossberg, "How Does a Brain Build a Cognitive Code?," Psychol. Rev. 87, 1 (1980).
21. S. Grossberg, Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition, and Motor Control (Reidel, Dordrecht, 1982).
22. S. Grossberg, "The Adaptive Self-Organization of Serial Order in Behavior: Speech, Language, and Motor Control," in Pattern Recognition by Humans and Machines, Vol. 1, Speech Perception, E. C. Schwab and H. C. Nusbaum, Eds. (Academic, New York, 1986), pp. 187–294.
23. S. Grossberg and G. O. Stone, "Neural Dynamics of Word Recognition and Recall: Attentional Priming, Learning, and Resonance," Psychol. Rev. 93, 46 (1986).
24. S. A. Hillyard, R. F. Hink, V. L. Schwent, and T. W. Picton, "Electrical Signs of Selective Attention in the Human Brain," Science 182, 177 (1973).
25. D. E. Broadbent, "Stimulus Set and Response Set: Two Kinds of Selective Attention," in Attention: Contemporary Theory and Analysis, D. Mostofsky, Ed. (Appleton-Century-Crofts, New York, 1970), pp. 51–60.
26. R. Naatanen and T. W. Picton, "N2 and Automatic Versus Controlled Processes," in Cerebral Psychophysiology: Studies in Event-Related Potentials, W. C. McCallum, R. Zappoli, and F. Denoth, Eds. (Elsevier, Amsterdam, 1986).
27. B. Renault and N. Lesevre, "Topographical Study of the Emitted Potential Obtained After the Omission of an Expected Visual Stimulus," in Multidisciplinary Perspectives in Event-Related Brain Potential Research, D. Otto, Ed. (U.S. Government Printing Office, Washington, DC, 1987), pp. 202–208.
28. R. Naatanen, M. Simpson, and N. E. Loveless, "Stimulus Deviance and Evoked Potentials," Biol. Psychiatry 14, 53 (1982).
29. R. Naatanen and A. W. Gaillard, "The Orienting Reflex and the N2 Deflection of the ERPs," in Tutorials in Event-Related Potential Research: Endogenous Components, A. W. K. Gaillard and W. Ritter, Eds. (North-Holland, Amsterdam, 1983), pp. 119–142.
30. S. Sutton, M. Braren, J. Zubin, and E. R. John, "Evoked Poten-

tial Correlates of Uncertainty," Science 150, 1187 (1965).

31. E. Donchin, "Surprise!...Surprise?," Psychophysiology 18, 493 (1981).

32. N. K. Squires, K. C. Squires, and S. A. Hillyard, "Two Varieties of Long-Latency Positive Waves Evoked by Unpredictable Auditory Stimuli in Man," Electroencephalogr. Clin. Neurophysiol. 38, 387 (1975).

33. E. Courchesne, S. A. Hillyard, and R. Galambos, "Stimulus Novelty, Task Relevance, and the Visual Evoked Potential in Man," Electroencephalogr. Clin. Neurophysiol. 39, 131 (1975).

34. E. Courchesne, "Changes in P3 Waves with Event Repetition: Long-Term Effects of Scalp Distribution and Amplitude," Electroencephalogr. Clin. Neurophysiol. 45, 754 (1978).

35. W. Ritter, H. G. Vaughan, Jr., and L. D. Costa, "Orienting and Habituation to Auditory Stimuli: A Study of Short Term Changes in Average Evoked Responses," Electroencephalogr. Clin. Neurophysiol. 25, 550 (1968).

36. W. T. Roth, "Auditory Evoked Responses to Unpredictable Stimuli," Psychophysiology 10, 125 (1973).

37. W. T. Roth and B. S. Kopell, "P300—An Orienting Reaction in the Human Auditory Evoked Response," Percept. Mot. Skills 36, 219 (1973).

38. E. Snyder and S. A. Hillyard, "Long-Latency Evoked Potentials to Irrelevant, Deviant Stimuli," Behav. Biol. 16, 319 (1976).

39. R. Munson, D. S. Ruchkin, W. Ritter, S. Sutton, and N. K. Squires, "The Relation of P3b to Prior Events and Future Behavior," Biol. Psychol. 19, 1 (1984).

40. R. Johnson, Jr., "Triarchic Model of P300 Amplitude," Psychophysiology 23, 367 (1986).

41. E. Donchin, W. Ritter, and W. C. McCallum, "Cognitive Psychophysiology: the Endogenous Components of the ERP," in Brain Event-Related Potentials in Man, E. Callaway, P. Tueting, and S. Koslov, Eds. (Academic, New York, 1978), pp. 349–441.

