TECHNICAL REPORT #27

TASK-EVOKED PUPILLARY RESPONSES AND COGNITIVE PROCESSING

JACKSON BEATTY
DEPARTMENT OF PSYCHOLOGY
UNIVERSITY OF CALIFORNIA AT LOS ANGELES

1 FEBRUARY 1980

PREPARED FOR OFFICE OF NAVAL RESEARCH
PHYSIOLOGY PROGRAM, ENVIRONMENTAL PHYSIOLOGY
CONTRACT N00014-76-C-0616

REPRODUCTION IN WHOLE OR IN PART IS PERMITTED FOR ANY PURPOSE OF THE UNITED STATES GOVERNMENT

This document has been approved for public release and sale; its distribution is unlimited.
Task-evoked pupillary responses and cognitive processing

**Invited address for the Society for Psychophysiological Research**
Cincinnati, Ohio 1979

**KEY WORDS** (Continue on reverse side if necessary and identify by block number)
- Pupillometry
- mental work
- processing resources
- reticular formation
- attention

**ABSTRACT** (Continue on reverse side if necessary and identify by block number)
The current status of the concept of processing resources in cognitive psychology is first presented. Difficulties with a purely behavioral theory of processing resources are then raised. Next the use of a physiological measure, task-evoked changes in pupillary diameter, as a reflection of general processing load is then proposed and the relevant research is reviewed. The measure is shown to yield consistent results across a range of qualitatively different mental tasks. Finally, a hierarchical physiological theory of processing resources is proposed.
TASK-EVOKED PUPILLARY RESPONSES AND COGNITIVE PROCESSING

Jackson Beatty
Department of Psychology and Brain Research Institute
University of California, Los Angeles

Invited Address
Society for Psychophysiological Research
Cincinnati, Ohio
1979
That the pupil of the eye dilates during mental activity has long been known in neurophysiology. For example, Bumke, the German neurologist, wrote nearly seven decades ago:

"every active intellectual process, every psychical effort, every exertion of attention, every active mental image, regardless of content, particularly every affect just as truly produces pupil enlargement as does every sensory stimulus..." (Bumke, 1911, as translated in Hess, 1975).

Only recently has this phenomenon been utilized as a tool in investigating the neurophysiological basis of human information processing. The pupillary dilations that accompany cognitive processes are indeed as pervasive a phenomenon as Bumke had indicated. They occur at short latencies following the onset of processing and subside quickly once processing is terminated. Perhaps most importantly, the degree of pupillary dilation appears to be a function of the information processing demands imposed by the particular task. For these reasons, the detailed investigation of pupillary movements in cognition is useful for the development of an understanding of the physiology of higher mental processes. In this paper, the physiological phenomenon of the pupillary response is employed to address a major issue in contemporary cognitive theory, the nature and structure of processing resources.

PROCESSING RESOURCES IN CONTEMPORARY COGNITIVE THEORY

Limitations on human information processing performance have remained a focus of interest in cognitive psychology since Welford (1952) and Broadbent (1958) proposed their single channel theories of processing limits. These models postulate that a central portion of the human information processing system is constrained to process only a single stream of information at a time. Such models were unable to account for apparent instances of parallel processing or resource sharing (Treisman, 1960, 1964, 1969). This situation suggested that a capacity theory might be more useful than single channel theory (Moray, 1967). In this view, there exists a general processing resource of fixed capacity that might be shared among several concurrent mental functions as long the joint demands do not exceed the available supply for the resource. Most influential among these early theories was that proposed by Kahneman (1973). Kahneman differed from Moray before him and Norman and Bobrow (1975) later by making some explicit assumptions concerning the biological properties of his general processing resource. Kahneman identified general capacity
with an attentional system, the amount of attention available at any moment in time being strictly limited, but the over time, the total attentional capacity of the system was postulated to vary with arousal. This formulation represented an attempt to interpret the early work of Kahneman and Beatty (1966, 1967; Beatty and Kahneman, 1966; Kahneman, Beatty and Pollack, 1967) who described systematic patterns of pupillary dilation that occurred during cognitive processing. Kahneman conceived of attention as being equivalent to "mental effort", which he approximately equated with the physiological concept of arousal. Kahneman did not distinguish between tonic and phasic arousal (Beatty, 1977, in press), nor was he concerned by the properties of the brain systems which regulated phasic activation changes during information processing. Despite these limitations, Kahneman's view of processing capacity was clearly physiological in nature, and not the abstract theoretical entity that was termed capacity by subsequent workers.

