Acta Psychologica 67 (1988) 191-257
North-Holland

Invited paper

DISCRETE AND CONTINUOUS MODELS OF HUMAN INFORMATION PROCESSING: THEORETICAL DISTINCTIONS AND EMPIRICAL RESULTS *

Jeff MILLER

University of California, San Diego, USA

Information-processing models of human performance are often categorized as discrete or continuous, with continuous models currently in vogue. This paper presents a three-part analysis of discrete and continuous models. Part 1 considers the definitions of the terms discrete and continuous, and concludes that a dichotomous categorization of models is a serious oversimplification for three reasons. First, discreteness and continuity are not themselves dichotomous properties, but rather two extremes between which there can be intermediate cases. Second, there are many different senses in which information-processing models can be described as discrete or continuous (or somewhere in between), and the same model may be relatively discrete in some senses and relatively continuous in others. Third, any complete model of an information-processing task assumes several different processing stages, of which some could be relatively discrete (in any given sense) and others relatively continuous. Part 2 reviews evidence commonly cited in support of continuous models. Many of the findings can be reconciled with fully discrete models, and the rest require continuity only in certain very limited senses. In particular, there is no compelling evidence against the discrete stage assumption of the Additive Factor Method (AFM) of Sternberg (1969a, b). Part 3 reviews recent evidence supporting discrete models, some of which specifically supports the discrete stage assumption. In the light of this evidence and the shortcomings of the evidence reviewed in Part 2, it is concluded that it is premature to abandon discrete models in favor of continuous ones.

Over the last 30 years, the information-processing approach has come to dominate cognitive psychologists' attempts to understand how organisms acquire and use knowledge about the world (J.R. Anderson 1980). Information-processing models have been developed to explain

---

* This research was supported by grant PHS-MH40733 from the National Institute of Mental Health. The first draft of this article was prepared while the author was a visitor in the Human Information Processing Research Dept. of Bell Laboratories, at Murray Hill, NJ.

I would like to thank David Bauer, Patricia Haden, James Johnston, Harold Pashler and Andries Sanders for helpful discussions and comments on the ideas and their exposition.

Author's address: J. Miller, Dept. of Psychology, C-009, University of California, San Diego, La Jolla, CA 92093, USA.

performance in myriad tasks, including perceptual discrimination (e.g., Vickers 1970, 1979), perceptual comparison (e.g., Posner 1978), reading (Just and Carpenter 1980), problem solving (e.g., Newell and Simon 1972), making decisions based on mental images (Kosslyn 1981), fact retrieval (J.R. Anderson and Bower 1980), and typing (Sternberg et al. 1978). Much of the success of the information-processing approach results from the generality of its form, nicely expressed by Bower:

> 'The information-processing approach assumes that perception and learning can be analyzed conceptually into a series of stages during which particular components ... perform certain transformations or recodings of the information coming into them. The subject's eventual response ... is considered to be the outcome of this lengthy series of operations. Each stage in the system receives as input the information as coded in its predecessor stage, operates upon it so as to condense, abstract, recode, or elaborate it, and then passes this product along to the next stage in the analysis. Since external stimuli cannot get inside an organism, the representation of them ... and their interactions ... is what we call "information", and this is the content we describe in our theories.' (Bower 1975: 33)

The information-processing approach encompasses an enormous class of models, because, in this very general form, it places no constraints on the system's representations, transformations, or component processing stages. [1] Indeed, the job of cognitive psychologists is to discover these constraints. Many specific models incorporating particular sets of constraints have been proposed within the approach, including information transmission through a single channel of limited capacity (Broadbent 1958; Hick 1952), interactive activation networks (Rumelhart and McClelland 1982), production systems (Newell 1973), serial stage systems (Sternberg 1969a, b), and hierarchical filters (Estes 1977). No one or combination of these models has gained overwhelming support, however, and many types of models are still under active investigation.

In recent years a number of researchers have distinguished between two broad categories of information-processing models known as *discrete* and *continuous* (e.g., Eriksen and Schultz 1979; McClelland 1979; Meyer et al. 1984, 1985; Miller 1982a, b, 1983; Norman and Bobrow 1975). In particular, it has become fashionable to reject discrete models as inconsistent with a great deal of behavioral evidence, physiologically

---

[1] It is standard in the literature to use the term 'stage' to refer to the major divisions or 'functional subunits' (Taylor 1976) in the processing needed to carry out a task. The term 'process' is commonly used to refer to the more specific operations of which the stages are composed (cf. Sanders 1980).

implausible, and just plain old-fashioned. Perhaps as a result of this trend, many empirical papers now include a continuous model that can account for their results, but do not consider how a discrete model might do so. Furthermore, the Additive Factor Method (AFM) of Sternberg (1969a, b), once considered a formidable tool in the exploration of cognition, seems now to be regarded as untenable because it assumes a serial succession of discrete processing stages.

We consider the *en masse* abandonment of discrete models in favor of continuous ones to be wholely unjustified given the evidence currently available, and thus scientifically premature. We believe that it is still very much an open question whether discrete or continuous models better characterize human information processing. In support of this position, we offer a three-part examination of the distinction between discrete and continuous models.

The first part considers in detail just what is meant by the terms discrete and continuous. It is suggested that an overarching binary distinction between discrete and continuous information-processing models is a tremendous oversimplification, because discreteness and continuity do not divide the models into two mutually exclusive and exhaustive classes. To support this thesis, it is argued that: (a) discreteness and continuity do not constitute a dichotomy, but rather two ends of a graded dimension with intermediates between the extremes, (b) there are at least three different senses in which information-processing models can be relatively discrete or relatively continuous, and models can be discrete in some senses but continuous in others, and (c) information-processing models involve several distinct processing stages, and different stages within a single model need not all be discrete or continuous, even within a single sense of that distinction.

The second part presents a critical review of evidence that has caused theorists to reject discrete models in favor of continuous ones. It is concluded that there is no decisive evidence in favor of fully continuous models, contrary to popular belief. In particular, most of the evidence cited in support of continuous models supports continuity only in a very limited sense, and this sense of continuity is compatible with models that are discrete in other important senses. In this part, special emphasis is given to evidence regarded as incompatible with the 'discrete processing stage' assumption that total processing time can be decomposed into a sum of times needed for component stages. This assumption is of special importance, because it underlies many discrete stage models and the AFM. It is argued that there is no persuasive evidence of continuity in the sense incompatible with the discrete stage assumption.

The third part presents a review of evidence in favor of discrete models. In some important senses of the terms, there seems to be more evidence for discrete models than continuous ones. This evidence has generally been ignored in the rush toward continuous models, and in some cases has even been misunderstood as support for those models.

Part 1:
On the meaning of 'discrete' and 'continuous'

1.1. The continuum of grain size

Before discussing the applications of discreteness and continuity to information-processing models, it is useful to review the dichotomy to which these terms refer in mathematics and statistics: a variable is discrete if its values cannot be arbitrarily similar to one another and continuous if they can. For example, population size of the U.S. is discrete, since values differ by a minimum unit of 1. Human height is continuous, since its values can differ by arbitrarily small amounts (i.e., no minimum unit).

Though the mathematical distinction between discreteness and continuity can be useful in comparing certain classes of models (e.g., Van Santen and Bamber 1981), the distinction between variables with and without minimum units will not always be the most useful distinction for cognitive psychologists. For one thing, we will often want to distinguish a variable with a small minimum unit from one with a large minimum unit, even though the mathematical distinction classifies both as discrete. For example, a system that represented perceived brightness with two or three or even ten discrete levels would be quite different, in most tasks, from one that represented perceived brightness in terms of minimum units corresponding to individual photoreceptors stimulated. For another thing, we will often wish to avoid distinguishing between variables with minimum units of zero and those with small but non-zero units, because either kind of variable can be used to model the other without any important loss of accuracy. In almost all perceptual analyses, for example, perceived brightness could be treated as a continuous variable, even if it actually varied in discrete units corresponding to single photoreceptors.

The example of brightness representation makes it clear that no dichotomy is sufficient to describe the minimum unit of a variable. If we start with a two-state model of brightness representation and gradually increase the number of brightness levels that can be represented, there is no magical cutoff number at which the brightness representation becomes continuous rather than discrete. The model of brightness representation can only be described accurately by stating the number of levels or the minimum difference between levels.

In practical terms, it makes sense to regard discreteness and continuity not as a dichotomy but rather as the ends of a quantitative dimension that might be called the 'grain size' dimension. The grain size of a variable is its minimum unit of change. A variable is more continuous to the extent that it has a small grain size and more discrete to the extent that it has a large one.

Both extremes of discreteness and continuity can be specified in terms of grain size. At the continuous extreme is a minimum unit of zero, so that no smaller grain size is possible. At the discrete extreme is a minimum unit that covers the entire range of variation, so that no value lies between two other values. For ordered variables, extreme discreteness means that the variable is binary; the minimum unit is the same as the range of variation. Unordered categorical variables also lie at the extreme of discreteness, because their range of variation cannot be subdivided into smaller minimum units in any meaningful way.

With discrete and continuous variables defined as the extremes of the grain-size dimension, intermediate variables are those for which grain size is neither very small nor very large relative to the range of interest.

In most cases it will be necessary to consider the grain size of a variable relative to the absolute size and range of values under consideration. For example, if the minimum unit of perceived brightness did correspond to the activation of an individual photoreceptor, then the grain size of brightness judgments would be very small relative to everyday ranges of illumination, but it might be considered relatively large (i.e., large enough to merit the label 'discrete') in certain psychophysical experiments with very low illuminations.

A further complication is that the grain size of a variable may not be constant across its range. Internal representations of brightness, for example, might vary in step sizes proportional to the absolute level of brightness, corresponding to Weber's law (Thurstone 1959). Thresholding is an even more extreme example; the representation of a thresholded variable could be fully discrete below the threshold and fully continuous above it.

In summary, it makes sense to apply the terms 'discrete' and 'continuous' to variables with very large or small grain sizes. Therefore, in this article we will use these terms as a short-hand for 'relatively discrete' and 'relatively continuous'. It is necessary to remember that this terminology conceals important complexities, however. For example, besides the intermediate possibilities, we have seen that the same variable may be continuous when considered over a large range of its values, but discrete when considered over a small range.

1.2. Three different senses of 'discrete' and 'continuous'

The terms 'discrete' and 'continuous' can be used in at least three different senses in describing the stages involved in information-processing models. It is important to distinguish these senses, because stages can be discrete in some and continuous in others. Furthermore, as argued in Parts 2 and 3, data may suggest discreteness or continuity in one sense without implying anything about the other senses, and failure to distinguish among the different senses can lead to misinterpretation of experimental results.

In the spirit of Bower's (1975) description, we will characterize information-processing models as composed of *representations* and *stages* (cf. J.R. Anderson 1980; Broadbent 1958; Neisser 1967; Sanders 1977, 1980; Simon 1969; Sternberg 1969a, b; Taylor 1976; Underwood 1978). In the class of models under discussion, representations are passive codes that hold information, and stages are active operators that use and modify representations to accomplish a more or less specific transformation.

A given information-processing model is typically composed of several distinct stages. These stages accomplish different transformations and may be influenced by different experimental factors. Each stage receives an input representation, performs some transformation on that representation, and produces an output representation. Although stages may use internal codes for holding information as they carry out their transformations, we will regard such local codes as characteristics of the transformations rather than as yet another type of information representation. This simplification does not obscure any important features of models, because the forms of the internal codes are determined by the rules governing the transformation (i.e., once one has specified the transformation, no additional information about the model is conveyed by describing these local codes). For example, if a mental rotation stage continuously transforms its input representation to upright, then the stage must have an internal code that is continuously varying. Likewise, if the stage were to transform its input representation discretely, then it adds nothing to say that within the stage orientation was represented discretely.

Within a model, stages are often ordered so that the output representation produced by one stage serves as the input representation to another stage, and representations can be described as being transmitted from one stage to the next. Simple examples of such models arise in the context of traditional choice reaction tasks (e.g., Sanders 1977, 1980; Smith 1968). For example, it may be assumed that an initial stage recognizes the stimulus. An output representation of stimulus identity is then transmitted to a decision stage, which selects the appropriate response. A representation of the response is then transmitted to the motor system, which activates the effectors needed to make that response. These stages must operate in order, because the decision is contingent upon the stimulus, and the response is contingent upon the decision.

The above characterization of information-processing models in terms of representations and stages is fairly standard, though unfortunately not very rigorous. Formal definitions of mental representations and stages might be very valuable in this context, but they are not yet available. Thus, in discussing current information-processing models, we will have to accept each model on its own terms. The representations and stages defined by each model will be regarded as primitives.

If stages take input representations and transform them to output representations, they can be described by the terms discrete and continuous in at least three different senses:

(1) Representation: What set of potential inputs is available to a stage, and what set of potential outputs it can select from.
(2) Transformation: How a particular output is determined.
(3) Transmission: How input is made available to a stage, and how its output is made available to the next stage.

These are not the only senses in which models have been said to be discrete or continuous, [2] but they are the ones most directly relevant to current models and most directly addressed by available data.

After elaborating these three senses in which information processing can be discrete or continuous, we will consider the constraints among them. It will be shown that a given stage can be discrete in some senses and continuous in others. Thus, the space of possible discrete and continuous models is complex and multidimensional. A later section will consider the additional opportunities for complexity that arise in multistage models. In particular, one stage within such a model can be discrete in a particular sense even though another stage in the model is continuous in that same sense.

1.2.1. Discrete versus continuous information representation

A stage can be said to be discrete or continuous depending on the information it receives as input or produces as output. Input or output representation is discrete if the available inputs or outputs include only highly distinct information codes. For example, some models assume that the speech recognition stage outputs a binary representation of voice onset time (e.g., Eimas and Corbit 1973). Representation is continuous (very small grain size) if the stage receives inputs or chooses outputs not from distinct codes but rather from codes that can be arbitrarily similar to one another. For example, psychophysical models of sensory discrimination assume that the output of the perceptual stage is a continuously varying representation of the sensory dimension, with the output representations of different inputs being arbitrarily similar to one another (e.g., Green and Swets 1966). Representation could also be intermediate between the discrete and continuous extremes (medium grain size). For example, a model of length discrimination could assume that length was coded as an integral multiple of a minimum unit (cf. Hartley 1981). If the minimum unit were of medium size relative to the range of variation, representation would be neither completely discrete nor completely continuous. Similarly, in

[2] For example, Duncan-Johnson and Donchin (1982) used the terms discrete and continuous in distinguishing between models in which the processing system is always active (continuous) and models in which processing starts with the stimulus and ends with the response (discrete). Others (e.g., Eriksen and St. James 1986; Jonides 1983) have used the terms to distinguish between models in which the boundaries of visual attention are sharp (discrete) or fuzzy (continuous).

any representation based on features (e.g., Tversky 1977), the grain size of representation is determined by the resolution of the feature set. [3]

Townsend and Ashby (1982) have provided good examples of various grain sizes of output representation in their comparison of three different models of letter recognition (cf. Smith and Spoehr 1974; Van Santen and Bamber 1981). First, the 'all-or-none' model uses discrete representations. It assumes identification output is either the correct letter name or a state of complete uncertainty, so output is one of several completely distinct codes. Second, the 'similarity-choice' model uses continuous representations. It assumes identification output is a set of activations indicating the perceptual support for each possible letter. Different representations can be arbitrarily similar to one another, because activations of the letter codes can vary in arbitrarily small units. Third, the 'overlap' model uses representations intermediate between the discrete and continuous extremes (though nearer the discrete end). It assumes identification output is either the correct letter name or a two-letter confusion set. This representation is not at the discrete extreme, because there are intermediate representations (e.g., the code for the BD confusion set is between the code for B and the code for D). It is not continuous either, since codes cannot be arbitrarily similar to one another (e.g., there is no code between the code for B and the code for the BD confusion set).

The grain size of representation can vary not only between letter identification models, but also between different parameterizations of a single model. Consider feature-based models of letter identification (e.g., Rumelhart 1970), in which the output of a letter recognition stage indicates the presence or absence of each visual feature. Grain size is determined by the complexity of the feature detectors, so feature representation can be relatively discrete, relatively continuous, or somewhere between. With extremely complex feature detectors, for example, each stimulus might activate one unique feature; representation would

---

[3] The issue of the grain size of representation is quite different from the issue of whether representation is analog or propositional (Pylyshyn 1979, 1981). The distinction between analog and propositional representation concerns the form of the information representation, whereas grain size refers to its content. Representations are regarded as analog to the extent that the representation space is isomorphic to the stimulus space (often Euclidean), and propositional to the extent that the representation space is different from the stimulus space (usually a feature space). Grain size is an index of similarity of alternative codes, however, and alternative codes can be continuous (i.e., arbitrarily similar to one another) or discrete (i.e., highly dissimilar) regardless of the relation between the representation space and the stimulus space.

be discrete because the output features would be distinct. With extremely simple feature detectors, different stimuli might activate different but overlapping sets of features; representation would be continuous because the sets of activated features could be arbitrarily similar to one another.