42. S. Grossberg, "A Neural Model of Attention, Reinforcement, and Discrimination Learning," Int. Rev. Neurobiol. 18, 263 (1975).

43. S. Grossberg, "A Theory of Human Memory: Self-Organization and Performance of Sensory-Motor Codes, Maps, and Plans," Prog. Theor. Biol. 5, 233 (1978).

44. D. Karis, M. Fabiani, and E. Donchin, "P300 and Memory: Individual Differences in the von Restorff Effect," Cognitive Psychol. 16, 177 (1984).

45. E. Donchin and E. Heffley, "Multivariate Analysis of Event-Related Potential Data: A Tutorial Review," in Multidisciplinary Perspectives in Event-Related Potential Research, D. A. Otto, Ed. (U.S. Government Printing Office, Washington, DC, 1978), pp. 552–572.

46. E. Donchin and J. B. Isreal, "Event Related Brain Potentials and Psychological Theory," in Motivation, Motor, and Sensory Processes of the Brain: Progress in Brain Research, H. H. Kornhuber and L. Deecke, Eds. (Elsevier, Amsterdam, 1980), pp. 697–715.

47. J.-P. Banquet and N. Lesevre, "Event-Related Potentials in Altered States of Consciousness," in Motivation, Motor, and Sensory Processes of the Brain: Progress in Brain Research, H. H. Kornhuber and L. Deecke, Eds. (Elsevier, Amsterdam, 1980), pp. 447–453.

48. W. Ritter, R. Simson, G. Herbert, H. G. Vaughan, Jr., and D. Friedman, "A Brain Event Related to Making a Sensory Discrimination," Science 203, 1358 (1979).

49. R. Simson, H. G. Vaughan, and W. Ritter, "Scalp Topography of Potentials in Auditory and Visual Discrimination Tasks," Electroencephalogr. Clin. Neurophysiol. 42, 528 (1977).

50. W. Ritter, R. Simson, H. G. Vaughan, Jr., and M. Macht, "Manipulation of Event-Related Potentials Manifestation of Information Processing Stages," Science 218, 909 (1982).

51. R. Naatanen, "Processing Negativity: an Evoked-Potential Reflection of Selective Attention," Psychol. Bull. 92, 605 (1982).

52. S. H. Curry, R. Cooper, W. C. McCallum, P. V. Popock, D. Papakostopoulos, S. Skidmore, and P. Newton, "The Principal Components of Auditory Target Detection," in Tutorials in Event-Related Potential Research: Endogenous Components, A. W. K. Gaillard and W. Ritter, Eds. (North-Holland, Amsterdam, 1983), pp. 79–117.

53. F. Rosler and D. Manzey, "Principal Components and Varimax-Rotated Components in Event-Related Potentials Research: Some Remarks on Their Interpretation," Biol. Psychol. 13, 3 (1981).

54. C. C. Duncan-Johnson and E. Donchin, "On Quantifying Surprise: the Variation in Event-Related Potentials with Subjective Probability," Psychophysiology 14, 456 (1977).

55. P. Tueting, S. Sutton, and J. Zubin, "Quantitative Evoked Potential Correlates of the Probability of Events," Psychophysiology 7, 385 (1971).

56. W. S. Pritchard, "Psychophysiology of P300," Psychol. Bull. 89, 506 (1981).

57. K. Alho, M. Sams, P. Paavilainen, and R. Naatanen, "Small Pitch Separation and the Selective-Attention Effect on the ERP," Psychophysiology 23, 189 (1986).

58. S. Sternberg, "High Speed Scanning in Human Memory," Science 153, 652 (1966).

59. A. Kramer, W. Schneider, A. Fisk, and E. Donchin, "The Effects of Practice and Task Structure on Components of the Event-Related Brain Potential," Psychophysiology 23, 33 (1986).

60. M. A. Cohen and S. Grossberg, "Neural Dynamics of Speech and Language Coding: Developmental Programs, Perceptual Grouping, and Competition for Short Term Memory," Hum. Neurobiol. 5, 1 (1986).

61. E. Donchin et al., "The Orienting Reflex and P300," in Brain and Information: Event-Related Potentials, R. Karrer, J. Cohen, and P. Tueting, Eds. (New York Academy of Sciences, 1984).

62. D. S. Ruchkin, R. Munson, and S. Sutton, "P300 and Slow Wave in a Message Consisting of Two Events," Psychophysiology 19, 629 (1982).