Norman and Bobrow (1975) provided a more formal statement of capacity theory and introduced the ideas of data and resource limits. For Norman and Bobrow, the pool of processing resources was purely an abstract construct devoid of surplus meanings which was used to explain tradeoffs in time sharing performance. They also emphasized the use of performance operating characteristics (POCs) as a method of determining the resource demands of an information processing tasks. A POC is a curve representing the performance of two time-shared tasks as the mix of available resources is shifted from one task to the other. Norman and Bobrow's elegant treatment of processing resources appeared to provide a solid basis for the study of resource demands, but the actual increase in understanding of information processing resources has been rather less than was hoped for. The problem is with the performance operating characteristics. Although a POC is analogous in some ways to the familiar receiver operating characteristic (Green and Swets, 1966), POCs are deceptive: Although they are easy to conceptualize, they are extremely difficult to determine empirically. In 1975, Norman and Bobrow were able to find very few examples of POCs in the published literature, but that could simply have reflected a general unawarness of the importance of these functions. However, Navon and Gopher (1979) could find few subsequent examples four years later.

The second difficulty with an analysis of processing resources that depends exclusively upon POC analysis is that these functions often appear ill behaved. Apparently trivial changes in the processing structure of one task, such as changing the sensory or response requirements, may have large effects on the pattern of two-task interference.
Such phenomena are suggestive of what Kahneman (1973) termed structural interference. The joint performance of such task pairs must be determined in important ways by factors other than the availability of the general processing resource. Moreover, other pairs of tasks when performed in combination show difficulty insensitivity: Changing the processing demands of one task has no effect upon the performance of the second task (Wickens, 1979). Difficulty insensitivity also indicates that a task is not limited by the availability of the general processing resource, since changing the resource demands of the second task is without an effect on performance.

Multiple Resource Models

The presence of substantial evidence for structural interference and task-specific performance limitations in two task interactions are difficult data for any simple capacity theory to explain. One solution that has been taken is to postulate multiple differentiated pools of processing resources, each pool being allocatable among competing tasks demands in a time sharing situation. Interference between tasks would result when the capacity of any one of the necessary pools is exhausted, even though capacity might remain in other, task-irrelevant pools. Such multiple resource capacity models have recently been proposed by Sanders (1979), Wickens (1979) and Navon and Gopher (1979). But the difficulty with the proposed multiple capacity theories is in the identification and specification of the multiple resource pools. Sanders (1979) quite reasonably suggests that three pools might exist, one having to do with encoding, one with choice and one with response execution. Sanders bases this hypothesis on the earlier work with the Sternberg additive factor method (Sanders, 1977). Wickens (1979) lists several possible processing resources but provides no empirical justification for their acceptance. Navon and Gopher (1979) avoid the substantive question of resource identification, proposing instead an elegant theoretical approach based upon the assumption that processing resources may be analysed by microeconomic theory. In their model, information about the several processing resources must be derived from a detailed analysis of multiple performance operating characteristics. However, given the difficulty in obtaining empirical POCs, any complex theory that depends solely upon such functions to define its essential features is in a difficult situation. For multiple capacity models, this problem is much more severe. To determine the structure of processing resources by POC analysis, the number of POCs required would seem to be proportional to the square of the number of relevant individual pools. This suggests that an alternative to POC analysis must be found if multiple resource models of the human information processing system.
are to be considered seriously.

PUPILLARY RESPONSES IN COGNITIVE PROCESSING

One approach to understanding processing resources is to utilize converging sources of information concerning cognitive operations. Neuroanatomy, neurophysiology and psychophysiology are all relevant, in that these disciplines provide biological data concerning the structure and function of the human nervous system, which supports cognition. In the present paper, information derived from changes in task-evoked pupillary responses that occurring during cognitive processing is employed to suggest a neurophysiological theory of processing resources.

Pupillary movements are typical of psychophysiological responses, in that they are systemic physiological signs that are related to psychological variables. Usually psychophysiological responses are of intermediate status in psychobiology, often lacking clear meaning for either neurophysiology or cognitive theory. In order to contribute to an understanding of brain mechanisms in human information processing, psychophysiological responses require two definitions, one at the level of cognitive processes and the other at the level of neurophysiology. When such definitions are firm, psychophysiological responses can be uniquely valuable in developing a deeper understanding of the physiology of higher mental functions.

An understanding of the meaning of task-evoked pupillary responses at the level of cognitive theory may be obtained by analysing the behavior of these responses in differing types of information processing functions. Task-evoked pupillary responses have now been measured in a wide range of cognitive tasks. This literature is reviewed in the following sections.