It should also be noted that information representations can be continuous in varying numbers of dimensions (cf. Van Santen and Bamber 1981), though a full consideration of these possibilities is beyond the scope of this paper. For example, a perceptual representation might contain the identity of the best guess as to the stimulus letter and a continuously varying degree of certainty as to the accuracy of the guess. This representation would be continuous along only one dimension: the certainty of the single best guess. Alternatively, a representation might contain continuously varying activations representing the amount of evidence for each possible letter, in which case the representation would be continuous along 26 dimensions at once. Distributed representations (e.g., McClelland and Rumelhart 1985) are even more extreme. They contain continuously varying activations for each unit in a multidimensional space, so the representation is continuous along potentially thousands of dimensions at once.

1.2.2. Transformations

A stage can also be described as discrete or continuous in another sense depending on whether it carries out its transformation abruptly or gradually. At the discrete extreme, a stage could perform its transformation using a single 'black box' operation that is considered elementary and indivisible. For example, holistic 'same'–'different' models assume that comparison of two shapes is an indivisible stage that results in an index of their overall similarity (e.g., Lockhead 1972). The transformation performed by such a comparison stage is discrete in the sense that it is not divided into a series of steps or operations, and the model does not distinguish among different states in which the transformation has been started but not yet finished.

At the continuous extreme, a stage could perform its transformation in many steps, using either many different elementary operations or the same operation many times. Such a stage is continuous in the sense that it goes through many states between starting and finishing the transformation, with small differences between successive states. For example, the stage that carries out the mental rotation of a shape has been assumed to use a continuous transformation (e.g., Cooper and Shepard 1973). This stage would rotate a shape 40 degrees by performing a sequence of arbitrarily small rotations on it, and the shape goes through all (or very many of) the intermediate orientations between 0 and 40 degrees. Thus, this stage performs its transformation gradually, going through many intermediate states between start and finish.

Intermediate between the discrete and continuous extremes, a stage could perform its transformation using a few distinct operations rather than just one or very many. Such a stage would go through several, but not too many, states between starting and finishing its transformation. If mental rotation of a shape proceeded in jumps of 10 degrees, for example, a 40-degree rotation would be intermediate between discreteness and continuity. Similarly, a 'same'–'different' comparison stage would be intermediate if it made holistic comparisons for a few distinct attributes of the shapes (e.g., Egeth 1966).[4]

Many models are flexible enough to allow discrete, continuous, or intermediate transformations, depending on certain parameters within the model. Consider the following hypothetical all-or-none letter identification stage, which produces a single letter identity as its discrete output representation. The stage monitors a set of sensory feature detectors, each of which becomes active when its target feature is present in the stimulus. The stage has a separate tally for each possible letter, and when any given feature detector becomes active, the stage increments the tally for each letter consistent with that feature. When the tally for one letter reaches a criterion, the identification stage terminates and that letter is selected as the output. In this model the grain size of transformations depends entirely on the size of the criterion. If the criterion is very small (i.e., only one or two features are needed for identification of a letter), as it might be if the subject had a strong expectation for a particular letter, the transformation from

---

[4] A complication for this analysis is that it is possible for a modeler to stipulate that a certain stage has one or more subprocesses. For example, one might postulate a mental rotation stage with a subprocess doing rotation in 10-degree steps, or a comparison stage with subprocesses to perform comparisons on each distinct attribute. In such a situation, it seems reasonable to analyze the more global stage in terms of its own state transitions, independently of how its subprocesses perform their transformations. Thus, one could regard a comparison stage as having intermediate transformations if it used subprocesses to make comparisons on each of several attributes. The subprocesses would then each have their own status of performing continuous, discrete, or intermediate transformations, depending on how they operated.

sensory input to letter identity would be performed discretely. If the criterion is very large (i.e., many features are needed), the transformation would proceed continuously. Finally, if the criterion were an intermediate number (say five), then the transformation would neither be completely discrete nor fully continuous, but somewhere in between. This example shows that transformations can be either discrete or continuous if units of information are accumulated until a criterion is reached, depending on the number of units needed to satisfy the criterion. Thus, many common stochastic models of decision making can have discrete or continuous transformations, including the random walk model (e.g., Link 1975) and the accumulator model (Vickers 1970).

1.2.3. Information transmission

A third sense in which models can be discrete or continuous concerns the temporal ordering of successive stages within the model. The label 'discrete stage models' is often applied to models in which two ordered stages must carry out their transformations in strict serial fashion, without any temporal overlap (e.g., McClelland 1979; Miller 1982a). In discrete stage models, if the output of stage N is the input to stage $N+1$, then stage $N+1$ cannot begin until stage N finishes. Models are called continuous if two successive stages can be carried out with some temporal overlap (i.e., if stage $N+1$ can start before stage N finishes).

This sense of the distinction between discrete and continuous models is closely related to the interface between stages: that is, to the manner in which the output representation of one stage is transmitted to the next stage. In the simplest form of discrete transmission, stage N produces a single chunk of output after it has finished. Naturally, then, stage $N+1$ cannot begin until stage N finishes. In the simplest form of continuous transmission, stage N gradually transmits whatever partial information it has acquired, without waiting to finish its transformation. Stage $N+1$ starts its transformation as soon as it received this partial information, so the two stages carry out some of their transformations at the same time.

Because of the obvious connection between the way a stage transmits its output and the possibility of overlap between successive stages, we will refer to this sense as discrete or continuous 'information transmission' (cf., Miller 1982a, 1983). It should be kept it mind, however, that our ultimate criterion is in terms of the temporal overlap of successive stages. Thus, discrete transmission includes a model in which stage N gradually sends out preliminary information but stage $N+1$ waits to receive complete information before it begins. For our purposes, it does not matter whether partial information is held in an output buffer in stage N or an input buffer in $N+1$, as long as there is no temporal overlap of the transformations.

Because our criterion for discrete or continuous transmission concerns the overlap between two stages, it is not strictly correct to speak of a single stage as transmitting information discretely or continuously. Nevertheless, in some cases we will focus discussion on the characteristics of a single stage, and in those cases it is very convenient to have labels for the interfaces of that stage to its preceding and following stages. Thus, we will use the term 'input transmission' to refer to the interface between a stage and the one that precedes it. This transmission will be regarded as discrete if the two stages do not overlap, and continuous if they do. Similarly, we will use 'output transmission' to refer to the interface between a stage and the one that follows it.

To explicate the intermediate models, it is helpful to state the distinction in terms of the grain size of information transmission. In completely discrete models the grain is the entire output of a stage, because no partial information can be transmitted. In completely continuous models the grain is arbitrarily small, because any partial information is transmitted as soon as it becomes available. In intermediate models the grain is smaller than the entire output of a stage, but not arbitrarily small. Intermediate models assume that a stage transmits output in several distinct chunks, not continuously (as output becomes available) or discretely (at the end of processing).

Different grain sizes of information transmission can be illustrated with contrasting models for a letter identification task with spoken responses. Assume, for the sake of discussion, that letters are represented internally in a semantic space of logogens (Morton 1969), and that the representation is continuous because the logogens vary continuously in strength. Assume also that the letter identification stage performs a continuous transformation of the sensory input, gradually identifying more and more of the stimulus features. The letter identification stage can be assumed to stop when it has enough features to identify the stimulus with a predetermined degree of certainty or when no further sensory information is present.

At the extreme of discrete transmission, the identification stage would not alter logogen strengths until it was completely finished identifying the stimulus letter. The identification stage might use an internal buffer to keep track of what features had been detected, and then update logogen strengths after it was done accumulating features. (This would be efficient if it were relatively costly to update logogen strengths.) Alternatively, the identification stage could update logogen strengths as it proceeded, as long as the following vocalization stage did not access the logogens until it received a 'DONE' signal from the identification stage or until the logogen strengths had stabilized.

At the extreme of continuous transmission, the identification stage would update logogen strengths as each new feature became available, perhaps incrementing the strength of each logogen consistent with the feature and decrementing the strength of each logogen inconsistent with it. In this case, letter identification transmits a series of outputs, each conveying a small additional piece of information. The subsequent vocalization stage would monitor logogens' strengths as they were adjusted, and would do some kind of preliminary work (e.g., response preparation) based on the pattern of logogen activations at each moment before identification had finished.

A model with transmission intermediate between the discrete and continuous extremes might update logogens at a few distinct times during processing, rather than continuously throughout processing or discretely at its end. For example, features might be accumulated until there were enough to rule out one stimulus letter, and logogen strengths might be updated then. In this model, the grain size of transmission is the number of features needed to rule out one letter. Each information grain would be transmitted as soon as it became available, so transmission would not be discrete with respect to the stimulus as a whole. On the other hand, no partial information about a grain could be transmitted before the whole grain was available, so transmission would not be fully continuous either. The vocalization stage would respond to the input of each new grain of information by making some appropriate preparatory motor adjustments, so it would go through several intermediate states while waiting for identification to finish.

The Asynchronous Discrete Coding (ADC) model suggested by Miller (1982a, 1983) postulates a very large grain size of information transmission, though the model is not quite at the discrete extreme. According to this model, perceptual information is made available to later stages only when a distinct perceptual code is fully activated. That is, a code is transmitted only when perceptual stages have finished constructing that code. If a stimulus is coded in terms of several distinct attributes (e.g., letter name and size) – either because several attributes are relevant to the response or because an irrelevant attribute is coded automatically (Stroop 1935) – then a code for one attribute can be transmitted to later stages even though perceptual stages are still trying to resolve the other attributes. This model is fully discrete with single-attribute stimuli, but it postulates a grain size less than the full stimulus for multiattribute stimuli.

The grain size of transmission is an especially important sense of the distinction between discrete and continuous models because of its implications for methodology in studies using reaction time (RT). Certain powerful methods for making inferences about particular mental stages from total RT (e.g., the AFM, Sternberg 1969a, b) rely on the assumption that total RT is the sum of the times needed for different stages (e.g., Taylor 1976), an assumption consistent with discrete but not continuous or intermediate transmission (McClelland 1979). If transmission were shown to be continuous, then these powerful experimental methods would have to be replaced with alternative methods that are both more difficult to use and less informative (e.g., McClelland 1979; Wickelgren 1977). Furthermore, a large body of literature is based on discrete methods, and many of the conclusions of this literature would essentially be invalidated if transmission were shown to be continuous.

The distinction between discrete and continuous transmission is closely related to the parallel/serial distinction (cf. Townsend 1971). Parallel mechanisms operate at the same time, while serial mechanisms operate one after another. Thus, models with continuous transmission could be said to allow ordered stages to operate in parallel, while models with discrete transmission require ordered stages to operate serially.

It is important to clarify the relationship between the parallel/serial and discrete/continuous transmission distinctions, because Townsend (1971) has shown that parallel and serial models can mimic each other with respect to mean RT in certain types of experiments. The parallel/serial distinction is different from the discrete/continuous distinction in that the former is generally applied to independent processes within a stage, whereas the latter is applied to different stages that are logically contingent. Townsend's (1971) proof applies only to independent processes, so it does not imply that discrete and continuous transmission cannot be separated experimentally.

1.2.4. Relationships among representation, transformation, and transmission [5]

The different senses of discreteness and continuity distinguished here are almost completely independent, because there are very few constraints among them. Certainly, some models have stages that are continuous in all three senses (e.g., Eriksen and Schultz 1979; McClelland 1979), and it is easy to imagine models with stages that are discrete in all three. [6] There are, however, more than just these two categories of stages, because it is possible to construct stages that are discrete in some senses and continuous in others.

Because stages have both input and output, one could in principle distinguish five different grain sizes needed to specify a stage uniquely in the present terms: (1) the representation it receives as input from the previous stage, (2) the transmission by which the input representation is made available to it from the previous stage, (3) the way it carries out its transformation, (4) the output representation it produces, and (5) the transmission by which it makes its output representation available to the next stage. In this analysis, however, we will consider specifically the constraints among the grain sizes of the transformation, output transmission, and output representation. The constraints (or lack thereof) on input and output seem symmetric, as will be noted at the appropriate points below, so it is largely redundant to consider input and output representations or transmissions separately. Furthermore, input and output representation are clearly independent, as are input and output transmission. Table 1 summarizes the discussion below by

---

[5] This section considers in detail the various kinds of stages which can be constructed by combining representations, transformations, and transmissions with different degrees of discreteness or continuity. Readers primarily interested in the implications of available data for discrete and continuous models may want to skip to section 1.3.

[6] To our knowledge, no one has actually proposed a model in which stages are discrete with respect to all three characteristics, though Meyer et al. (1985) have considered predictions of such models. The model underlying the AFM (Sternberg 1969a, b) is normally regarded as discrete in the extreme, but it actually only assumes that stages produce discrete transmissions (see section 1.4).

Table 1
Different combinations of discrete and continuous senses in a given process.

Type	Input or output representation	Transformation	Input or output transmission	Process possible?
1	Discrete	Discrete	Discrete	Yes
2	Continuous	Discrete	Discrete	Yes
3	Discrete	Continuous	Discrete	Yes
4	Continuous	Continuous	Discrete	Yes
5	Discrete	Discrete	Continuous	No
6	Continuous	Discrete	Continuous	No
7	Discrete	Continuous	Continuous	Yes
8	Continuous	Continuous	Continuous	Yes

showing which combinations of the various senses of discreteness and continuity are possible.

1.2.4.1. Transformation and representation. The nature of an output representation is not constrained by the number of steps involved in producing it, so the grain size of output representation is independent of the grain size of transformation. To illustrate, we present examples of the four possible combinations, using mechanisms already discussed.

(1) Continuous transformation and output representation: A letter identification stage could tally arbitrarily small features, ultimately outputting the amount of perceptual support for each alternative letter.
(2) Discrete transformation, continuous output representation: A holistic shape comparison stage might result in a continuous representation of similarity between two shapes, even though it performed a discrete, black-box transformation.
(3) Continuous transformation, discrete output representation: A random walk stage might choose one of two possible outputs in a binary decision, even though it required many small steps to reach the decision criterion.
(4) Discrete transformation and output representation: A random walk stage might choose one of two possible outputs in a binary decision, but it might require only one or two pieces of information to reach its decision criterion.

The independence of transformation and input representation is analogous. Regardless of whether the input is graded or categorical, the transformation could proceed in many small steps or a single indivisible one.

1.2.4.2. Representation and transmission. Three of the four possible combinations are quite straightforward. It is easy for a stage to discretely transmit either a discrete or a continuous representation. All that is required is for the stage to hold its output, of either form, until it is done, possibly using some sort of output buffer to accumulate small features for certain types of continuous representations. Alternatively, the stage could continuously update an output representation accessible to the next stage, but the latter stage could wait for a 'DONE' signal before it started. It is equally easy for a stage to continuously transmit a continuous representation. Continuous transmission requires preliminary output, and the continuity of the representation would provide for a rich vocabulary of graded output representations to be passed from one stage to the next; for example, many small features could be transmitted individually as they were identified. The fourth combination is not so clear: Is it possible to have continuous transmission of a discrete output representation? At first, it appears not, because discrete representations do not seem to provide a way to represent the partial information inherent in continuous transmission. However, our definition of continuous transmission is that preliminary information is transmitted, allowing the later stage to begin before the earlier one is finished. This does not require that the information necessarily be transmitted in small units. Thus, one could construct, with some difficulty, a model in which there was continuous transmission of a discrete representation. Suppose, for example, that a stage was assumed to transmit its current best guess (from among the discrete representations available for its output) at many points in time before it was finished. As long as it did not assign a graded certainty level to each best guess, this would be a discrete output representation. A subsequent stage could potentially begin working given this current best guess, but would have to remain sensitive to its input buffer in case the best guess changed before the subsequent stage was finished. Thus, preliminary transmissions could discretely represent the partial information extracted by a stage. There may be no current models of this somewhat baroque nature, but the point is that one could be devised. Again, the arguments are analogous concerning input transmission and representation. This is easy to see, because the outputs from one stage are the inputs to another, so if all combinations of outputs exist, then all combinations of inputs must also.

1.2.4.3. Transformation and transmission. Though each is independent of representations, these two senses are not completely independent of each other. Specifically, output transmission cannot be continuous if a transformation is discrete. If a transformation operates as a single indivisible step, it does not produce any preliminary information to be transmitted early. We might consider a model in which a stage transmits random outputs before its discrete transformation is completed. Technically, however, this model does not satisfy our definition of continuous transmission, which required preliminary *information* to be passed from one stage to the next. Of course the other three combinations are all possible. Output transmission can be discrete regardless of the type of transformation, and continuous transmission is consistent with continuous transformation. Again, similar arguments apply to the relationship between transformation and input transmission. Input transmission cannot be continuous if the transformation is discrete, because a discrete transformation cannot do part of its work while waiting for full information to become available. The other three combinations, however, are easily imagined.