63. D. S. Ruchkin and S. Sutton, "Positive Slow Wave and P300: Association and Dissociation," in Tutorials in Event-Related Potential Research: Endogenous Components, A. W. K. Gaillard and W. Ritter, Eds. (North-Holland, Amsterdam, 1983), pp. 233–250.

64. D. S. Ruchkin, S. Sutton, M. L. Kietzman, and K. Silver, "Slow Wave and P300 in Signal Detection," Electroencephalogr. Clin. Neurophysiol. 50, 35 (1980).

65. D. S. Ruchkin, S. Sutton, and M. Stega, "Emitted P300 and Slow Wave Event-Related Potentials in Guessing and Detection Tasks," Electroencephalogr. Clin. Neurophysiol. 49, 1 (1980).

66. S. Grossberg, "Processing of Expected and Unexpected Events During Conditioning and Attention: a Psychophysiological Theory," Psychol. Rev. 89, 529 (1982).

67. E. Halgren, N. K. Squires, C. L. Wilson, J. W. Rohrbaugh, T. L. Babb, and P. H. Crandall, "Endogenous Potentials Generated in the Human Hippocampal Formation and Amygdala by Infrequent Events," Science 210, 803 (1980).

68. N. K. Squires, E. Halgren, C. Wilson, and P. Crandall, "Human Endogenous Limbic Potentials: Cross-Modality and Depth-Surface Comparison in Epileptic Subjects," in Tutorials in Event-Related Potential Research: Endogenous Components, A. W. K. Gaillard and W. Ritter, Eds. (North-Holland, Amsterdam, 1983), pp. 217–232.

69. M. Wood, G. McCarthy, N. K. Squires, H. G. Vaughan, D. L. Woods, and W. C. McCallum, "Anatomical and Physiological Substrates of Event-Related Potentials," in Brain and Information: Event-Related Potentials, R. Karrer, J. Cohen, and P. Tueting, Eds. (New York Academy of Sciences, 1984).

70. Y. C. Okada, L. Kaufman, and S. J. Williamson, "The Hippocampal Formation as a Source of the Slow Endogenous Potentials," Electroencephalogr. Clin. Neurophysiol. 55, 417 (1983).

71. S. Grossberg, Ed., The Adaptive Brain II: Vision, Speech, Language, and Motor Control (Elsevier/North-Holland, Amsterdam, 1987).

72. S. Grossberg and G. O. Stone, "Neural Dynamics of Attention Switching and Temporal Order Information in Short Term Memory," Mem. Cognit. 14, 451 (1986).

73. J.-P. Banquet and W. Guenther, "Intuitive Statistics and Related Memory Models," in Cognitiva 85, Cesta-Afcet, Ed. (publisher, Paris, 1985), pp. 49–56.

74. S. Grossberg, "Cortical Dynamics of Three-Dimensional Form, Color, and Brightness Perception: I. Monocular Theory," Percept. Psychophys. 41, 87 (1987).

75. S. Grossberg and E. Mingolla, "Neural Dynamics of Perceptual Grouping: Textures, Boundaries, and Emergent Segmentations," Percept. Psychophys. 38, 141 (1985).

76. M. A. Cohen and S. Grossberg, "Unitized Recognition Codes for Parts and Wholes: the Unique Cue in Configural Discriminations," in Pattern Recognition and Concepts in Animals, People, and Machines, M. L. Commons, S. M. Kosslyn, and R. J. Herrnstein, Eds. (Erlbaum, Hillsdale, NJ, 1987).

77. S. Grossberg and D. S. Levine, "Neural Dynamics of Attentionally Modulated Pavlovian Conditioning: Blocking, Interstimulus Interval, and Secondary Reinforcement," Appl. Opt. 26, 5015 (1987).

78. S. Grossberg and N. Schmajuk, "Neural Dynamics of Attentionally-Modulated Pavlovian Conditioning: Conditioned Reinforcement, Inhibition, and Opponent Processing," Psychobiology (1987), in press.

79. M. A. Cohen and S. Grossberg, "Masking Fields: a Massively Parallel Neural Architecture for Learning, Recognizing, and Predicting Multiple Groupings of Patterned Data," Appl. Opt. 26, 1866 (1987).

80. S. Grossberg and W. Gutowski, "Neural Dynamics of Decision Making Under Risk: Affective Balance and Cognitive-Emotional Interactions," Psychol. Rev. 94, 300 (1987), in press.