Short-term Memory

The demands placed upon short-term memory formed an initial and enduring problem in the pupillometric study of information processing. Kahneman and Beatty (1966) presented the first pupillometric analysis of the processing demands encountered in a short-term memory task. Strings of 3 to 7 digits were auditorily presented at the rate of 1 per sec. After a two second pause, subjects were required to repeat the digit string at the same rate. Under these conditions, pupillary diameter increases with the presentation of each digit, reaching a maximum in the pause preceding report. During report, pupillary diameter decreases with each digit reported, reaching baseline levels
The magnitude of the peak pupillary dilation during the pause between input and output is an increasing function of string length. In unpublished work, Kahneman and Beatty observed that if the subject were requested to repeat the string a second time immediately after reporting the final digit, the pupil immediately redilates to the peak diameter for that string and then decreases with each digit spoken until the entire string has been reported for the second time. Beatty and Kahneman (1966) demonstrated that a similar pupillary function is obtained when a string of items is recalled from long-term memory for report: on request to report, a large pupillary dilation is observed as information is retrieved from long-term memory and the response is organized. As each digit in the string is spoken, pupillary diameter decreases, reaching baseline levels at report of the last digit.

The slope of this task-evoked pupillary response is determined by the difficulty of the to-be-remembered information as indexed by memory span for different types of items. Kahneman and Beatty (1966) tested three conditions: recall of four digits, recall of four unrelated nouns, and transformation of a four digit string by adding one to each item. The slope of the pupillary response during input was lowest for the least difficult items, the strings of four digits that were to be simply repeated. A steeper slope was observed for the strings of four words. The greatest slope was obtained for the most difficult task, digit string transformation. Thus both item difficulty and number of items affect the pupillary response in short-term memory tasks.

The idea that the task-evoked pupillary response provides a physiological measure of demand for a general processing resource received direct support in a subsequent experiment by Kahneman, Beatty, and Pollack (1967) in which both pupillometric and behavioral interference methods were utilized to assess processing load in the four-digit add-one memory transformation task. Using a secondary task of visual target detection, it was found that the behavioral estimate of processing load and the amplitude of the pupillary response were in close agreement. A series of controls ruled out any peripheral interference of the pupillary dilations themselves on performance of the secondary task.

The use of rehearsal strategies that improve performance on a short-term memory task acts to reduce the amplitude of the task-evoked pupillary response. Kahneman, Onuska, and Wolman (1968) presented subjects with strings of nine digits, either at a uniform rate of 1 per second, or with a temporally-imposed three-digit grouping (.5 sec separating digits within a group and 2.0 sec separating
groups). The grouped mode of presentation had previously been shown to materially increase digit span (Ryan, 1967), presumably by breaking the string into more codable units or chunks (Miller, 1956). The pupillometric data reflected the experimentally induced differences in processing strategy: a steady monotonic increase in pupillary diameter accompanied presentation of the digits at the uniform rate, whereas waves of dilation during presentation and constriction during the inter-group pauses characterized the grouped presentation condition. Thus, the task-evoked pupillary response appears to reflect changes in information processing demands induced by processing strategies that affect performance.

This hypothesis found further support in Peavler's (1974) study of information overload. The capacity of short-term memory for strings of unrelated digits is approximately seven (Miller, 1956). Peavler measured the task-evoked pupillary response for strings of 5, 9 and 13 digits, which were randomly intermixed in presentation. During presentation of the strings pupillary diameter increases as an increasing function of memory load for digits 1 through 7. At the seventh or eighth digit, the pupillary response asymptoted; no further dilation was observed. Peavler interpreted these data as indicating that as long as some information-processing capacity remains, increasing memory load is reflected by increasing pupillary dilation. But once the limits of capacity are exceeded further increases in task demands no longer yield increased pupillary dilation.

Language Processing

Several levels of language processing have been studied pupillometrically. At the most molecular level, Beatty and Wagoner (1978) used an experimental method developed by Posner (Posner & Mitchell, 1967; Posner & Boles, 1971) to study the visual encoding of single letters. In their first experiment subjects were required to judge whether or not a pair of visually-presented letters had the same name. Individual letters were presented in either upper or lower case type. Thus, two types of letter pairs could be judged to be the same by the name criterion. If both letters are presented in the same case (e.g. AA or aa), only the physical features of the letters need by analyzed to reach the correct judgment. If they differ in case (e.g. Aa or aB) then, in addition to analyzing the physical features, a second step of name code extraction must be performed. Although the task-evoked pupillary responses were small in this simple task (on the order of $.1 mm), they did reflect the extra processing step required for name code extraction. Significantly larger responses were obtained for letter
pairs that differed in case.