1.3. Additional flexibility in multistage models

Having established that individual stages can be discrete or continuous in various senses, we must now emphasize the additional theoretical possibilities that arise when a model has more than one stage. Different stages within the same model need not all be equally discrete or continuous, even with respect to one particular sense of the distinction. For example, some stages within a model might perform continuous transformations and some might perform discrete ones. Likewise, the transmission of information from stage A to stage B could be continuous whereas the transmission from B to C was discrete. The complexities are even more striking with respect to representations. Not only could the representation of some information be continuous and that of other information be discrete, but it is possible that some stages within a model would use a given representation continuously, while other stages used the same representation discretely (e.g., by applying thresholds to it).

It is beyond the scope of this analysis to determine exactly what constraints exist between the different senses of discreteness and continuity of successive stages. One such constraint follows directly from the previous discussion of transformation and transmission: if stage N transmits information continuously, then stage $N+1$ must perform a continuous transformation. A second constraint is probably that if the output representation of stage N is discrete, then the input representation to $N+1$ must also be discrete (though the converse is not true because of the possibility of thresholding, as noted above).

Even if there are a few such constraints, however, it seems clear that an enormous number of possibilities will remain after all of the constraints have been taken into account. Thus, with multistage models, the distinction between discrete and continuous models becomes even more complex and multidimensional because of the possibility of using different types of stages within a single model.

1.4. Senses of discreteness required by the Additive Factor Method

Two main assumptions underly the AFM (Pachella 1974), and it is worthwhile to consider briefly how the three different senses of discreteness and continuity distinguished here relate to these assumptions. One assumption is that of strictly serial processing. According to this assumption, total RT is the sum of the times needed for the different stages involved in a task. The other assumption is that of invariant stage output, according to which the character of a stage's output does not vary with the stage's duration.

Clearly, models with continuous transmissions violate the assumption of strictly serial processing, and models with discrete transmission satisfy it. The former models allow temporal overlap among successive stages, so total RT is less than the sum of the times needed for the individual stages. Models with discrete transmissions require earlier stages to finish before later stages can begin, as required by the assumption of seriality.

Consider, however, a hybrid case in which some stages transmit continuously and others transmit discretely. One could then define 'superstages' as groups of stages whose overall input and output were transmitted discretely. For example, in a model with discrete transmission from stage 3 to stage 4, and continuous transmission between all other pairs of stages, there would be two superstages corresponding to stages 1–3 and stages 4–end. In such hybrid cases total RT would be the sum of the times for the different superstages, and one could use the AFM to study the operations of the superstages. Obviously, then, evidence for continuous transmission between two specific stages would not completely invalidate the AFM. In fact, as long as there is at least one point of discrete transmission anywhere in the overall processing system, the AFM would have some utility.

The distinction between discrete and continuous transformations is not critical for the assumptions of the AFM. The grain size of the transformations describes a property of processing within a stage, and the AFM requires no assumptions about how abruptly a stage transforms its input to its output. Thus, the AFM is equally applicable with either discrete or continuous transformations.

The distinction between discrete and continuous representations may appear to be related to the assumption of invariant output, but it is not. The assumption is needed by the AFM in order to eliminate the possibility that a factor has an indirect effect on the duration of stage $N+1$ by changing the character of the output of stage N (Sternberg 1969a). Thus, whether a stage produces a discrete or continuous output representation, it is compatible with the AFM as long as the nature of the variation in outputs does not influence the duration of the subsequent stage. In a brightness judgment task, for example, the encoding stage might produce a continuous output code varying along the single dimension of perceived brightness. As long as the duration of the subsequent stage did not depend on the amount of perceived brightness, the invariance assumption would be satisfied. This could be true even if the *accuracy* of the brightness code depended on the duration of the encoding stage. The subsequent stage would only receive the coded brightness value, and it need not be affected by whether that value was very accurate or was only approximate (indeed, how would it 'know' which values were which?). Thus, the AFM could be applicable with either discrete or continuous representations.

1.5. Summary

To summarize Part 1, it is an enormous oversimplification to categorize models of human information processing into the two classes of discrete and continuous. It is not only impossible to divide the models in a binary fashion, it is even impossible to arrange models on a unidimensional scale from discreteness to continuity. First, discreteness and continuity are a matter of degree. The two terms define the ends of the grain size dimension, and there are intermediate possibilities between these extremes. Second, there are at least three different senses in which stages can be discrete or continuous, and stages can be relatively discrete in some senses and relatively continuous in others. These senses include information representation, transformation, and transmission. Third, multistage models allow the combinations of different types of stages in a great variety of ways. A full model may include some stages that are relatively discrete and other stages that are relatively continuous, even in one particular sense of that distinction. Finally, the AFM depends only on the assumption that information transmission is discrete, and then only between the particular stages or groups of stages influenced by the factors under study.

Part 2:
Critique of evidence thought to support continuous models

This section presents an analysis of four broad categories of evidence previously cited in support of continuous models. For each type of evidence offered, we will consider whether it really supports continuous models, and, if so, in what sense. We will place the most emphasis on the implications of experimental results regarded as incompatible with the assumption underlying discrete stage models – that information transmission is discrete – because that is where the distinctions made in Part 1 shed the most new light. As mentioned above, this assumption is important not only because many models are based on it, but also because of its methodological implications (Sternberg 1969a, b). Regretfully, this means that much interesting research relevant to other senses of discreteness and continuity must be omitted. Particularly noteworthy are the studies of Foster (e.g., Ferraro and Foster 1986; Foster 1979, 1982, 1983) and Massaro (e.g., Massaro and Cohen 1983) relevant to the issue of discreteness or continuity of representations.

2.1. Neuroanatomy and neurophysiology

Probably the most common and superficially compelling argument for continuous models stems from the anatomy of the brain. It is well known that the brain is made up of highly interconnected systems of neurons (e.g., Hubel 1979), and it is often said that there are no black boxes in the head. How, then, can anyone seriously entertain discrete models?

It is possible to question the force of this argument on a number of grounds, even without adopting the dualist position that mental phenomena are not tied directly to neural tissue. Taken to its logical extreme, a 'homogeneous brain' position would deny the entire information processing approach, not just discrete models. After all, if the brain were just one big network of interconnections, then there would be no distinct stages or representations to be explained in the first place.

In any case, neurological and neuropsychological work has underscored the modularity of neural systems (e.g., Coltheart 1985). For example, anatomical evidence indicates that the brain is far from homogeneous. At the level of gross neuroanatomy, the cortex is composed of two distinct hemispheres, each divided into four lobes, which in turn have various fissures, gyri, and sulci (e.g., Hassett 1978; Kolb and Whishaw 1985). At a finer scale, projection maps show great variation in the density of neural connections within and between different areas (Kolb and Whishaw 1985).

More importantly, there is evidence that in many cases distinct anatomical areas carry out distinct stages of mental processing (e.g., Hassett 1978; Kinsbourne 1982). This specialization is most obvious in the case of perceptual and motor functions. Specific regions of neural tissue have been associated with different sensory modalities (e.g., Hubel 1963; Kolb and Whishaw 1985), and the areas controlling different effectors have in some cases been mapped with precision on the order of millimeters (e.g., Penfield and Rasmussen 1950).

Complex cognitive functions also show clear localization. Patients with various types of injuries show highly specific deficits in facial recognition (Benton 1980; Damasio 1985), memory (e.g., Cohen and Squire 1980; Graf and Schacter 1985), verbal number production (McClosky et al. 1986), visual attention (e.g., Posner et al. 1984), and specific language skills (e.g., Coltheart 1985). This suggests the existence of specific brain subsystems – though they need not necessarily be narrowly localized physically – responsible for these specific cognitive functions.

The notion of modular neural systems is completely consistent with discreteness of transmission. It is easy to imagine a system in which individual modules transmit their results only after they have settled into a state of equilibrium – i.e., after they have finished. Of course, it is also possible that transmission begins before equilibrium is reached. But if nature wanted all the modules to be continuously interfaced to each other, why would it bother to build modules in the first place?

A more elaborate argument for continuous transmission could be based on evidence that different neural modules are highly interconnected. The implicit basis of this argument is that high interconnectivity allows transmission of partial information from one area to another, thereby violating the assumption of discrete transmission. [7] The sheer number of fibers connecting two areas is not direct evidence of partial output, however. Indeed, we might just as well argue that so many connections are needed precisely because a great deal of information must be transmitted all at once. Fewer 'lines' would be needed if the message could be transmitted gradually.

The modular architecture of the brain does not seem relevant to the issues of representation and transformation. Instead, our notions of representation would be influenced by evidence about how information was passed from one neural module to another, and our notions of transformation would be influenced by evidence about the workings within a module.

Information about the workings of individual neurons does seem potentially relevant to the discreteness or continuity of representation and transformation. In particular, one might argue for continuity in both senses based on evidence that single neurons seem to code information continuously (e.g., J.A. Anderson 1977) – perhaps by their firing rates (e.g., Terzuolo 1970) – rather than in the all-or-none fashion of digital logic.

---

[7] A more sophisticated version of the high connectivity argument would also include evidence of simultaneous activity in two interconnected systems. The simultaneous activity could be interpreted as evidence of temporal overlap, contrary to discrete transmission. The implications of temporal overlap also arise in the interpretation of other types of psychophysiological and behavioral data, so discussion of the overlap argument will be deferred until to section 2.4 (especially 2.4.4).

That individual neurons vary continuously in firing rate is not very persuasive evidence of continuity in either sense, because the system could use thresholding extensively. Much continuous variation in individual neuron activity may be due to internal noise, and thresholding would be an easy way to filter out this noise. In fact, almost all neurally-based models include significant nonlinearity at the level of individual units (e.g., Grossberg 1978; Rumelhart and McClelland 1982), and discreteness could just result from a high degree of nonlinearity. Furthermore, there is evidence of such thresholding. The strengths of behavioral responses can be discrete even when the responses are based on continuously varying activity in sensory neurons (e.g., Kettner et al. 1980). This evidence clearly suggests the possibility that neural activity is thresholded at some point in the system, and of course if thresholding occurs in one place it could occur in many places.

In summary, there is no compelling evidence for the common notion that discrete models are inconsistent with known neurophysiology. If anything, evidence of modularity suggests that nature has done its best to build black boxes out of biological material.

2.2. *Evidence of graded effects*

We next consider in detail two types of experiments in which gradual changes in an experimental variable produce gradual changes in performance. Without distinguishing between different senses of continuity or different stages within a model, some researchers have argued for continuous models on the basis of such graded effects of experimental variables (e.g., Eriksen and Schultz 1979; Flowers and Wilcox 1982; Norman 1984; Norman and Bobrow 1975). On the surface, this type of argument has considerable appeal, because the existence of graded performance levels suggests an array of graded states within the information processing system. Furthermore, there are continuous models that quite naturally produce graded performance changes in response to graded conditions (e.g., Grice et al. 1974; McClelland 1979).

The two experimental variables to be discussed are the processing time for a relevant stimulus and the salience of an irrelevant stimulus. In both cases, a careful analysis indicates that evidence of graded effects seems to require only that transformation be continuous – not representation or transmission – and that even transformation need only be continuous within one stage, possibly perception or decision. Thus, overarching conclusions of continuity from graded effects are too strong, and they tend to obscure the true implications of the phenomena.

2.2.1. Processing time for a relevant stimulus

The percentage of correct responses (PC) gradually increases with the amount time available for processing the relevant stimulus, up to a maximum level of accuracy attainable in the assigned task. This effect has been observed in a variety of paradigms using either masking (see Breitmeyer (1984), or Kahneman (1968) for a review) or speed stress (see Pachella (1974), Wickelgren (1977), or Wood and Jennings (1976) for a review) to control the time available to perform the task. We will refer to the relationship between processing time and accuracy as the speed–accuracy tradeoff (SAT). [8] The fact that the increase in PC is gradual suggests continuous processing, because it seems to indicate that the amount of processing changes gradually as time is increased.

The gradual SAT can, however, be explained by a model which is discrete in all of the senses distinguished above. As has been pointed out before (e.g., Dosher 1982; Ollman 1966; Wickelgren 1977), a gradual increase in PC can be explained as a gradually changing probability mixture (Everitt and Hand 1981) of two processing outcomes. The simplest example is the Fast Guess Model (Ollman 1966; Yellott 1967, 1971), in which every response is generated from one of two processing outcomes. One outcome is that the fully discrete system has finished completely, in which case it generates a response with maximal accuracy. The other outcome is that the fully discrete system has not finished completely, so its response must be a complete guess. In this model it need only be assumed that the time needed for the discrete system to finish varies randomly from trial to trial. Because of trial-to-trial variability, a gradual increase in the available time (i.e., greater mask delay or less speed stress) causes a gradual increase in the proportion of trials on which processing finishes. Thus, PC must increase gradually with processing time, even if there are only two

---

[8] The term 'speed–accuracy tradeoff' is normally applied to the time–accuracy functions obtained in speed-stress but not masking paradigms. Nevertheless, the paradigms have comparable implications for discrete and continuous models, so it is convenient to refer to both with a single term for the purposes of the present analysis. This use is not intended to imply that the same mechanisms limit performance in the two paradigms.

distinct outcomes of discrete processing.[9] The Fast Guess Model may be rejectable on other grounds (e.g., Pachella 1974), but it is clearly capable of generating gradual SATs.

Further studies of the SAT suggest two principles of information processing that cannot be explained with the two-state discrete model described above. The first is that there are some responses based on partial information, not just guesses and responses based on full information. The second is that information processing can be interrupted at unpredictable times without complete disruption of performance. We will argue, however, that these principles are only suggestive of continuous transformations, not continuous representations or transmissions.

2.2.1.1. Responses based on partial information.

It seems clear that there are some responses based on partial information, not just guesses and responses based on complete analysis (cf. Massaro et al. 1979). This principle is demonstrated by studies showing that the type of errors, as well as the proportion, changes with increases in processing time (e.g., Grice et al. 1982; Pachella 1974; Stanovich 1979; Stanovich et al. 1977). For example, Stanovich et al. (1977) found that with very little processing time subjects simply guessed among stimulus letters, whereas with more time responses were strongly influenced by pairwise letter similarities. This finding is evidence that incorrect responses can still be based on partial information, because complete guesses could not be influenced by stimulus similarities.

The existence of individual responses based on partial information rules out the simple two-state discrete model considered above, because that model does not allow any states with partial information. In what sense does this evidence require continuous processing, however?

Clearly, information representation need not be continuous. A discrete code can be selected on the basis of partial processing just as easily as on the basis of full processing or guessing. The defining characteristic of discrete representation is that the output is one dis-

---

[9] Meyer and Irwin (1981) examined RT distributions in an attempt to evaluate this two-state RT model. They concluded that, at least in their task, a three-state model (i.e., some responses based on a fixed amount of partial information) was capable of explaining the dependence on processing time not only of overall PC but also of the probability distribution of RT. Ratcliff (1985) argued that continuous models would give the same results, however. In any case, other evidence discussed next seems to indicate that there are responses based on varying amounts of information about the stimulus, not just responses based on only two or three informational states.

tinct alternative, and it does not matter how much processing went into its selection. For example, the perceptual stage might simply provide a discrete representation of its best guess as to the stimulus, even if this best guess were based on different amounts of information in conditions with different amounts of processing time.

Information transmission need not be continuous, either. No information need be transmitted until a stage is finished, whether it does a cursory analysis, thorough analysis, or something in between. Transmission can occur discretely at the point when the code is selected, regardless of how long the stage worked. For example, the perceptual stage need not output its best guess until it has used up all of its processing time.[10] When time has elapsed, this stage could send a single discrete stimulus representation to later stages, which could then initiate the response. In such a model, response accuracy would grow gradually as a function of processing time. Furthermore, with intermediate amounts of processing time, responses would be based on partial information (i.e., whatever information the perceptual stage had had time to accrue).

It seems clear that the existence of responses based on partial information demonstrates continuous *transformations*. At least one stage must perform its function in small steps, gradually accumulating whatever information constitutes its output. The number of steps taken by the stage must be adjustable (at least on average) in response to experimenter-imposed limitations on processing time. When its allotted processing time had expired on any one trial, such a stage could simply output its single best guess from among its alternative outputs, thereby using both a discrete output representation and a discrete transmission. For example, the perceptual stage could gradually extract information about the stimulus, stopping after varying amounts of time and transmitting a single discretely coded output based on varying amounts of information. The important point is that a set of such discrete outputs would vary continuously in accuracy, thereby producing the usual SAT functions. Furthermore, even a little processing would allow the set of outputs to reflect similarities among stimuli (or responses), in accordance with the findings of Stanovich et al. (1977).

[10] In this example and the next we are assuming that the perceptual stage is curtailed when task time is limited, as it would be in many tasks. Some results of Sanders and Houtmans (1985) suggest that the perceptual stage may not be the one to be curtailed in tasks with pressure to respond quickly, however. In tasks where decision or response preparation time is curtailed, analogous arguments could be applied to these later stages.