In a second similar experiment, Beatty and Wagoner examined three levels of character encoding by requiring the use of a third, higher order category for classification (vowels and consonants). Thus some letter pairs could either be physically identical, identical in name, or identical in category membership (e.g. Ae or BK). Again, the task-evoked pupillary response was responsive to the additional processing load required to perform the letter-matching task at each level.

Ahern (Ahern, 1978; Ahern & Beatty, in press) undertook two experimental investigations involving language processing as part of a larger research program on individual differences in intelligence. The first of these experiments examined task-evoked pupillary responses in the perception and comprehension of words. Subjects were required to judge pairs of words as similar or different in meaning. The first word of each pair was drawn from either the easiest or the most difficult items of one of three psychometric vocabulary tests. The second word, presented two seconds later, was either a synonym of the first, or quite different in meaning. In this experiment, the ease of accessing lexical information was reflected in the pupillary response. A dilation of approximately .1 mm followed the presentation of the easy target words, whereas the dilation to the difficult target words was twice as large. A second dilation followed the presentation of the comparison word, yielding pupillary dilations of .30 and .35 mm, respectively during the judgment period. Thus, larger pupillary dilations accompany the processing of psychometrically more difficult vocabulary items.

At the most molar level, task-evoked pupillary responses have been studied as subjects processed meaningful sentences of differing complexity. Ahern (Ahern, 1978; Ahern & Beatty, 1979), employing Baddeley's Grammatical Reasoning Task (Baddeley, 1968), presented sentences of the form "A follows B" or "B precedes A" after which an exemplar "AB" or "BA" was given. The task was to determine whether the sentence correctly described the exemplar. Sentences differed in grammatical complexity, being active-positive, active-negative, passive-positive or passive-negative. Although these sentences differed in length, sentence duration was held constant by using computer presentation of digitized natural speech. In this task, increasing dilation was observed during the presentation of the sentence and the exemplar, which peaked during the decision interval. The amplitude of these responses averaged approximately .40 mm and differed significantly as a function of grammatical complexity, with the longer, more complex sentences eliciting larger pupillary responses. Thus, in processing
sentences for meaning, the task-evoked pupillary response is sensitive to variations of sentence complexity.

Wright and Kahneman (1971) have also applied pupillometric measurements in a sentence processing task. Subjects were presented with complex sentences of the form: "The qualified managing director was recently sensibly appointed by the expanding successful company." Subjects were either required to repeat the sentence or to answer a question of the form: "Who appointed the director in this sentence?" The query was posed either before or after the sentence was presented. When the task was to repeat the sentence, the task-evoked pupillary response increased as the sentence was presented, and peaked during the retention interval (3 or 7 sec), reaching a maximum dilation of approximately .30 mm. When the question was posed after the pause, peak dilation during the pause was approximately .20 mm and was followed by another dilation as the answer to the question was formed. The peak of this dilation was approximately .40 mm with respect to pre-sentence baseline. When the question was posed before sentence presentation, the task-evoked pupillary response rose more gradually, but increased rapidly when the relevant portion of the sentence was presented, indicating organization and processing of the answer to the query. No evidence of processing of phrase boundaries was observed, but, as Wright and Kahneman commented, their sentences were not representative of those naturally occurring in spoken English.

Reasoning

Mental arithmetic has been used as an example of a complex reasoning problem by several investigators. Hess and Polt (1964), in their initial and influential article on pupillary signs of mental activity, measured pupillary diameter as 5 subjects solved 4 multiplication problems, ranging in difficulty from 7 X 8 to 16 X 23. For each of the subjects and each of the problems, pupillary diameter increased from the moment of problem presentation until the point of solution. Hess and Polt reported these data as percentage dilation, not as absolute values. Across subjects, the percentage dilation was perfectly ordered by presumed problem difficulty.

Bradshaw (1968) reported similar results for 6 subjects performing mental division problems at two levels of difficulty. Pupillary diameter increased during problem solving until the point of solution, peak dilation being larger for the more difficult problems. Similarly, Payne, Parry, and Harasymiw (1968) described a monotonic relation between mean pupillary diameter and problem difficulty, but noted that this relationship is markedly nonlinear with
respect to difficulty scales based upon percent correct solution, time to solution or subjective rating of difficulty. Pupillary diameter in mental multiplication appears to peak rapidly as a function of difficulty, with more difficult problems requiring more time until solution is reached.

These results were subsequently replicated by Ahern (Ahern, 1978; Ahern & Beatty, 1979a, in press). Three levels of problem difficulty were employed, ranging from multiplying pairs of 1-digit numbers to multiplying pairs of 2-digit numbers ranging between 11 to 20. In this task an initial dilation of approximately .15 mm accompanies the encoding and storage of the multiplicand. The second and major dilation follows presentation of the multiplier and continues through problem solution. Both the amplitude and latency of this latter dilation increase as a function of problem difficulty. In the most difficult condition, the response appears to asymptote at approximately .50 mm.

Perception and Attention

Perceptual processes appear to proceed quite effortlessly and place rather little demand upon the limited capacity of the human information processing system (Kahneman, 1973). Thus, Wickens (1974) was unable to observe a secondary task decrement when a sensory signal-detection task was imposed as the primary task in an experiment investigating the distribution of time-shared processing capacity. The information processing load involved in the detection of weak signals is evidently small.

It is therefore of interest that small but reliable pupillary dilations accompany the detection of both visual and acoustic signals at near threshold intensities. Hakerem and Sutton (1966) provided the first pupillometric analysis of processing load in perceptual detection. Subjects viewed a uniform visual field upon which brief increments in luminance could be imposed with the left eye as pupillary diameter of the right eye was measured. When the magnitude of the intensity increment was adjusted to yield 50% correct detection, all vestiges of the flash-induced light reflex disappeared. Under these conditions, a clear pupillary dilation of approximately .10 mm was observed if and only if a presented target was detected. Although of small magnitude, these task-evoked responses to detection were highly reliable.

Beatty and Wagoner (1975) extended Hakerem and Sutton's finding to audition, using weak 100 msec 1kHz sinusoidal acoustic signals presented against a background of white
noise. Signals were presented on each trial with a probability of .50. After each trial, the subject rated his certainty that a target had or had not been presented (Green & Swets, 1966). For signal present trials, the magnitude of the task-evoked pupillary response was largest for signals judged present with high certainty and smaller for signals judged present with low certainty. The task-evoked pupillary response was absent on no-signal trials. These results fully confirm those reported for visual detection by Hakerem and Sutton.

Task-evoked pupillary responses are also observed in perceptual discrimination tasks, in which a presented stimulus must be compared against memory and a judgment rendered. For difficult discriminations, the amplitude of the response and hence the inferred processing load may be larger than that observed for simple detection. Kahneman and Beatty (1967) reported the first study of the pupillary response in perceptual discrimination. On each trial the subject first heard a standard tone of 850 Hz, which was followed 4 sec later by a comparison tone. The comparison was one of eleven frequencies, ranging between 820 and 880 Hz in 6 Hz steps. The subject's task was to judge whether the comparison tone was higher or lower in pitch than the standard. The difficulty of the discrimination is inversely related to the difference in pitch between the comparison and the standard. The amplitude of the response to the comparison tone varies as a function of discrimination difficulty, from approximately .10 mm for easy to .20 mm for difficult discriminations.

Processing efficiency in memory-dependent perceptual discrimination tasks is known to deteriorate if the task is prolonged and the number of discriminations required per minute is relatively high (Parasuraman, 1979; Parasuraman & Davies, 1977). One theory to explain this vigilance decrement is that the quality of information processing deteriorates with boredom under such conditions. To test this idea, Beatty (in press) measured task-evoked pupillary responses to non-target stimuli in an auditory vigilance task. Non-target stimuli were 50 msec 1 kHz tone bursts, presented at intervals of 3.2 sec. Randomly intermixed were target stimuli, which were attenuated by 3.5 db. Subjects reported the detection of targets by depressing a microswitch. Under these conditions, the efficiency of target/non-target discrimination dropped as a function of time over the 48 minutes of the task (from 84% to 67%). The amplitude of the task-evoked pupillary response showed a similar reduction, from approximately .07 mm in the first third of the task to .04 mm in the last. This reduction in the task-evoked pupillary response was interpreted as indicating that the adequacy of processing deteriorates with time in the vigilance task, thereby reducing processing
demands imposed by the discrimination task.

In summary, perceptual processes appear to place small but pupillometrically measurable demands for general processing capacity. Variations in task requirements result in reliable changes of the task-evoked pupillary response. In tasks requiring sustained attention, the pupillary response indexes the reduction in processing quality that results in performance decrements under conditions of prolonged, monotonous watch-keeping. Thus, in perceptual tasks, where behavioral methods are insensitive to variations of processing load, the task-evoked pupillary response reveals systematic changes with task variables.