Most stochastic models of decision making developed to account for the SAT are entirely compatible with the idea of stages continuous only in transformation, not representation or transmission. For example, the random walk model (Link 1975) posits a variable information criterion, set in advance of a trial, controlling how long a stage will work toward resolution of its discrete (in most cases binary) output. Though the point has not been emphasized by those developing these models, it is clear that a stage operating as a random walk need not transmit its output to later stages until it has finished (i.e., reached its boundary criterion) on any trial.

Meyer et al. (1985) mentioned the random walk model as an example of a continuous model, and it is clear that we do not agree with that classification. Our view is that the random walk is a model of a stage or even a process, and that the output of the random walk could be a discrete representation, discretely transmitted. Furthermore, the transformation carried out by the random walk could also be relatively discrete, if the boundary criterion required taking only a few steps. Apparently, their use of the term 'continuous' refers mainly to the type of transformations carried out, and we shall argue in Part 3 that their experimental results also address this sense of discreteness versus continuity.

2.2.1.2. Interruptability. The second principle is that information processing can be interrupted at unpredictable times, and performance will still improve gradually as more processing time is allowed. For example, Reed (1973, 1976) had subjects make recognition responses to visual stimuli. Subjects were to respond immediately after a response signal was given (in this case, stimulus offset), even if they had to guess to do so. The time available for processing (i.e., time from stimulus onset to its offset) was varied randomly from trial to trial over a range of several seconds, and the dependent measure was the probability of a correct response at each value of processing time. The response signal procedure differs from the more traditional deadline procedure (e.g., Pachella and Fisher 1969) in that the subjects in the response signal procedure do not know in advance of the trial how much processing time will be available on that trial.

With a little practice, subjects perform fairly well in the response signal procedure, and the resulting SATs have the same general shape obtained when deadlines are used. Thus, it seems to make little qualitative difference whether the system is set for partial processing in advance or simply interrupted part way through. Unfortunately, there have been few direct evaluations of the effect of foreknowledge of processing time, so it is impossible to know whether it makes any quantitative difference. Likewise, it is not clear whether this procedure elicits responses based on partial information (e.g., Stanovich, et al. 1977) or just guesses and responses based on full information.

The fact that processing is interruptable argues against models in which a processing criterion must be determined in advance of a trial, but it is not evidence against discreteness in either representation or transmission. Certainly, output could be a discrete code (e.g., best guess) whether a stage terminates expectedly or unexpectedly, and information could be transmitted in a single information grain once a stage has finished, regardless of what caused the termination.

Therefore, interruptable processing merely reinforces the conclusion that there must be continuous transformations in at least one stage. Whereas the existence of responses containing partial information shows that some stage can terminate after intermediate amounts of processing, interruptability shows that it must actually go through the states of partial processing on each trial. A stage with discrete transformations might terminate after partial processing, but only if the precision of the transformation were set in advance. Partial information would not be available from a discrete transformation that was interrupted at unpredictable times, because the transformation would either be finished or it would not.

In summary, SAT phenomena suggest at most that information transformation is continuous in at least one stage. Furthermore, it appears that the stage with continuous transformations can be interrupted, at which point it will use its current state to select a best guess for an output code. Nevertheless, the phenomena are compatible with discrete information representation and transmission, so they cannot be regarded as evidence in favor of fully continuous models. The output code of a continuously transforming stage may be one of a discrete set of alternatives, and this code may be transmitted discretely to the next stage at the termination of the continuous transformation. Since processing can be interrupted at continuously varying points, the amount of information used to select the output code will vary continuously. Thus, the accuracy of the output codes will vary continuously even though the codes themselves are discrete.

2.2.2. Salience of irrelevant stimuli

Another graded effect cited in support of continuous models is the effect of unattended flankers in focused attention tasks. Eriksen and Schultz (1979), for example, had subjects make two-alternative forced-choice responses to target letters presented in the relevant middle position of a display. Irrelevant flanker letters presented on both sides of the relevant letter were to be ignored. As in earlier studies (e.g., Eriksen and Eriksen 1974), the results indicated that flankers produced a response compatibility effect. Specifically, responses were relatively fast when the flankers were target letters assigned to the same response as the relevant center target, and responses were relatively slow when the flankers were target letters assigned to the opposite response.

In three experiments Eriksen and Schultz (1979) found that the response compatibility effect varied in a graded fashion as a function of several manipulations of the relative processing time of flankers and targets: the size and figure/ground contrast of the targets, holding flankers constant (experiment 1); the size of the flankers, holding targets constant (experiment 2); and the stimulus onset asynchrony (SOA) of the target and the flankers (experiment 3; see also Flowers and Wilcox 1982). In each case, they found that the response compatibility effect gradually increased as processing of the target was delayed relative to processing of the flankers, whether delay was due to smaller size, lower contrast or later presentation.

Eriksen and Schultz (1979) argued that the compatibility effect is difficult to reconcile with discrete models, particularly because gradual manipulations of processing time produce gradual variations in the size of the effect. They argued that their results support a 'continuous flow model', in which (1) information about stimuli accumulates gradually in the perceptual stage, activating form units corresponding to the different stimulus alternatives (continuous transformations), (2) the output from each form unit is a continuously varying activation level indicating the perceptual support for that form (continuous representation), and (3) activation building up in form units is fed continuously through the decision stage to the response system, where responses are primed as perceptual evidence accumulates in support of them (continuous transmission).

The results of Eriksen and Schultz (1979) are quite consistent with their fully continuous model, as are some more recent findings obtained with psychophysiological measures. These findings indicate that flankers influence the motor activations of the effectors used to execute each of the two possible responses. In particular, there is measurable motor activation of the incorrect response hand on trials with response-incompatible flankers, and this activation is correlated with the RT inhibition produced by the flankers (Coles and Gratton 1986; Coles et al. 1985; Eriksen et al. 1985; O'Hara et al. 1981). This is strong evidence that flankers cause activation of response effectors, and that activation of the incorrect effector is at least partly responsible for slowing correct responses.

The graded effects of irrelevant flankers seem to be evidence of continuous processing in all senses, but they are actually not difficult to explain with models that are fully or at least primarily discrete. Because the effects on RT are averages across a number of trials within a condition, it is possible to explain their gradual change with a two-state model in which the flankers sometimes have a large effect and sometimes have a much smaller effect or no effect at all, just as the gradual form of the SAT might be explained with a two-state model. Similarly, a two-state model could be invoked to explain effects on average values of the electrophysiological measures.

One might challenge this two-state model, arguing that the psychophysiological data demonstrate continuous variation in the activation of individual responses on a trial-by-trial basis. When looking at single trials, for example, one does see continuous variations in electromyogram activity on the incorrect response hand. Like the continuous trial-to-trial variability in RT, however, this does not rule out a model with two underlying informational states and a lot of superimposed random noise. Obviously, continuous noise does not rule out discrete information processing in any interesting sense.

One might also argue against the two-state model based on the analysis of RT distributions provided by Eriksen et al. (1986). They found that responses to incompatible displays were slower than responses to compatible ones by about the same amount throughout the entire distribution of RTs. According to any model in which the size of the compatibility effect varies from trial to trial, including a two-state mixture model, the compatibility effect should be larger at the high end of the RT distribution than at the low end. Thus, assuming adequate statistical power to reject the mixture distribution, the finding of Eriksen et al. (1986) suggests the surprising conclusion that the incompatibility effect is the same size on all trials, at least under the conditions of this experiment. Extrapolating to other conditions, one would then conclude that graded changes in compatibility effects result from continuous variation in the size of the flanker compatibility effect on a trial-by-trial basis, not from changes in mixture proportions.

Though we doubt the conclusion that the compatibility effect is the same on all trials within a condition, and therefore question the statistical power of the analysis conducted by Eriksen et al. (1986), we will grant for the sake of argument that a two-state model can be rejected, and that it can be shown that the continuous variation in the salience of the irrelevant flankers (rather than noise) produces continuous variation in the activation of responses. Even in this case, we will argue, the graded flanker effects would only require models with continuous transformations, not models with continuous representations or transmissions.

Consider the following model, which is continuous only in the transformation carried out by the response activation stage. Stimulus letters are identified in parallel across the different display positions, both relevant and irrelevant, by mechanisms tied to individual letter positions. Each mechanism performs a discrete transformation leading to the selection of a discrete code for the name of the letter in its position, and no information about any letter is transmitted to the decision stage until identification of that letter is complete (discrete transmission). In the decision stage, each letter is evaluated with respect to its relevance (i.e., position) and its response assignment. Letters are evaluated one at a time, in the order of their arrival from letter identification mechanisms, with queueing if one letter arrives before the previous one has been completely processed. After each letter has been analyzed, the decision stage sends a discrete output to the motor system. This output indicates a response to be activated (left or right), and a degree of activation (large or small) depending on letter relevance.[11] The motor system activates the indicated response to the

---

[11] It is certainly not optimal to send output derived from irrelevant letters, yet the response compatibility effect shows that something nonoptimal is taking place. Eriksen et al. (1985) suggested that the nonoptimality was itself evidence against discrete transmission, because a discrete system would not send the small activations produced by irrelevant flankers. However, there is no logical reason that a discrete system cannot transmit both large and small activations, even if such action is not optimal for performance in this particular task. In fact, there need only be two levels of activation – one for relevant letters and one for irrelevant – not continuously graded amounts of activation, so the basic result is consistent with discrete representations as well.

indicated degree, but does so with a gradual transformation, initiating the response when a threshold activation is reached. Furthermore, the buildup of motor activation is slow enough, relative to letter recognition and decision times, that information about several different letters is received during activation – not just information about the target. When more than one transmission has been received from the decision stage, motor activation builds at a rate proportional to the sum of the activations.

In this model both relevant and irrelevant stimulus letters are encoded discretely, and discretely transmitted from stage to stage. Because *multiple stimulus codes* are processed in parallel, the model can account for the basic compatibility effect quite easily. Flankers influence the rate at which activation of the correct response builds up, thereby producing the compatibility effect, because they weakly activate motor responses at the end of their discrete journey through the information processing system.

The model can also explain why the compatibility effect depends on the relative processing time of flankers and targets, as demonstrated by Eriksen and Schultz (1979). If the flankers are identified before the target, the flanker names will be transmitted to the decision stage first and have more time to influence response activation.[12] If flankers are recognized after the relevant letter, they will have less time to influence the buildup of response activation, and they may even be too late altogether.

The above model, which accounts well for response compatibility phenomena, is discrete in most ways, but it does have two features that are not fully discrete. First, it is type of ADC model (Miller 1982a), in that stimulus displays are not treated as unitary wholes, but rather as amalgamations of separate letters. This feature of the model allows discrete transmissions of the different letters at different times, thus providing a convenient explanation of how the effects of flankers depend on their relative salience. As argued by Miller (1982a), ADC models are nearer the discrete than the continuous end of the possibilities for information transmission. Second, response activation was assumed to operate as a continuous transformation, slowly building the energy required to initiate a response. This feature of the model allows

---

[12] There is probably a limit to this. If the flanker arrived too much earlier than the target name, the motor system might be reset by the time the target arrived.

response activations to be influenced by both center letters and flankers on a given trial, to varying extents depending on the relative recognition times of the different letters. Therein lies the model's capability to generate graded activations of the incorrect response from trial to trial.

In summary, it appears that the graded effects of unattended flanker letters, like the basic SAT, do not rule out two-state models with changing probability mixtures of the two discrete processing outcomes. Even if such models were ruled out, the graded effects would at most be evidence for models in which at least one stage carries out its transformation relatively continuously.

2.3. Effects of featural similarity

Two papers have attempted to use effects of featural similarity to discriminate between discrete and continuous models, both using extensions of Eriksen's focused attention paradigm. Both appear relevant to the information transmission sense of the debate.

Within the focused attention paradigm, Yeh and Eriksen (1984) used the upper case letters A and E as the target letters for one response, and used upper case G and Q as targets for the other. These target letters appeared both in the relevant middle position of the display and in the irrelevant flanker positions. To test between discrete and continuous models, lower case versions of these same letters (matched with the upper case letters in size) were also presented as flankers on some trials. The lower case flankers 'a' and 'e' have important visual features in common with the targets G and Q. According to continuous models this featural information would cause some partial priming of the response associated with the targets G and Q, because response priming begins before flanker name recognition has finished. This priming should slow responses with the displays 'aAa', 'eEe', 'aEa', and 'eAe', and speed them with the displays 'aGa', 'eGe', 'aQa', and 'eQe'. Some evidence of the predicted effects was obtained.

In an analogous test, Miller (1982b) looked for effects of nontarget letters that were visually similar to target letters. Sometimes, for example, C and I were target letters assigned to one response, K and V were target letters assigned to the other response, and the featurally similar letters G, T, R, and U were used as flankers. With such stimuli many continuous models (e.g., Eriksen and Schultz 1979) predict that responses should be faster when flankers are visually similar to the target letter assigned to the same response as the relevant middle letter (e.g., TCT) than when flankers are visually similar to a target letter assigned to the opposite response (e.g., RCR, UCU). The visually similar flanker would provide some partial activation in support of the target letter, and this partial activation would influence responses before the perceptual stage had successfully classified the flanker as a response-irrelevant letter. No significant effects in the direction predicted by continuous models were obtained when the visually similar letters appeared in irrelevant flanker positions. However, when the task was changed to visual search, so that all display positions were relevant, small but significant effects of similarity were obtained as predicted by continuous models.

The similarity effects obtained by Yeh and Eriksen (1984) and Miller (1982b) suggest that features sometimes prime responses. Assuming that responses are selected on the basis of letter names rather than visual features, it would thus appear that featural information must be made available to response stages before perceptual recognition is complete. Thus, to the extent that similarity effects were obtained, both papers provide support for models with continuous information transmission.

Any number of replies are available in defense of discrete transmission, however. First, it is possible that subjects code the stimuli in terms of visual features rather than letter names, at least on some trials. [13] If so, visual feature analysis could finish completely before any response priming occurred, and yet latency could still be influenced by the number of features favoring each response. Second, the observed effects may reflect only the speed of perceptual processing. Perhaps the perceptual stage finishes sooner when all of the available featural information supports the same decision, perhaps by adjusting an internal certainty criterion (cf. Miller and Bauer 1981). This model accounts for similarity effects entirely within the perceptual system, so it does not require continuous transmission of perceptual information to response stages. Third, it may be that flankers similar to targets were occasionally incorrectly recognized as targets. If so, then discrete but

---

[13] Dr. Sanders pointed out that this possibility is especially likely when visually similar stimulus pairs are assigned to common responses. On the other hand, Miller and Bauer (1981) found evidence that the perceptual system does distinguish, at least on some trials, between visually similar stimuli assigned to the same response.

incorrect name codes could be responsible for the response priming, as discussed in section 2.2.2. Miller (1982b) argued against this explanation, but it is very difficult to exclude the possibility completely. Fourth, it is possible that the results were stimulus artifacts, since the pool of appropriate stimuli was rather small.

2.4. *Evidence of overlapping stages*

Many researchers have argued for continuous models on the basis of evidence suggesting that different stages operate at the same time (i.e., temporal overlap). Such evidence specifically supports continuous transmission, because a later stage needs partial output if it is to begin before the stage providing its input has finished. As emphasized earlier, arguments for continuous transmission are particularly important because of their consequences for the interpretation of RT and the validity of the AFM. Thus, they bear careful analysis.

Even if overlap can be demonstrated, two auxiliary criteria must also be considered before discrete transmission can be challenged seriously. First, one stage must be *contingent* upon the other one. That is, it must be clear on logical or empirical grounds that one stage is earlier in the processing sequence than the other, and that the output of the earlier stage is the sole input to the later stage. [14] If the later stage can begin without any input from the earlier one, then the fact that they overlap does not imply continuous transmission of partial information.

In copy typing, for example, it seems clear that reading of new text takes place at the same time as typing of previously read text, so perception and responding must overlap (e.g., Shaffer 1971). This finding is evidence against models in which a single-channel, limited-capacity processor can only be allocated to one task at a time, because two distinct functions are being carried out at once. It is not evidence of continuous transmission, however, because the contingency criterion is not satisfied. That is, typing of old text is not contingent upon recognition of new text, nor vice versa. Discrete models can easily explain the overlap by appealing to the metaphor of an assembly line.

---

[14] Note that if the later stage also gets input from some third stage, then it could start before the earlier stage had finished if a discrete output from the third stage had already been made available. To challenge discrete transmission, one must show not only that the later stage begins before the earlier one has finished, but also that what it begins doing is processing preliminary information provided by that earlier stage.