Inter-task Comparisons

Task-evoked pupillary responses have been measured in a wide variety of cognitive tasks by a number of investigators, yielding a remarkably consistent body of data. Tasks which place large demands on the information processing system, judged behaviorally, subjectively or by an analysis of task requirements, elicit large task-evoked pupillary responses. Less demanding tasks elicit smaller responses. An intriguing possibility is that the measurement of task-evoked pupillary responses associated with cognitive function might provide a common metric for the assessment and comparison of information-processing load in tasks that differ substantially in their functional characteristics. Underlying this possibility is the idea that CNS activation is a general resource that is allocated among cortical processes as required. It therefore seems reasonable to compare qualitatively different mental processes in terms of activation requirements.

This possibility is strengthened by the finding that the magnitude of the task-evoked pupillary responses during cognitive processing is independent of baseline pupillary diameter over a physiologically reasonable range of values (Bradshaw, 1969, 1970; Kahneman, Beatty and Pollack, 1967). It is therefore possible to compare the absolute values of the task-evoked dilations reported from different laboratories for qualitatively different tasks. Figure I presents such a quantitative comparison, giving the approximate peak amplitude of the task-evoked pupillary response measured from published figures for each of the tasks detailed above, subject only to the constraint that the data are not confounded by the effects of overt responding.

The leftmost panel of Figure I presents peak dilations for short-term memory tasks. The data for short-term retention of digits are the average of the values obtained
The next panel summarizes the literature on language processing. The peak value for the letter matching task (Posner and Mitchell, 1967) is the average of both experiments published by Beatty and Wagoner (1978). Sentence encoding is from Wright and Kahneman (1971). The other data from the Wright and Kahneman experiment are not included, as the peak of the task-evoked pupillary response occurred during verbal responding. All other values for language processing tasks are taken from Ahern (1978). Word encoding is the response to the presentation of the first word in the synonyms judgment task. The values for easy and hard word matching are the peak response during the judgment period following presentation of the second word in that task. The value for grammatical reasoning is the average of the four types of sentences in Baddeley’s Grammatical Reasoning Task (Baddeley, 1968).

The third panel presents data from the mental multiplication task used as an example of complex reasoning. Only Ahern (1978) has presented task-evoked pupillary responses for this task which are necessary for comparative peak measurement. Multiplicand storage is the amplitude of the peak response to the first item in the multiplication task. The other three values are the peak amplitudes attained during problem solution.

The rightmost panel presents data for perceptual tasks. The visual detection data are from Hakerem and Sutton (1966) and the auditory detection data are from Beatty and Wagoner (1975). The discrimination data are taken from Kahneman and Beatty (1967).

Several points concerning these data deserve mention. First, the data are very tolerant of the exceedingly stringent demands placed upon them in comparing absolute dilation values across experiments. Usually rescaling of some sort is required for physiological data to remove individual differences in responsivity (Johnson and Lubin, 1972). No such rescaling was undertaken here. The data plotted are absolute peak dilations obtained from different groups of subjects performing a wide range of cognitive tasks under varying experimental conditions in different laboratories. The data base is the entire relevant published literature. The times at which these experiments were performed span a period of at least 12 years. Second, the data plotted in Figure 1 are internally consistent. No abnormally large or small values are present. Third, the ordering of these values corresponds quite closely to an ordering of these tasks using other criteria of information-processing load. The short-term memory tasks
cover a large range of values, depending on the number of items held for recall. Similarly in language processing the sentence comprehension tasks yield large pupillary dilations whereas the simpler word and letter matching tasks elicit much smaller values. The mental multiplication tasks again span a wide range of values, each appropriate to the difficulty of the particular problem. Finally, the perceptual tasks, which behavioral techniques indicate impose negligible processing load, are associated with small task-evoked pupillary responses.

Figure 1 provides strong evidence that the task-evoked pupillary responses measures what cognitive theorists term the utilization of processing resources (Norman and Bobrow, 1975; Sanders, 1979). Such a definition of the response as an indicator of cognitive processing seems relatively secure. In fact, this physiological phenomenon provides a primary reason for retaining some form of a concept of a unitary processing resource, an idea which has recently come under attack because of a failure of single resource models to adequately predict two-task interactions when timesharing (Navon and Gopher, 1979). The evidence outlined above supports the idea that a global indicator of central nervous system function varies directly with changes in task parameters that affect information processing load. However, several alternative hypotheses have been suggested which require consideration.

Alternative Hypotheses

The definition of the task-evoked pupillary response as an indicator of demand for information processing resources depends not only upon the clear demonstration that the response varies with relevant task parameters, but also upon evidence excluding the involvement of other, potentially confounding variables. In this context, it is important to distinguish between factors affecting the task evoked pupillary response and those that affect tonic or baseline pupillary diameter. Whereas basal diameter is strongly influenced by a wide variety of factors, including ambient illumination, general arousal, emotional activation, fatigue and similar systemic variables, the task evoked pupillary response remains relatively uninfluenced by these factors. The effects of a number of such possibly confounding variables upon the task evoked pupillary response in information processing have been investigated experimentally.