In an assembly line different discrete stages work on different stimuli at the same time, and no stage passes its output along to the next stage in line until it has completely finished processing a given input. [15]

The second criterion concerns the range of processing over which overlap is demonstrated. The mere existence of overlapping processes is not sufficient to reject discrete transmission unless those processes seem to belong to different processing stages (Sanders 1980). Even then, evidence of overlap implies continuous transmission only from the earliest to the latest of the overlapping stages. The point here is simply that it is not valid to extrapolate beyond the range of the observed overlap. The fact that a few particular stages overlap does not imply that transmission is continuous throughout the entire range of stages involved in the task.

Consider, for example, the interactive activation model of word recognition (e.g., McClelland and Rumelhart 1981; Rumelhart and McClelland 1982), in which letter detectors continuously transmit partial output to word detectors. If we were to accept the success of this model as evidence of overlap between letter and word recognition, with the latter contingent upon the former, then we would have evidence of continuous transmission from letter to word recognition. Even if we were to regard these as two different stages, however, the range criterion would prohibit us from also concluding that the word recognition stage transmits its output continuously (say to a decision or speech production system) or that any subsequent stages involved in the task transmit continuously.

We shall next examine evidence of overlapping stages – some of which has been regarded as critical evidence against models with discrete transmission and/or against the validity of the AFM – in light of the contingency and range criteria.

2.4.1. *Turvey's (1973) evidence for concurrent-contingent visual information processing*

Turvey (1973) conducted an elegant series of masking experiments that revealed two kinds of processes underlying visual masking: periph-

---

[15] Indeed, the metaphor of an assembly line seems much more appropriate for copy typing than the metaphor of cascading stages. In an assembly line model, there would be no mutual interference between stages, because the intermediate states of different stages are insulated from each other by discrete transmission. Models with continuous transmission would seem to predict considerable interference among stages working on different stimuli, since the different stages are assumed to be communicating constantly.

eral and central. A variety of results suggested that the peripheral process involves activation of context-independent (i.e., hard-wired) feature detectors by visual input, whereas the central process appears to assemble a representation of overall stimulus shape and identifies the stimulus by consulting long-term memory. Furthermore, the characteristics of the peripheral and central processes strongly suggest that they are successive, with the output from peripheral processes being the input to the central process. Output of the central process could be described as being a letter identity placed in the visual icon, from which later stages could retrieve stimulus identity.

The results of major interest here were those indicating temporal overlap of central and peripheral processes. The critical observation was that the time from stimulus onset to the end of central processing was not affected by certain variables that did affect the time from stimulus onset to the end of peripheral processing (Turvey 1973). This finding is inconsistent with a model in which peripheral and central processes operate in strict succession (i.e., with no overlap), because such a model requires that the time to the end of central processing is the sum of the times needed by the peripheral and central processes. With additive process times, a variable affecting the duration of the former process would obviously also affect the sum of the durations.

Turvey (1973) suggested that his results could be explained by a 'concurrent-contingent' model that allows continuous information transmission from the peripheral process to the central process. According to this model, the peripheral process extracts different features at different times, and each feature is transmitted to the central process as soon as it is extracted. The central process starts as soon as it receives the first feature from the peripheral process, and it may deal with features serially. Thus, peripheral and central processes overlap temporally, assuming that multiple features are required for letter identification. The model also assumes that the central processing of each feature is relatively slow compared to peripheral feature extraction. This assumption implies that the peripheral process is always able to extract at least one more feature while the central process is dealing with the previously extracted one. Once extracted, the feature(s) are held in a buffer until needed by the central process. Thus, the central process only has to wait for the peripheral process to extract the first feature; subsequent features are available to the central process as soon as it is ready for them. This model explains the critical observation that a factor can affect peripheral processing time without influencing the total time needed for letter identification. Since the peripheral process extracts additional features before the central process is ready for them, the peripheral process can be slowed somewhat without increasing the total time for letter identification, though of course a large enough slowing would increase the time for letter identification.[16]

As Turvey (1973) noted, it is also possible to account for his data with a model in which central processes do not receive output from peripheral processes, thus preserving discrete transmission. However, he regarded the required mechanisms as so improbable that he argued that his data supported continuous transmission from peripheral to central letter identification processes.

Even if it is granted that Turvey's results support continuous transmission from peripheral to central processes in letter identification, it must be emphasized that this has extremely limited implications regarding transmission for the information processing system as a whole. Taking the range criterion into account, it is clear that the evidence supports continuous transmission at a very specific point in the perceptual system, as did the success of the Interactive Activation model (Rumelhart and McClelland 1982). The point is that one cannot rule out discrete transmission of an identified stimulus to a decision stage, or of a selected response to the response preparation stage, from evidence of overlapping perceptual processes.

Turvey's results suggest a gradual accumulation of peripheral information overlapping with a gradual resolution of overall form at a more central perceptual level, ultimately leading to letter recognition. For the many models in which letter recognition is regarded as a single complex stage (e.g., Sternberg 1969a, b), however, this is evidence of continuous transformations within that stage, not evidence of continuous transmission between stages. The AFM might still be a valid tool for examining stages in such models, because these stages could operate in strict succession even if there were continuous and partially overlapping processes within each one. That is, the AFM is a valuable tool as long

---

[16] In principle, this model predicts some effect on total letter identification time whenever there is a change in the time needed for the first feature to be extracted. To explain the critical observation, then, it must also be assumed that the experimental manipulation produces little or none of its effect on the peripheral processing needed to extract the first feature.

as there is *at least one* point of discrete transmission in the overall model.

2.4.2. Overlap of encoding and response selection

Miller (1976) and Stanovich and Pachella (1977) conducted experiments using the AFM to examine the initial stage of perceptual processing (often termed 'encoding') in information processing tasks. Both obtained patterns of interaction that were interpreted as support for temporal overlap between perception and response selection. Since in both cases response selection was based on the output of encoding, the contingency criterion was satisfied. Furthermore, the hypothesized overlap occurs over a wide enough range of information processing stages (perception and response selection) to provide a serious challenge for almost any model with discrete transmission. Thus, these two papers warrant careful analysis.

Miller (1976) studied the interaction of stimulus probability and visual quality. In previous studies these factors had been found to interact in some tasks and with some stimuli but to be additive with other tasks and/or stimuli (Miller and Pachella 1973, 1976; Pachella and Miller 1976). Miller (1976) suggested that the interaction of probability and quality was in fact always present in the time needed for stimulus encoding. The interaction sometimes disappeared in total RT, however, when it was concealed by a slow response selection stage going on in parallel with encoding. The response selection stage was hypothesized to receive partial output from the encoding stage (i.e., continuous transmission), and to use this partial output to begin its work. If response selection were very slow, then its duration would determine RT (Schweickert 1978), and the duration of the encoding stage carried out at the same time would not influence RT. In particular, this means that the interactive effect of probability and quality on encoding duration would not be seen in RT if response selection were sufficiently slow. As predicted from this overlapping stages model, Miller (1976) found that the interaction of probability and quality was obtained in all tasks if stimulus quality was sufficiently reduced (i.e., encoding was slowed enough).

Subsequently, however, Miller (1979) concluded that the interaction of probability and stimulus quality is actually controlled by the type of representation that the subject seeks to encode in a given task. In particular, activation of abstract internal representations or logogens

(Morton 1969) is influenced by both visual quality and stimulus probability, producing the interaction of the two factors. However, if subjects use a direct mapping of visual features to responses, without performing any intermediate coding, then no stage is affected by both factors, and they are additive. This view reconciles discrete models with the changing pattern of interaction versus additivity. If the encoding stage performs different functions with different tasks and/or stimuli, then it can easily be influenced differently by experimental factors in the different cases.

In a related series of studies, Stanovich and Pachella (1977) found that degrading the visual quality of a stimulus has a larger effect when stimulus–response (S–R) compatibility is high than when it is low. This underadditive interaction of compatibility and stimulus quality is somewhat surprising in view of other experiments finding additivity of the two factors (e.g., Blackman 1975; Frowein and Sanders 1978; Hardzinski 1980; Shwartz et al. 1977; Sternberg 1969a), and it may result from an artifact in the experimental setting (Hardzinski 1980; Sanders 1980). Nonetheless, it is interesting to consider Stanovich and Pachella's (1977) interpretation of the results.

Like Miller (1976), Stanovich and Pachella (1977) suggested that preliminary perceptual information is used to begin response selection while identification finishes, so the two stages are carried out mostly in parallel. When compatibility is low, response selection takes so long to process preliminary perceptual information that more perceptual information is always available by the time the response stage needs it, even if stimulus quality is low. Thus, prolonging the encoding stage has little effect because it goes on in parallel with a slow response selection stage (cf. Schweickert 1978). When compatibility is high, however, response selection is so fast that it processes perceptual information as quickly as it becomes available. Thus, prolonging the encoding stage has a large effect, because it slows down the entire system.

Although the explanation based on overlapping stages handles the underadditivity nicely, there are also ways to reconcile the results with a fully discrete model. For example, Sanders (1980; De Jong and Sanders 1986) noted that Stanovich and Pachella (1977) obtained an unusually large visual quality effect from their contrast manipulation, and he suggested that the input to the response choice stage might have been slightly distorted in the low quality condition. By changing the representation of the stimuli, this would essentially reduce S–R compatibility. It is reasonable that a small change from optimal S–R compatibility would have more effect than a small change from suboptimal S–R compatibility, thus explaining the underadditivity.

The theoretical approach taken by Miller (1979) can also explain Stanovich and Pachella's (1977) results with a fully discrete model. It must simply be assumed that the encoding stage produces a different representation of the stimulus when S–R compatibility is low than when it is high, possibly because the response selection stage is more complicated (cf. Holender 1980). Naturally, if a different representation is produced, it would not be surprising to obtain different effects of experimental variables, even if the output of encoding is transmitted discretely. This explanation, while preserving discrete transmission, is incompatible with the constant stage output assumption of the AFM.

2.4.3. Other behavioral evidence of overlapping stages

There are many other studies presenting behavioral evidence of overlapping stages, at least some of which have been taken to support continuous transmission. Space limitations preclude an in-depth analysis of every one, but it is possible to indicate briefly the reasons why most do not seriously challenge discrete models.

In most demonstrated cases of overlapping stages, the contingency criterion is not satisfied. For example, there is evidence that form and size discriminations can overlap (Ellis and Chase 1971), form and color discriminations can overlap (Ellis and Chase 1971), tone pitch and duration discriminations can overlap (Hansen and Hillyard 1983), both color and pitch discriminations can overlap with switching attention between modalities (LaBerge 1973), the motor programming of an eye movement can overlap with perceptual processing of information in the current fixation (e.g., Rayner and Pollatsek 1981; but see Sanders and Houtmans 1985), the programming of different movement features (e.g., limb, extent) can occur partly in parallel (Rosenbaum 1980), and the execution of early responses in a sequence can overlap with the planning of later responses (Rosenbaum et al. 1987). In none of these cases, however, was the output of one overlapping stage or process needed as the input to another. It may also be noted that in most cases the range criterion is not satisfied either, because the overlapping stages tend to be performing functions at about the same point in the information processing sequence.

The contingency criterion also weakens the conclusions of Flowers and Wilcox (1982), who argued for continuous transmission partly on the basis of evidence that facilitory and inhibitory stages overlap in the Eriksen flanker paradigm. Though both of these stages are contingent upon perception of the flanker, they need not have a contingent relationship with each other, so continuous transmission is not implied.

Fletcher (1983) presented evidence that recognition of word identity and category is accomplished by partially separate processes carried out to some degree in parallel. This need not be taken as strong evidence against discrete models, because the overlapping processes are both perceptual in nature. It is not clear whether the contingency criterion is satisfied, because it is not known whether category judgments are contingent upon perception of identity.

2.4.4. Psychophysiological evidence concerning overlapping stages of brain processing

There is growing recognition that psychophysiological measures may be useful in discriminating between information-processing models originally developed to explain behavioral data (e.g., Posner and McLeod 1982; Vaughan and Ritter 1973). In particular, many researchers have studied event-related potentials (ERPs), which are time-locked averages of the electroencephalographic activity evoked in response to specific stimuli in specific tasks, and some have found evidence relevant to the issue of discrete versus continuous information transmission.

Vaughan and Ritter (1973) presented what is perhaps the simplest and most direct argument for overlapping stages:

> 'in simple RT tasks, the motor response begins long before the termination of central neural activity generated by the stimulus. Due to this extensive overlap of physiologic activity, notions of information processing involving additive, exclusive serial processes must be oversimple.' (Vaughan and Ritter 1973: 133)

Thus, the argument is that overlapping physiological activity implies continuous transmission (see Renault et al. (1982) for another example of this argument).

There are two related flaws in this argument. First, it is very difficult to show that a particular physiological response arises from a particular type of mental processing. Even if the response is influenced by variables associated with that type of processing, it is impossible to be sure that the physiological response does not arise later, from consequences of that processing rather than from the processing itself. As Ritter et al. (1983b) put it: 'The circumstance that the peak latency of a component is differentially affected by changes in a particular processing demand does not distinguish between the possibility that the component reflects the relevant underlying process or an outcome of that process' (1983b: 144).

The second flaw is more serious, because it arises even in cases when the first flaw can be overcome. Specifically, the argument from overlapping physiological activity ignores the contingency criterion. We do not know whether a later stage of physiological processing is contingent upon all of the activity carried out by an earlier one, or perhaps just on the first part of that activity. For example, Vaughan and Ritter (1973) state that a flash of light lasting less than 1 msec can produce activity in the optic nerve that lasts up to 100 msec. Perhaps the initial 25 msec of this activity results in the discrete detection of the light onset, and the remaining 75 msec does something else such as extracting brightness information or settling of the nerve back to its stimulus-absent state as a result of flash offset. If so, central processing for a detection task would not be contingent upon the later activity in the optic nerve. The fact that central or motor stages overlapped with optic nerve processes on which they were not contingent is, of course, consistent with discrete transmission. The point is that even if some physiological activity arises from neural mechanisms responsible for a specific informational function, it is virtually impossible to be sure that *all* of that activity arises in connection with that function.

Psychophysiologists have offered a number of other, more indirect arguments concerning discrete versus continuous transmission. For example, Ford et al. (1980) varied positive set size and stimulus quality in a memory scanning task (Sternberg 1969b). In addition to RT, they measured the peak latency of the P300, a component of ERPs thought to index stimulus evaluation and decision making time (e.g., Kutas et al. 1977). If P300 peak latency in this task indicates the instant at which the response has been selected, then the additional response time beyond that instant (RT–P300 latency) should be independent of both memory set size and visual quality, according to Sternberg's (1969b) four-stage serial model. Contrary to this prediction, the RT–P300 difference increased both with reduced stimulus quality and with increased memory set size. An alternative model was suggested in which response preparation was carried out in parallel with decision making. Mulder et al. (1984) also observed that the RT–P300 difference increased with memory set size, and seconded the idea that response preparation overlapped with decision making.[17]

Clearly, the results of Ford et al. (1980) only indicate overlap under the assumption that the P300 peak latency indicates the time at which the decision is finalized.[18] Previous research has not established this strong conclusion, and there are even disputes about whether P300 results from a stage before or after S–R compatibility has its effect (e.g., McCarthy and Donchin 1981; Ragot and Renault 1981). Even if we grant the weaker conclusion that P300 peak latency increases with decision latency, however, the results do not require overlapping stages. Since waveforms generally get broader as their peak latencies increase (e.g., Chase et al. 1984), the choice to measure peak latency was seemingly an arbitrary, but not inconsequential, one. Furthermore, it seems quite unlikely *a priori* that the P300 peak, as opposed to some other point on the waveform, would indicate the instant at which the decision was finalized. It is interesting to note that Ford et al.'s (1980) assumption – that processing is finished at the peak – is quite different from Vaughan and Ritter's (1973) – that processing is not finished until activity is over. The objection to both is the same: It is just not known what physiological index might correspond to the end of processing for specific information.

Even if response execution time is affected by memory set size or stimulus quality, as argued by Ford et al. (1980) and Mulder et al. (1984), this finding is not automatically incompatible with discrete transmission. For example, one could imagine a discrete model in which different stages had to share a limited pool of resources. When

---

[17] Neither paper spelled out the parallel model in any detail, and we do not see how overlapping stages can be used to explain the effects. Normally, overlapping stages have been invoked to explain the reduction or disappearance of effects on one stage, based on the idea that they are covered up by a slower stage going on in parallel (e.g., Miller 1976; Schweickert 1978; Stanovich and Pachella 1977; Turvey 1973). By that logic, a model with overlapping decision and response stages would seem to predict that the response stage would need less time to finish, not more, after the end of a slower decision stage (i.e. when set size was larger or stimulus quality reduced).

[18] For their arguments, the peak latency need not be exactly the point at which the decision is made, as long as it is off by a constant. It is not sufficient, however, for the peak latency and decision latency to be perfectly correlated.

the memory set size was large, the comparison stage might get more resources and the response stage less. When the stimulus was degraded, the identification stage might get more resources at the expense of the response stage. In either case, the response stage would operate more slowly as a result of the reduction in its resources. [19] Thus, effects of memory set size and stimulus quality on the duration of the response stage are evidence against Sternberg's (1969b) particular model, not against discrete transmission in general.

It should be noted that other psychophysiologists have obtained evidence they regarded as more consistent with discrete than continuous transmission. Ritter et al. (1983b; 1982; 1983a) isolated two ERP components by performing subtractions between pairs of observed ERP waveforms. Subjects were tested in a simple RT task and in a choice RT task with one stimulus presented on 80% of the trials and another stimulus presented on 20%. One component (termed 'Na') was isolated by subtracting the waveform elicited in the simple RT task from that elicited by the 80% choice stimulus, and evidence suggested that this component reflected the onset of a pattern recognition stage. A later component ('N2') was isolated by subtracting the waveform elicited by the 80% stimulus from the waveform elicited by the 20% stimulus, and evidence suggested that this component reflected the onset of a stimulus categorization stage. Furthermore, the N2 component seemed to be contingent upon the Na component.

The effects of discrimination difficulty on the latencies of the Na and N2 components provided some suggestion of discrete transmission. Across a number of different stimulus sets, the onset latency of the Na component did not vary as a function of the difficulty of the discrimination required by the choice task, but the peak latency did. This is consistent with the idea that pattern recognition can begin as soon as any stimulus is presented, but that more recognition processing is needed when the discrimination is harder. Onset and peak latency of the N2 component increased with discrimination difficulty by about the same amount, and this amount was greater than or equal to the increase in Na peak latency. Thus, all phases of the stimulus classification stage, including its onset, were delayed by the full amount of time

---

[19] This model is difficult to reconcile with current views on resource sharing (e.g., Gopher and Sanders 1984), and we do not wish to argue strongly for it. The purpose of describing it is to emphasize how hard it is to provide an unequivocal demonstration of overlapping stages.

needed for extra pattern recognition. This suggests that there is little or no overlap between the stages generating the Na and N2 components. If there were much overlap, then N2 onset latency should thereby show much less effect of discrimination difficulty than N2 peak latency.

Ritter et al. (1983b) emphasized the consistency of their data with sequential pattern recognition and stimulus classification stages, but they also concluded that pattern recognition and stimulus classification were partially overlapping, primarily because N2 onset occurred earlier than Na peak. However, based on the earlier argument that overlapping physiological activity is not necessarily indicative of overlapping mental stages, we find the evidence of sequentiality more convincing than the evidence of overlap.

Part 3:
Evidence in favor of discrete models

One very general class of evidence in favor of discrete models is the consistent, systematic, and readily interpretable patterns of results obtained by researchers using the AFM. Because this method assumes discrete transmission, it seems likely that the method would eventually lead to serious contradictions and inconsistencies if transmission were really continuous. Although the results from any one study can almost always be interpreted within the framework of the AFM, consistent patterns of additivity and interaction across a number of studies not only provide compelling demonstrations of the utility of the method but also raise one's confidence in its assumptions (Sanders 1980).

In fact, the AFM has been used very widely, and reviews of the evidence obtained with this method indicate fairly consistent results. For example, Sternberg (1975) reviewed research on memory scanning, Sanders (1980, 1983) reviewed studies of choice reaction time tasks and the effects of stress, and Sanders et al. (1982) considered the effects of sleep loss. In no case were there obvious contradictions between studies, as would be expected if an invalid method were being used. Sanders (1980, 1983) has argued particularly effectively that the available results make a strong case for the utility of the method, especially because the obtained patterns of additivity and interaction are largely independent of the demands placed on individual stages.

Of course, the existence of many results consistent with a theory does not prove the theory or any of its assumptions, because alternative theories might also be able to account for the results. Therefore, the patterns of consistent results emerging from use of the AFM cannot be regarded as decisive evidence against continuous models, just as the many results consistent with continuous models could not necessarily be regarded as inconsistent with discrete ones (Section 2). For this reason and because they have already been thoroughly reviewed (e.g., Sanders 1980, 1983), we will not consider in detail the consistent patterns emerging from studies using the AFM.

Instead, we will in this section concentrate on a few critical experiments designed explicitly to test between discrete and continuous models. These studies have been designed to provide critical tests by examining response preparation, and the results generally favor discrete or nearly discrete models. In these studies, the subject is given certain critical stimulus information that is correlated with the correct response, and the resulting response preparation is monitored in different ways, as discussed below.

Although the distinction has not previously been emphasized, it will be argued here that response preparation studies are critical tests of very different senses of discreteness and continuity depending on how the critical stimulus information is presented. In some studies the critical information is given in a cue presented prior to the imperative stimulus (e.g., Meyer et al. 1985; Sanders 1971), and these studies tend to address the issue of discrete versus continuous transformation. In other studies the critical information is given in a distinct attribute or set of features of the imperative stimulus itself (e.g., Miller 1982a), and these studies are relevant to the issue of discrete versus continuous transmission.

3.1. Effects of response-correlated cues

Many experiments have included response-correlated cues presented prior to the imperative stimulus, and it is beyond the scope of this paper to review them. Instead, we will discuss two specific paradigms which were originally presented in terms of the issues of temporal overlap and discrete vs. continuous models.

In the most straightforward study of this type, Sanders (1971) sought to determine whether there was any overlap in the processing of an informative cue and an imperative stimulus. Imperative stimuli were rectangles appearing to the left or right of fixation, to which the subject made spatially compatible keypress responses. A prior informative cue indicated which response was more probable, with 80% cue validity, and SOA between cue and imperative stimulus was varied.

Sanders (1971) found that informative cues could speed responses to imperative stimuli and that their facilitatory effects were heavily dependent on SOA. Facilitation was absent for SOAs of 150 msec or less, appearing suddenly and at virtually 'full strength' for SOAs of 200 msec or more. Sanders (1971) argued that if processing of cues and imperative stimuli did overlap, then facilitation should begin at an SOA of about 50 msec (the estimated encoding time of the cue), and it should develop gradually. Thus, he rejected the idea of processing overlap between informative cues and imperative stimuli.

In what sense do these results support discrete models? First, the results do not support discrete representation, even for the states of response preparation found in the system. The degree of attained response preparation might vary continuously from trial to trial, for example, and the effect on RT could index the average benefit produced by that preparation. Even if the same degree of preparation was attained on every trial in this experiment, preparation might still vary continuously as a function of the predictive validity (79%, 80%, 81% ...) of the cue. The results suggest invariance in the *time* at which preparation is attained, not in the *extent* of that preparation. Second, the results do not support discrete transmission, because the cue was physically and temporally separate from the test stimulus. Showing that there is no temporal overlap in processing of two separate stimuli does not imply that there is no overlap of different stages processing a single stimulus, as would be required to infer discrete transmission. Third, the results do support discrete transformation within the response preparation stage.[20] Assuming that the RT to the imperative stimulus reflects the extent of response preparation at the instant the imperative stimulus is presented, the abrupt change in the cuing effect across SOA indicates an abrupt transformation from the unprepared to the prepared state.

---

[20] In this case, we would define as 'discrete' any transformation abrupt enough to show up as a single transition given the SOA values used.

Meyer et al. (1984, 1985; 1982) have developed a more elaborate version of the cuing paradigm, including a novel analysis of RT distributions. As will be seen, their methods address the same sense of discreteness and continuity (transformation) as those of Sanders (1971).

Meyer et al. (1984, 1985) sought to discriminate between discrete and continuous models by determining the number of intermediate states involved in response preparation. In their experiments subjects were first given a cue indicating the correct response (100% validity), and then given a test stimulus to which that response was required. Subjects had extensive practice with the invariant relationship between cue and response, and they were given incentives to use the information in the cue to speed their responses as much as possible.

Meyer et al. (1984, 1985) characterized the most extreme discrete model as one in which a cuing stimulus results in either full response preparation or no preparation at all, with no possible states of intermediate preparation. In a fully continuous model, on the other hand, a full range of partially prepared intermediate states is possible. They also considered the possibility of multi-state discrete models, in which there may be one or two intermediate states of response preparation (i.e., a large grain size of response preparation).

The main independent variable of Meyer et al. (1984, 1985) was the SOA between cue and test stimulus, which was either short, medium, or long. With short SOAs, subjects would not have time to do any response preparation using the cue, so they should be in a fully unprepared state at the onset of the test stimulus. With very long SOAs, subjects would have plenty of time to do complete response preparation, so they should be in a fully prepared state.

The important question was what would happen with a medium SOA. According to the two-state discrete model, the subject would attain either full preparation or none at all. There could of course be random trial-to-trial variation in how long it took the subject to attain the state of full preparation, so with a medium SOA some responses would be made from the unprepared state and some would be made from the fully prepared state. However, no responses could be made from a state of intermediate preparation, because no such state is allowed by the model. Therefore, the two-state discrete model predicts that the RT distribution observed with the medium SOA should be a probability mixture of the RT distributions observed with the long and short SOAs.

Continuous models, as defined by Meyer et al. (1984, 1985), allow states of intermediate response preparation to arise when the subject is given a medium amount of time between cue and test stimulus. According to these models, then, the RT distribution observed with a medium SOA need not be a probability mixture of the RT distributions observed with long and short SOAs.

To determine whether RT distributions obtained with the medium SOA could be described as probability mixtures of the distributions obtained with long and short SOAs, new statistical techniques beyond the scope of this article were devised (Smith et al. 1982). Using these techniques, the results of several experiments were found to be roughly in accord with discrete models. Specifically, the RT distribution obtained with medium SOAs seemed to be a probability mixture of two, or at most a few, discrete states of response preparation. The discrete model could be rejected, however, when the task involved relatively many different responses and relatively low S–R compatibility (Meyer et al. 1985).

Like the simpler cuing paradigm of Sanders (1971), the experimental paradigm and analytic techniques developed by Meyer et al. (1984, 1985) seem primarily relevant to the issue of discrete versus continuous transformations within the response preparation stage. By probing the response preparation stage with a test stimulus at an intermediate time, their methods are clearly designed to measure the number of intermediate states through which the response preparation stage passes. This corresponds exactly to our notion of the grain size of transformations within response preparation. The main advantage of the sophisticated distributional analysis of Meyer et al. (1984, 1985) is that it allows detection of a discrete transformation even if there is variability in the time at which it occurs.

Unlike Sanders (1971), Meyer et al. (1984, 1985) found evidence of intermediate states of response preparation in some cases. However, just as Sanders' evidence of a discrete shift from the unprepared state to the prepared state does not imply discrete representation or transmission, Meyer et al.'s evidence of a few cases with intermediate states of response preparation does not imply continuous representation or transmission even for those few cases. Even a very large number of intermediate states of response preparation (i.e., highly continuous transformations) would be compatible with models that were discrete in these other two senses. One way to establish this point is to exhibit a model with discrete information representation and discrete transmissions between all pairs of stages, but which can account for many possible intermediate states of response preparation by using a small grain size of transformation within the response stage. Such a model does exist: Suppose the cue is discretely represented as belonging to one of two alternative classes, with one class associated with each response. Furthermore, processing of the cue eventually results in the discrete transmission of one of two response codes to the response stage, indicating which response should be prepared. At that point, response preparation begins, and it continues until the test stimulus is presented. As long as the response preparation stage transforms continuously from the unprepared to the prepared state, the amount of response preparation at stimulus onset would vary continuously, even though all other aspects of the model are discrete.

In summary, cuing effects may be useful for discovering whether the response preparation stage carries out its transformations discretely or continuously, but they are not directly relevant to the issues of discrete or continuous representations and transmissions. It is instructive to consider explicitly the reason why cuing paradigms (e.g. Meyer et al. 1984, 1985; Sanders 1971) do not demonstrate continuous transmission even if evidence of continuous response preparation is obtained. Quite simply, the reason is that a cue is physically and temporally separate from a test stimulus. Therefore, even discrete transmission of information provided by the cue would allow response preparation to begin before the test stimulus is recognized – possibly even before the test stimulus is presented. Discretely transmitted information about the cue could produce a continuously variable response activation, possibly proportional to the continuously varying time between cue and test stimulus. This activation, in turn, could produce a graded effect on RT to the test stimulus. In order to falsify discrete transmission, it is necessary to show that response stages receive partial information from one stimulus before perceptual processing of that *same stimulus* is complete. Cuing effects, however, only show that response stages receive partial information from one stimulus before perceptual processing of *another stimulus* is complete. Thus, the separation of the two stimuli is what prevents graded cuing effects from suggesting continuous transmission. Even if it could be shown that response preparation using the cue went on in parallel with perceptual analysis of the test stimulus – a possibility which Meyer et al. (1984, 1985) explicitly disallow as being incompatible with the assumptions of their analysis – there would still be no evidence that partial information about the cue was transmitted before perceptual analysis of the cue was finished.

3.2. *Effects of response-correlated stimulus attributes*

When response-correlated information is presented as part of the test stimulus itself, effects of this information can provide evidence against fully discrete transmission. In particular, evidence that response preparation begins before analysis of the test stimulus is complete (i.e., based on information which is only imperfectly correlated with the correct response) would suggest that the response preparation stage received at least two transmissions of information about the test stimulus. An earlier transmission would convey the correlated information and enable response preparation to begin. A later one would actually determine the response, allowing it to be executed. The presence of multiple transmissions from a single test stimulus (unlike the existence of two transmissions from a cue and a test stimulus) would be incompatible with fully discrete models in which the response information from a single stimulus was transmitted in a discrete chunk after all of that stimulus' information about the response had been ascertained. Multiple transmissions would be compatible with models in which transmission was continuous, of course. If the response-correlated information were presented in a distinct, separately codable attribute of the test stimulus (e.g., its name), multiple transmissions would also be compatible with the ADC model (Miller 1982a, 1983).

Dumas et al. (1972) were the first to present response-correlated information in an attribute of the test stimulus itself. They studied the memory scanning task (Sternberg 1969b), in which subjects make a positive response if a test stimulus belongs to a previously memorized set and a negative response otherwise. Multiattribute stimuli were used, and one critical attribute (e.g., red) was more common among stimuli associated with the positive response than among stimuli associated with the negative response. The results showed that positive responses were faster to stimuli with the critical attribute than to stimuli without it, and the reverse was true for negative responses. This finding suggests that the presence of the critical attribute caused preparation of the positive response, and that response preparation took place in parallel with memory comparison. If this is indeed the explanation of the effect, then it implies that the response system gets partial information about the test stimulus (i.e., the critical attribute) before getting full information about it (i.e., whether it belongs to the positive set).

Although the results of Dumas et al. (1972) are consistent with continuous transmission of information to the response preparation stage, they are not strong evidence of it for two reasons. First, the effects could plausibly arise in the memory comparison stage. Dumas et al. assume that comparison is an exhaustive, item-by-item process, but this may not be true with multiattribute stimuli, especially when attributes are differentially correlated with responses. Other sorts of comparison stages (e.g., hierarchical) could easily yield faster positive and slower negative responses for test items with the critical feature. If the effects do arise in the comparison stage, then they are not evidence that response preparation begins before it has complete information. Second, even if response preparation does take place in parallel with comparison, this finding is perfectly consistent with the ADC model. Because the stimuli were made up of several distinct attributes, it is possible that information about each attribute was discretely transmitted to the response preparation stage. This would allow response preparation to occur before memory comparison was done, even though each attribute was processed discretely.

Miller (1982a, 1983, 1985a, 1987) performed a series of experiments to see whether partial information available early in perceptual processing of a stimulus (i.e., preliminary information) could be used to begin preparing responses. He manipulated pairwise stimulus discriminabilities to control the availability of preliminary information. For example, one stimulus set consisted of a large and small 'S' and a large and small 'T', with the size discrimination chosen to be much more difficult than the letter name discrimination. With this stimulus set it was expected that letter name could be recognized by the perceptual system before size, and the question was whether preliminary name information could be used to prepare responses before size recognition was complete. Another stimulus set consisted of the letters M, N, U, and V, for which pilot work had shown that the pairs MN and UV had high interpair and low intrapair discriminability. With this set, preliminary information would indicate to which of the two highly confusable pairs a stimulus letter belonged.

Several different paradigms were used to try to determine whether response preparation began as soon as preliminary information became available, thereby overlapping with the later perceptual processing. Miller (1982a) had subjects respond to the four stimuli by pressing buttons with the index and middle fingers of the two hands, and the assignment of stimuli to responses was varied. In the same-hand condition two similar stimuli (e.g., small & large S, or M & N) were assigned to response fingers on the same hand, and in the different-hand condition they were not. Other results suggest that, with this response set, preparation of two response fingers on the same hand leads to faster responding than preparation of two response fingers on different hands (Miller 1982a, 1985b, 1987, 1988). Based on these results, responses in the same-hand condition should be faster than those in the different-hand condition if preliminary information is used to prepare responses, and Miller (1982a) argued that there should be no difference if it is not. The results indicated a difference between the same- and different-hand conditions (i.e., evidence of preliminary response preparation) only when preliminary information corresponded to a discretely coded stimulus attribute like letter name, not when it corresponded to difficult to code information such as which of two pairs (e.g., MN versus UV) a stimulus was from. Thus, these results were interpreted as support for the ADC model, not for fully continuous transmission.