The possibility that the task-evoked pupillary response may reflect emotional rather than attentional or cognitive factors has received considerable attention and may be rejected on several grounds. First, task-evoked pupillary
Task-Evoked Pupillary Response

responses have been reliably observed in tasks in which it is difficult to hypothesize emotional involvement. For example, in an auditory selective attention task (Beatty, in press), small but consistent task-evoked pupillary responses of approximately .02 mm were observed following presentation of non-signal tone on the attended channel, whereas no responses were present to tones on the unattended channel. Since the event rate in that experiment was 3/sec, an average of 1.5 dilations were obtained each second for the duration of the testing procedure (approximately 15 minutes). To consider these dilations to be a long string of stimulus selective, high speed anxiety reactions seems implausible. Similar arguments might be made for a variety of other simple cognitive tasks that would not appear to be anxiety-provoking for any subject.

A second reason to reject an interpretation of the task-evoked pupillary responses as an expression of emotional factors is that these responses do not habituate as long as the task performance of the subject is stable. For example, in an auditory signal detection task, the amplitude of the task-evoked pupillary response to detected signals (Beatty, unpublished data), remains stable from the beginning to the end of a one hour experimental test. This suggests that the pupillary responses reflect brain processes that are obligatory and necessary for task performance, rather than any emotional factors such as task anxiety.

Another reason to reject an emotion hypothesis as an explanation of task-evoked pupillary responses is based upon an investigation of individual differences in pupillary response amplitude. Ahern (Ahern, 1978; Ahern and Beatty, 1979a, in press) obtained pupillary responses in 40 university undergraduates tested in four cognitive tasks. Subjects were divided into two groups designed to differ in cognitive ability (as indexed by scores on the Scholastic Aptitude Test). There was a significant correlation between a psychometric measure of subject intelligence and the amplitude of the task-evoked pupillary responses in the cognitive tasks. More intelligent individuals were able to perform these tasks with less cognitive loading. There was no significant correlation between the amplitude of the pupillary response and either state or trait anxiety (Spielberger, 1968). Differences in amplitude of the task-evoked pupillary response between individuals appears to be a function of differences in cognitive ability rather than emotionality.

Fourth, in his study of information overload, Peavler (1974) also addressed the question of interpreting pupillary dilation as an indication of emotional factors. Peavler reasoned that his data are incompatible with any
interpretation of the task-evoked pupillary response as a reflection of task anxiety or other emotional responses to the testing situation. If the task-evoked pupillary response reflected emotional factors due to fear of performance failure, then a large dilation should accompany the presentation of the later digits in the 13 digit strings, since only at this time could the subjects know that the limits of capacity would be exceeded and that their performance could not be perfect. No such dilations to information overload were observed.

These four lines of argument suggest that emotional factors are relatively unimportant as determinants of the pupillary responses observed in carefully controlled information processing tasks. Although emotional factors are well known for their expression in the autonomic nervous system, the effects of emotional arousal are longer lasting than the brief phasic responses evoked by cognitive activity. Thus changes in emotionality are more likely to affect the tonic or basal pupillary diameter and not the phasic responses under discussion here.

The effect of motor responding as a confounding variable has also received attention. Kahneman, Peavler and Onuska (1968, Experiment 1) examined the effects of verbalization on the task-evoked pupillary response in a short-term memory task at two levels of difficulty. Subjects listened to a string of 4 digits which they were to repeat or transform by adding one (Kahneman & Beatty, 1966). After presentation, they either repeated the appropriate response twice at the rate of 1 digit per second, or mentally produced the response in the first interval and verbally produced it in the second. The more difficult digit transformation task yielded systematically larger pupillary dilations regardless of verbalization condition. The form of the response was unaltered in the absence of verbalization. The effect of verbalization was to increment the amplitude of the task-evoked pupillary response. These data are consistent with the hypothesis that the organization of an overt motor act places demands upon information processing capacity that are reflected in the task-evoked pupillary response.

In a second experiment, Kahneman et al. tested the effects of motivation on the pupillary response to the same short-term memory task by varying monetary incentives associated with correct performance on different trials (Kahneman, Peavler & Onuska, 1968, Experiment 2). Increasing the incentives had no effect on performance, nor did it affect the task-evoked pupillary response during the performance of either of the short-term memory tasks.

Clark and Johnson (1970) tested the possibility that
the task-evoked pupillary response in short-term memory experiments might result from the subject's knowledge about the results of previous pupillometric studies and the demand characteristics of the experiment. Varying these expectations had no effect on the pupillary response, which conformed to the pattern previously reported by Kahneman and Beatty (1966).