Reeve and Proctor (1984, 1985; Proctor and Reeve 1985) have suggested that the difference between same- and different-hand conditions does not really reflect response preparation. Using a divided attention task, however, Miller (1987) obtained strong evidence that it does. Targets on one channel were small and large S's and T's, assigned to four response keys using either the same- or different-hand condition. Medium-sized S's and T's were sometimes presented as distractors on this channel. Targets on the other channel were bright squares appearing in one of four locations, and locations were mapped compatibly onto the same four response keys used to respond to letters. When distractors were presented on the letter channel, the names of these distractors were found to influence responses to bright squares. Specifically, responses were facilitated if the bright square called for one of the two responses assigned to target letters with the same name as the distractor, and inhibited otherwise. Thus, the names of the distractor letters caused some response preparation. Furthermore, the size of this response preparation effect increased with the difficulty of the size discrimination needed to identify distractor letters, indicating that the preparation took place before distractor letters had been fully recognized. Interestingly, Miller (1985a) found no evidence of response preparation in an analogous two-choice divided attention situation. It seems plausible that preliminary information might not be used to prepare single responses, because, whereas two prepared responses can hold each other in check, a single prepared response may be executed prematurely. Further research will be needed to clarify the reason for this difference in results.

Miller (1983) developed another paradigm to monitor response preparation enabled by preliminary information. Basically, the same-hand condition of Miller (1982a) was used, and relative stimulus discriminabilities were varied to control the time available for response preparation (i.e., the time between perception of preliminary information and perception of exact stimulus identity). Increasing the time available for response preparation was expected to increase its effect. In addition, a cue was sometimes given prior to stimulus onset, and the cue conveyed the same information as the preliminary stimulus discrimination. When a cue was given, response preparation could be carried out in advance of the stimulus, thus decreasing the effect of response preparation enabled by preliminary stimulus information (it was assumed that response preparation was subject to diminishing returns). In this paradigm, then, preliminary response preparation is revealed by an interaction of discriminability and cuing. If response preparation occurs, discriminability should have a larger effect when cues are given than when they are not, because in the latter case some of the extra discrimination time is offset by increased preparation. If response preparation does not occur, the effects of discriminability and cuing should be additive. Again, evidence for response preparation was obtained only when preliminary information was discretely codable, supporting the ADC model rather than continuous transmission.

In summary, the evidence obtained from studies of response preparation is more compatible with the ADC model than with fully continuous information transmission. There is evidence that preliminary information about a stimulus can be used to prepare responses when that information is highly codable, but in general preliminary information does not seem to be so used. It would appear that the grain size of information transmission is fairly large, including all of each internal code used to represent the stimulus. These findings are generally compatible with discrete stage models, except when stimulus sets are represented with multiple internal codes.

3.3. Relationship between the two types of response preparation studies

There appears at first to be a conflict between the conclusions drawn from studies of response-correlated cues and those of response-correlated stimulus attributes. The results of Meyer et al. (1984, 1985) indicated that response preparation operates discretely, passing through at most one intermediate state between the unprepared and fully prepared conditions. The results of Miller (1982a, 1983, 1985a, 1987), however, indicate that preliminary information about a codable stimulus attribute can be used to prepare responses before full stimulus information is available. The apparent contradiction is that the response preparation implied by Miller's results requires partially prepared responses, which seem incompatible with the conclusion of Meyer et al. (1984, 1985) that there are only a few different states of response preparation.

On closer analysis, the apparent conflict disappears. One reason is that Miller's results do not require continuous degrees of response preparation. Instead, they require one state of preparation produced by the single bit of preliminary stimulus information. The response can be made either from an unprepared state or this prepared state, and the probability of being in the prepared state can vary across conditions to produce the obtained effects. It could be the case, then, that the prepared state implied by Miller's results is analogous to the prepared state in Meyer et al.'s two-state model of response preparation. A second reason is that most of Meyer et al.'s (1984, 1985) evidence for discrete response preparation comes from two-choice tasks, whereas Miller's evidence of an intermediate level of response preparation comes from four-choice tasks. It is easy to imagine that intermediate states of response preparation could correspond to activation of subsets of responses. Thus, even if Miller's results were produced by intermediate levels of response preparation, this important task difference could explain the discrepant number of intermediate states of response preparation.

Conclusions

The issue of whether human information-processing is discrete or continuous is enormously complex theoretically, for the three reasons outlined in Part 1. First, discreteness and continuity are not a dichotomy, but rather a dimension on which there are intermediate possibilities. Second, stages can be discrete or continuous in at least three different senses: information representation, transformation, and transmission. Third, some of the stages used to perform a task may be discrete and others continuous, even in a given sense. Thus, there is a graded multidimensional space of information-processing models differing with respect to discreteness and continuity. Future comparisons of models, both theoretical and empirical, must acknowledge the great complexity of this space to avoid overgeneralization and equivocation.

A review of evidence cited in support of continuous models turned up no decisive evidence of continuity. On careful analysis, functional neuroanatomy, behavioral responses, and psychophysiological responses are all consistent with primarily discrete models. For example, gradual speed–accuracy tradeoffs and other graded effects can be explained in terms of changing mixtures of two distinct states. Even if they could not, they would support continuity only in the limited sense of continuous transformations within at least one specific stage. But even this limited sense of continuity is not supported in direct examinations of the transformations carried out in preparing responses (Meyer et al. 1984, 1985).

Evidence previously cited as contradicting the AFM (Sternberg 1969a, b) is remarkably weak. So far, evidence of overlapping mental stages is indirect at best, consisting mostly of results that can be explained with only slight elaborations of existing models in which transmission is discrete. Psychophysiological measures do not yet provide direct indices of the offset of mental events, which would be needed to show that two events overlap. In many cases the putative overlapping stages would not constitute general evidence against discrete transmissions even if they had been demonstrated unequivocally. In some of these cases there is little reason to believe that the second stage is contingent upon the first, and in others the two overlapping stages perform such closely related functions that they would never have been logical candidates for separate, discrete stages in the first place.

If anything, the available evidence supports discrete transmission, consistent with the AFM. Studies of the preparation produced by response-correlated stimulus attributes (e.g., Miller 1982a) suggest that this preparation occurs only when response-correlated information is provided by a discretely codable stimulus attribute. Other information available early in perceptual processing seems to be held until recognition is complete rather than being continuously transmitted to response stages.

Although the AFM can be questioned on other grounds (e.g., Pieters 1983; Sanders 1980; Taylor 1976), it appears that the considerable shift in opinion against it has been motivated mainly by the unwarranted belief that there is strong evidence against discrete transmission. Given the balance of evidence in favor of discrete transmission, it seems quite unjustified to reject sequential stage models and discard the AFM at this time. Sequential stage models have had considerable success in explaining human performance (e.g., Chase 1984; Sanders et al. 1982; Sternberg 1969a, b) – far too much success to be dismissed without solid grounds.

References

Anderson, J.A., 1977. 'Neural models with cognitive implications'. In: D. LaBerge and S. Samuels (eds.), Basic processes in reading: perception and comprehension. Hillsdale, NJ: Erlbaum. pp. 27–90.

Anderson, J.R., 1980. Cognitive psychology and its implications. San Francisco, CA: W.H. Freeman.

Anderson, J.R. and G.H. Bower, 1980. Human associative memory: a brief edition. Hillsdale, NJ: Erlbaum.

Benton, A., 1980. The neuropsychology of facial recognition. American Psychologist 35, 176–186.

Blackman, A.R., 1975. Test of the additive-factor method of choice reaction time analysis. Perceptual and Motor Skills 41, 607–613.

Bower, G.H., 1975. 'Cognitive psychology: an introduction'. In: W.K. Estes (ed.), Handbook of learning and cognitive processes, Vol. I. Introduction to concepts and issues. Hillsdale, NJ: Erlbaum. pp. 25–80.

Breitmeyer, B.G., 1984. Visual masking: an integrative approach. New York: Oxford University Press.

Broadbent, D.E., 1958. Perception and communication. London: Pergamon Press.

Chase, W.G., 1984. 'The timing of mental acts'. In: E. Donchin (ed.), Cognitive psychophysiology: event-related potentials and the study of cognition. Hillsdale, NJ: Erlbaum. pp. 221–247.

Chase, W.G., G. McCarthy, K.C. Squires and R.W. Schvaneveldt, 1984. 'Report of panel IV: mental chronometry'. In: E. Donchin (ed.), Cognitive psychophysiology: event-related potentials and the study of cognition. The Carmel Conferences, Vol. 1. Hillsdale, NJ: Erlbaum. pp. 249–301.

Cohen, N.J. and L.R. Squire, 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. Science 210, 207–210.

Coles, M.G.H. and G. Gratton, 1986. 'Cognitive psychophysiology and the study of states and processes'. In: G.R.J. Hockey, A.W.K. Gaillard and M.G.H. Coles (eds.), Energetics and human information processing. Dordrecht: Martinus Nijhoff. pp. 409–424.

Coles, M.G.H., G. Gratton, T.R. Bashore, C.W. Eriksen and E. Donchin, 1985. A psychophysiological investigation of the continuous flow model of human information processing. Journal of Experimental Psychology: Human Perception and Performance 11, 529–553.

Coltheart, M., 1985. 'Cognitive neuropsychology and the study of reading'. In: M.I. Posner and O.S.M. Marin (eds.), Mechanisms of attention: attention & performance XI. Hillsdale, NJ: Erlbaum. pp. 3–37.

Cooper, L.A. and R.N. Shepard, 1973. 'Chronometric studies of the rotation of mental images'. In: W.G. Chase (ed.), Visual information processing. New York: Academic Press. pp. 75–176.

Damasio, A.R., 1985. Prosopagnosia. Trends in Neurosciences 8, 132–135.

De Jong, F. and A.F. Sanders, 1986. Relative signal frequency imbalance does not affect perceptual encoding in choice reactions. Acta Psychologica 62, 211–223.

Dosher, B.A., 1982. Effect of sentence size and network distance on retrieval speed. Journal of Experimental Psychology: Learning, Memory and Cognition 8, 173–207.

Dumas, J., E. Gross and S.F. Checkosky, 1972. Effects of attribute probability in a memory search task. Journal of Experimental Psychology 93, 327–332.

Duncan-Johnson, C.C. and E. Donchin, 1982. The P300 component of the event-related brain potential as an index of information processing. Biological Psychology 14, 1–52.

Egeth, H.E., 1966. Parallel vs. serial processes in multidimensional stimulus discrimination. Perception & Psychophysics 1, 245–252.

Eimas, P.D. and J.D. Corbit, 1973. Selective adaptation of linguistic feature detectors. Cognitive Psychology 4, 99–109.

Ellis, S.H. and W.G. Chase, 1971. Parallel processing in item recognition. Perception & Psychophysics 10, 379–384.

Eriksen, B.A. and C.W. Eriksen, 1974. Effects of noise letters upon the identification of a target letter in a nonsearch task. Perception & Psychophysics 16, 143–149.

Eriksen, B.A., C.W. Eriksen and J.E. Hoffman, 1986. Recognition memory and attentional selection: serial scanning is not enough. Journal of Experimental Psychology: Human Perception and Performance 12, 476–483.

Eriksen, C.W. and D.W. Schultz, 1979. Information processing in visual search: a continuous flow conception and experimental results. Perception & Psychophysics 25, 249–263.

Eriksen, C.W. and J.D. St. James, 1986. Visual attention within and around the field of focal attention: a zoom lens model. Perception & Psychophysics 40, 225–240.

Eriksen, C.W., M.G.H. Coles, L.R. Morris and W.P. O'Hara, 1985. An electromyographic examination of response competition. Bulletin of the Psychonomic Society 23, 165–168.

Estes, W.K., 1977. 'On the interaction of perception and memory in reading'. In: D. LaBerge and S.J. Samuels (eds.), Basic processes in reading: perception and comprehension. Hillsdale, NJ: Erlbaum. pp. 1–25.

Everitt, B.S. and B.J. Hand, 1981. Finite mixture distributions. New York: Chapman & Hall.

Ferraro, M. and D.H. Foster, 1986. Discrete and continuous modes of curved-line discrimination controlled by effective stimulus duration. Spatial Vision 1, 219–230.

Fletcher, B.C., 1983. The role of category information in word identification: a parallel decision model. Memory & Cognition 11, 237–250.

Flowers, J.H. and N. Wilcox, 1982. The effect of flanking context on visual classification: the joint contribution of interactions at different processing levels. Perception & Psychophysics 32, 581–591.

Ford, J.M., R.C. Mohs, A. Pfefferbaum and B.S. Kopell, 1980. 'On the utility of P3 latency and RT for studying cognitive processes'. In: H.H. Kornhuber and L. Deecke (eds.), Motivation, motor and sensory processes of the brain: electrical potentials, behavior and clinical use. Progress in brain research, Vol. 5. Amsterdam: North-Holland. pp. 661-667.

Foster, D.H., 1979. Discrete internal pattern representations and visual detection of small changes in pattern shape. Perception & Psychophysics 26, 459-468.

Foster, D.H., 1982. 'Analysis of discrete internal representation of visual pattern stimuli'. In: J. Beck (ed.), Organization and representation in perception. Hillsdale, NJ: Erlbaum. pp. 319-341.

Foster, D.H., 1983. Visual discrimination, categorical identification, and categorical rating in brief displays of curved lines: implications for discrete encoding processes. Journal of Experimental Psychology: Human Perception and Performance 9, 785-806.

Frowein, H.W. and A.F. Sanders, 1978. Effects of visual stimulus degradation, S-R compatibility, and foreperiod duration on choice reaction time and movement time. Bulletin of the Psychonomic Society 12, 106-108.

Graf, P. and D.L. Schacter, 1985. Implicit and explicit memory for new associations in normal and amnesic subjects. Journal of Experimental Psychology: Learning, Memory and Cognition 11, 501-518.

Green, D.M. and J. Swets, 1966. Signal detection theory and psychophysics. New York: Wiley.

Grice, G.R., L. Canham and C. Schafer, 1982. Development of associative and perceptual interference. Perception & Psychophysics 32, 375-387.

Grice, G.R., R. Hunt, B. Kushner and C. Morrow, 1974. Stimulus intensity, catch trial effects, and the speed-accuracy tradeoff in reaction time: a variable criterion theory interpretation. Memory & Cognition 2, 758-770.

Grossberg, S., 1978. Do all neural models really look alike? A comment on Anderson, Silverstein, Ritz, and Jones. Psychological Review 85, 592-596.

Gopher, D. and A.F. Sanders, 1984. 'S-Oh-R: Oh stages! Oh resources!' In: W. Prinz and A.F. Sanders (eds.), Cognition and motor behavior. Heidelberg: Springer. pp. 231-253.

Hansen, J.C. and S.A. Hillyard, 1983. Selective attention to multidimensional auditory stimuli. Journal of Experimental Psychology: Human Perception and Performance 9, 1-19.

Hardzinski, M.L., 1980. A critical path analysis of negative interactions in choice reaction time experiments. Unpublished Ph.D. Thesis, University of Michigan, Ann Arbor, MI.

Hartley, A., 1981. Mental measurement of line length: the role of the standard. Journal of Experimental Psychology: Human Perception and Performance 7, 309-317.

Hassett, J., 1978. A primer of psychophysiology. San Francisco, CA: W.H. Freeman.

Hick, W.E., 1952. On the rate of gain of information. Quarterly Journal of Experimental Psychology 4, 11-26.

Holender, D., 1980. 'Interference between a vocal and a manual response to the same stimulus'. In: G.E. Stelmach and J. Requin (eds.), Tutorials in motor behavior. Amsterdam: North-Holland. pp. 421-431.

Hubel, D.H., 1963. The visual cortex of the brain. Scientific American 209, 54-62.

Hubel, D.H., 1979. The brain. Scientific American 241, 39-47.

Jonides, J., 1983. Further toward a model of the mind's eye's movement. Bulletin of the Psychonomic Society 21, 247-250.

Just, M.A. and P. Carpenter, 1980. A theory of reading: from eye fixations to comprehension. Psychological Review 87, 329-354.

Kahneman, D., 1968. Method, finding, and theory in studies of visual masking. Psychological Bulletin 20, 404-425.

Kettner, R., R. Shannon, R. Nguyen and R. Thompson, 1980. Simultaneous behavioral and neural (cochlear nucleus) measurement during signal detection in the rabbit. Perception & Psychophysics 28, 504-513.

Kinsbourne, M., 1982. Hemispheric specialization and the growth of human understanding. American Psychologist 37, 411–420.

Kolb, B. and I.Q. Whishaw, 1985. Fundamentals of human neuropsychology. New York: W.H. Freeman.

Kosslyn, S.M., 1981. The medium and the message in mental imagery: a theory. Psychological Review 88, 46–66.

Kutas, M., G. McCarthy and E. Donchin, 1977. Augmenting mental chronometry: the P300 as a measure of stimulus evaluation time. Science 197, 792–795.

LaBerge, D., 1973. 'Identification of two components of the time to switch attention: a test of a serial and a parallel model of attention'. In: S. Kornblum (ed.), Attention and performance, Vol. 4. New York: Academic Press. pp. 71–85.

Link, S.W., 1975. The relative judgment theory of two choice response time. Journal of Mathematical Psychology 12, 114–135.

Lockhead, G.R., 1972. Processing dimensional stimuli: a note. Psychological Review 79, 410–419.

Massaro, D.W. and M.M. Cohen, 1983. Evaluation and integration of visual and auditory information in speech perception. Journal of Experimental Psychology: Human Perception and Performance 9, 753–771.

Massaro, D.W., R.L. Venezky and G.A. Taylor, 1979. Orthographic regularity, positional frequency, and visual processing of letter strings. Journal of Experimental Psychology: General 108, 107–124.

McCarthy, G. and E. Donchin, 1981. A metric for thought: a comparison of P300 latency and reaction time. Science 211, 77–80.

McClelland, J.L., 1979. On the time relations of mental processes: a framework for analyzing processes in cascade. Psychological Review 86, 287–330.

McClelland, J.L. and D.E. Rumelhart, 1981. An interactive activation model of context effects in letter perception: part 1. An account of basic findings. Psychological Review 88, 375–407.

McClelland, J.L. and D.E. Rumelhart, 1985. Distributed memory and the representation of general and specific information. Journal of Experimental Psychology: General 114, 159–188.

McClosky, M., S.M. Sokol and R.A. Goodman, 1986. Cognitive processes in verbal-number production: inferences from the performance of brain-damaged subjects. Journal of Experimental Psychology: General 115, 307–330.

Meyer, D.E. and D.E. Irwin, 1981. On the time-course of rapid information processing. Paper presented at the annual meeting of the Psychonomic Society, Philadelphia, PA, Nov. Available as Tech. Rep. No. 43, Cognitive Science Program, University of Michigan, MI, June, 1982.

Meyer, D.E., A. Osman and S. Yantis, 1982. RT mixture distributions: evidence for finite-state models of response preparation. Paper presented at the annual meeting of the Psychonomic Society, Minneapolis, MN, Nov.

Meyer, D.E., S. Yantis, A.M. Osman and J.E.K. Smith, 1984. 'Discrete versus continuous models of response preparation: a reaction-time analysis'. In: S. Kornblum and J. Requin (eds.), Preparatory states and processes. Hillsdale, NJ: Erlbaum. pp. 69–94.

Meyer, D.E., S. Yantis, A.M. Osman and J.E.K. Smith, 1985. Temporal properties of human information processing: tests of discrete versus continuous models. Cognitive Psychology 17, 445–518.

Miller, J.O., 1976. Effects of stimulus probability on encoding mechanisms in information processing tasks. Unpublished Ph.D. thesis, University of Michigan, Ann Arbor, MI. Available as Human Performance Center Technical Report No. 56, University of Michigan, Ann Arbor, MI.

Miller, J.O., 1979. Cognitive influences on perceptual processing. Journal of Experimental Psychology: Human Perception and Performance 5, 546–562.

Miller, J.O., 1982a. Discrete versus continuous stage models of human information processing: in search of partial output. Journal of Experimental Psychology: Human Perception and Performance 8, 273–296.

Miller, J.O., 1982b. Effects of noise letters on decisions: discrete or continuous flow of information? Perception & Psychophysics 31, 227–236.

Miller, J.O., 1983. Can response preparation begin before stimulus recognition finishes? Journal of Experimental Psychology: Human Perception and Performance 9, 161–182.

Miller, J.O., 1985a. 'Discrete and continuous models of divided attention'. In: M.I. Posner and O.S.M. Marin (eds.), Mechanisms of attention: attention & performance XI. Hillsdale, NJ: Erlbaum. pp. 513–528.

Miller, J.O., 1985b. A hand advantage in preparation of simple keypress responses: reply to Reeve and Proctor, 1984. Journal of Experimental Psychology: Human Perception and Performance 11, 221–233.

Miller, J.O., 1987. Evidence of preliminary response preparation from a divided attention task. Journal of Experimental Psychology: Human Perception and Performance 13, 425–434.

Miller, J.O., 1988. Response compatibility effects in focused attention tasks: a same-hand advantage in response preparation. Perception & Psychophysics 43, 83–89.

Miller, J.O. and D.W. Bauer, 1981. Visual similarity and discrimination demands. Journal of Experimental Psychology: General 110, 39–55.

Miller, J.O. and R.G. Pachella, 1973. On the locus of the stimulus probability effect. Journal of Experimental Psychology 101, 227–231.

Miller, J.O. and R.G. Pachella, 1976. Encoding processes in memory scanning tasks. Memory & Cognition 4, 501–506.

Morton, J., 1969. Interaction of information in word recognition. Psychological Review 76, 165–178.

Mulder, G., A.B.M. Gloerich, K.A. Brookhuis, H.J. van Dellen and L.J.M. Mulder, 1984. Stage analysis of the reaction process using brain-evoked potentials and reaction time. Psychological Research 46, 15–32.

Neisser, U., 1967. Cognitive psychology. New York: Appleton-Century-Crofts.

Newell, A., 1973. 'Production systems: models of control structures'. In: W.G. Chase (ed.), Visual information processing. New York: Academic Press. pp. 463–526.

Newell, A. and H.A. Simon, 1972. Human problem solving. Englewood Cliffs, NJ: Prentice-Hall.

Norman, D.A., 1984. 'Theories and models in cognitive psychology'. In: E. Donchin (ed.), Cognitive psychophysiology: event-related potentials and the study of cognition. The Carmel Conferences, Vol. 1. Hillsdale, NJ: Erlbaum. pp. 119–138.

Norman, D.A. and D.G. Bobrow, 1975. On data-limited and resource-limited processes. Cognitive Psychology 7, 44–64.

O'Hara, W.P., L.R. Morris, M.G.H. Coles, C.W. Eriksen and N.M. Morris, 1981. Stimulus incompatibility and response competition: an EMG/RT analysis. Psychophysiology 18, 170–171.

Ollman, R., 1966. Fast guesses in choice reaction time. Psychonomic Science 6, 155–156.

Pachella, R.G., 1974. 'The interpretation of reaction time in information-processing research'. In: B. Kantowitz (ed.), Human information processing: tutorials in performance and cognition. Hillsdale, NJ: Erlbaum. pp. 41–82.

Pachella, R.G. and D.F. Fisher, 1969. Effect of stimulus degradation and similarity on the trade-off between speed and accuracy in absolute judgments. Journal of Experimental Psychology 81, 7–9.

Pachella, R.G. and J.O. Miller, 1976. Stimulus probability and same–different classification. Perception & Psychophysics 19, 29–34.

Penfield, W. and T. Rasmussen, 1950. The cerebral cortex of man. New York: Macmillan.

Pieters, J.P.M., 1983. Sternberg's additive factor method and underlying psychological processes: some theoretical considerations. Psychological Bulletin 93, 411–426.

Posner, M.I., 1978. Chronometric explorations of mind. Hillsdale, NJ: Erlbaum.

Posner, M.I. and P. McLeod, 1982. 'Information processing models – in search of elementary operations'. In: M. Rosenzweig and L. Porter (eds.), Annual review of psychology, Vol. 33. Palo Alto, CA: Annual Reviews Inc. pp. 477–514.

Posner, M.I., J.A. Walker, F.J. Friedrich and R.D. Rafal, 1984. Effects of parietal injury on covert orienting of attention. Journal of Neuroscience 4, 1863–1874.

Proctor, R.W. and T.G. Reeve, 1985. Compatibility effects in the assignment of symbolic stimuli to discrete finger responses. Journal of Experimental Psychology: Human Perception and Performance 11, 623–639.

Pylyshyn, Z.W., 1979. The rate of 'mental rotation' of images: a test of a holistic analogue hypothesis. Memory & Cognition 7, 19–28.

Pylyshyn, Z.W., 1981. The imagery debate; analogue media versus tacit knowledge. Psychological Review 88, 16–45.

Ragot, R. and B. Renault, 1981. P300 as a function of S–R compatibility and motor programming. Biological Psychology 13, 289–294.

Ratcliff, R., 1985. Continuous vs discrete processing: accumulation of partial information. Paper presented at the annual meeting of the Society for Mathematical Psychology, San Diego, CA, August.

Rayner, K. and A. Pollatsek, 1981. Eye movement control during reading: evidence for direct control. Quarterly Journal of Experimental Psychology 33A, 351–373.

Reed, A.V., 1973. Speed accuracy trade-off in recognition memory. Science 181, 574–576.

Reed, A.V., 1976. List length and the time-course of recognition in immediate memory. Memory & Cognition 4, 16–30.

Reeve, T.G. and R.W. Proctor, 1984. On the advance preparation of discrete finger responses. Journal of Experimental Psychology: Human Perception and Performance 10, 541–553.

Reeve, T.G. and R.W. Proctor, 1985. Nonmotoric translation processes in the preparation of discrete finger responses: a rebuttal of Miller's (1985) analysis. Journal of Experimental Psychology: Human Perception and Performance 11, 234–241.

Renault, B., R. Ragot, N. Lesevre and A. Remond, 1982. Onset and offset of brain events as indices of mental chronometry. Science 215, 1413–1415.

Ritter, W., R. Simson and H.G. Vaughan, Jr., 1983a. Event-related potential correlates of two stages of information processing in physical and semantic discrimination tasks. Psychophysiology 20, 168–179.

Ritter, W., H.G. Vaughan Jr. and R. Simson, 1983b. 'On relating event-related potential components to stages of information processing'. In: A.W.K. Gaillard and W. Ritter (eds.), Tutorials in event-related potential research: endogenous components. Amsterdam: North-Holland. pp. 143–158.

Ritter, W., R. Simson, H.G. Vaughan Jr. and M. Macht, 1982. Manipulation of event-related potential manifestations of information processing stages. Science 218, 909–911.

Rosenbaum, D.A., 1980. Human movement initiation: specification of arm, direction, and extent. Journal of Experimental Psychology: General 109, 444–474.

Rosenbaum, D.A., V. Hindorff and E.M. Munro, 1987. Scheduling and programming of rapid finger sequences: tests and elaborations of the hierarchical editor model. Journal of Experimental Psychology: Human Perception and Performance 13, 193–203.

Rumelhart, D.E., 1970. A multicomponent theory of the perception of briefly exposed displays. Journal of Mathematical Psychology 7, 191–218.

Rumelhart, D.E. and J.L. McClelland, 1982. An interactive activation model of context effects in letter perception: part 2. The contextual enhancement effect and some tests and extensions of the model. Psychological Review 89, 60–94.

Sanders, A.F., 1971. Probabilistic advance information and the psychological refractory period. Acta Psychologica 35, 128–137.

Sanders, A.F., 1977. 'Structural and functional aspects of the reaction process'. In: S. Dornic (ed.), Attention and performance, Vol. 6. Hillsdale, NJ: Erlbaum. pp. 3–25.

Sanders, A.F., 1980. 'Stage analysis of reaction processes'. In: G.E. Stelmach and J. Requin (eds.), Tutorials in motor behavior. Amsterdam: North-Holland. pp. 331–354.

Sanders, A.F., 1983. Towards a model of stress and human performance. Acta Psychologica 53, 61–97.

Sanders, A.F. and M.J.M. Houtmans, 1985. There is no central stimulus encoding during saccadic eye shifts: a case against general parallel processing notions. Acta Psychologica 60, 323–338.

Sanders, A.F., J.L.C. Wijnen and A.E. van Arkel, 1982. An additive factor analysis of the effects of sleep loss on reaction processes. Acta Psychologica 51, 41–59.

Schweickert, R., 1978. A critical path generalization of the additive factor method: analysis of a Stroop task. Journal of Mathematical Psychology 18, 105–139.

Shaffer, L.H., 1971. Attention in transcription skill. Quarterly Journal of Experimental Psychology 23, 107–112.

Shwartz, S., J.R. Pomerantz and H.E. Egeth, 1977. State and process limitations in information processing: an additive factors analysis. Journal of Experimental Psychology: Human Perception and Performance 3, 402–410.

Simon, H.A., 1969. The sciences of the artificial. Cambridge, MA: MIT Press.

Smith, E.E., 1968. Choice reaction time: an analysis of the major theoretical positions. Psychological Bulletin 69, 77–110.

Smith, E.E. and K.T. Spoehr, 1974. 'The perception of printed English: a theoretical perspective'. In: B.H. Kantowitz (ed.), Human information processing: tutorials in performance and cognition. Hillsdale, NJ: Erlbaum.

Smith, J.E.K., D.E. Meyer, S. Yantis and A. Osman, 1982. Finite-state models of reaction time: estimation of latency distributions. Paper presented at the annual meeting of the Society for Mathematical Psychology, Princeton, NJ.

Stanovich, K.E., 1979. Studies of letter identification using qualitative error analysis: effects of speed stress, tachistoscopic presentation, and word context. Journal of Experimental Psychology: Human Perception and Performance 5, 713–733.

Stanovich, K.E. and R.G. Pachella, 1977. Encoding, stimulus–response compatibility, and stages of processing. Journal of Experimental Psychology: Human Perception and Performance 3, 411–421.

Stanovich, K.E., R.G. Pachella and J.E.K. Smith, 1977. An analysis of confusion errors in naming letters under speed stress. Perception & Psychophysics 21, 545–552.

Sternberg, S., 1969a. The discovery of processing stages: extensions of Donders' method. Acta Psychologica 30, 276–315.

Sternberg, S., 1969b. Memory scanning: mental processes revealed by reaction-time experiments. American Scientist 57, 421–457.

Sternberg, S., 1975. Memory scanning: new findings and current controversies. Quarterly Journal of Experimental Psychology 27, 1–32.

Sternberg, S., S. Monsell, R.L. Knoll and C.E. Wright, 1978. 'The latency and duration of rapid movement sequences: comparisons of speech and typewriting'. In: G.E. Stelmach (ed.), Information processing in motor control and learning. New York: Academic Press.

Stroop, J.R., 1935. Studies of interference in serial verbal reactions. Journal of Experimental Psychology 18, 643–662.

Taylor, D.A., 1976. Stage analysis of reaction time. Psychological Bulletin 83, 161–191.

Terzuolo, C.A., 1970. 'Data transmission by spike trains'. In: F.O. Schmitt (ed.), The neurosciences: second study program. New York: Rockefeller University Press. pp. 661–671.

Thurstone, L.L., 1959. Measurement of values. Chicago, IL: University of Chicago Press.

Townsend, J.T., 1971. A note on the identifiability of parallel and serial processes. Perception & Psychophysics 10, 161–163.

Townsend, J.T. and F.G. Ashby, 1982. Experimental tests of contemporary mathematical models of visual letter recognition. Journal of Experimental Psychology: Human Perception and Performance 8, 834–864.

Turvey, M.T., 1973. On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. Psychological Review 80, 1–52.

Tversky, A., 1977. Features of similarity. Psychological Review 8, 327–352.

Underwood, G., 1978. 'Concepts in information processing theory'. In: G. Underwood (ed.), Strategies of information processing. London: Academic Press.

Van Santen, J.P.H. and D. Bamber, 1981. Finite and infinite state confusion models. Journal of Mathematical Psychology 24, 101–111.

Vaughan, H.G., Jr. and W. Ritter, 1973. 'Physiologic approaches to the analysis of attention and performance: tutorial review'. In: S. Kornblum (ed.), Attention and performance, Vol. 4. New York: Academic Press. pp. 129–154.

Vickers, D., 1970. Evidence for an accumulator model of psychophysical discrimination. Ergonomics 13, 37–58.

Vickers, D., 1979. Decision processes in visual perception. London: Academic Press.

Wickelgren, W.A., 1977. Speed–accuracy tradeoff and information processing dynamics. Acta Psychologica 41, 67–85.

Wood, C. and J.R. Jennings, 1976. Speed–accuracy tradeoff functions in choice reaction time: experimental designs and computational procedures. Perception & Psychophysics 19, 92–101.

Yeh, Y.Y. and C.W. Eriksen, 1984. Name codes and features in the discrimination of letter forms. Perception & Psychophysics 36, 225–233.

Yellott, J.I., 1967. Correction for guessing in choice reaction times. Psychonomic Science 8, 321–322.

Yellott, J.I., 1971. Correction for fast guessing and the speed–accuracy tradeoff in choice reaction time. Journal of Mathematical Psychology 8, 159–199.