Taken together, these lines of evidence lend support to the original hypothesis of Kahneman and Beatty (1966) that the task-evoked pupillary response reflects the momentary level of processing demands and not an artifact of non-cognitive confounding factors.

A NEUROPHYSIOLOGICAL THEORY
OF THE STRUCTURE OF PROCESSING RESOURCES

The search for a neurophysiological definition of the task-evoked pupillary response should begin with a consideration of the musculature of the human iris and the pathways that innervate it. The iris surrounding the pupil contains two opposing muscle groups. The dilator pupillae is a set of radially oriented fibers that act to dilate the pupil when they contract. In mechanical opposition to these fibers are the sphincter pupillae, which act to constrict the pupil as they contract. Pupillary diameter is determined by the differential excitation of these two sets of muscles.

Both muscle groups are controlled peripherally by the autonomic nervous system. The dilator pupillae are sympathetically innervated. Sympathetic efferents originating in the ciliospinal center of Budge (located in the thoracic spinal cord) exit the spinal column and ascend to the superior cervical ganglion, where they synapse with fibers that innervate the dilator pupillae directly. Increased sympathetic discharge along this pathway acts to contract the dilator pupillae and thereby dilate the pupil.

In contrast, the sphincter pupillae are parasympathetically innervated. Peripheral control of these muscles originates in the nucleus Edinger-Westphal (located within the third nerve nucleus at the level of the midbrain), which sends parasympathetic efferents to the ciliary ganglion. The ciliary nerves proceed from the ciliary ganglion to the sphincter pupillae, where they synapse. Increased parasympathetic activity along this pathway acts to contract the sphincter pupillae and constrict the pupil.

Although the peripheral innervation of the pupil is
mediated by the autonomic nervous system, it would be a mistake to conclude that only traditional autonomic functions can affect pupillary diameter. Pathways affecting pupillary diameter have been traced from a large number of brain regions, including cerebral cortex, diencephalon and brainstem (Lowenstein and Lowenfeld, 1962). Thus, a variety of brain systems could affect pupillary diameter in the course of their operation. For this reason, the next step in attempting a neurophysiological definition of the task-evoked pupillary responses is to search for plausible neurophysiological systems that both could function as a processing resource and also be anatomically and physiologically capable of affecting the pupillary control system in the manner described in the previous section. Such an hypothesis would provide a plausible mechanism for relating cognitive processing and pupillary movements at the level of neurophysiology.

The Reticular Hypothesis

The reticular formation is a diffusely organized densely interconnected region of the brainstem core, extending from the medulla though the midbrain. Phylogenetically very old, this system is interconnected with both afferent sensory and efferent motor systems. Long ignored by both neuroanatomists and neurophysiologists, the reticular core became the subject of intensive research following Moruzzi and Magoun's (1949) discovery that electrical stimulation in this structure activates and alerts the cerebral cortex. This finding led to the idea of an ascending reticular activation system that is intimately involved the attentional processes and the regulation of conscious activity (Lindsley, 1961). It is of interest that one of the classical effects of reticular activation is pupillary dilatation, and that pupillary measures are used extensively in determining the state of reticular function in experimental animals (Moruzzi, 1972).

Despite abundant evidence that the reticular system exerts a profound influence on the activity of the cerebral cortex, until recently little was known concerning either the effects of reticular stimulation on single cortical units or the mechanism of reticulo-cortical interaction (Schlag, 1974). This situation was changed when Skinner (1979; Skinner and Yingling, 1977; Yingling and Skinner, 1977) described a mechanism for reticular modulation of single unit activity in primary sensory cortex that may be extended to provide a general view of reticulo-cortical interactions. This model is quite relevant for a neurophysiological theory of processing resources, since it provides a mechanism for regulating the dynamics of forebrain information processing functions.
The sense of Skinner's suggestion is that the ascending reticular activating system exerts its effect on the forebrain by modulating the excitability of thalamocortical projections. Most, if not all, afferent fibers entering the cerebral cortex originate in the thalamus. Correspondingly, each region of cerebral cortex sends efferents back to the thalamus, terminating in the same thalamic region that gives rise to its afferents. Thus, a fundamental mode of connection in the forebrain is the highly specific and reciprocal system of thalamocortical connection. Wrapped around the lateral, anterosuperior and anteroinferior aspects of the thalamus is a thin neuronal shell, the reticular nucleus of the thalamus (Carpenter, 1978). Most of the axons from this shell enter the dorsal thalamus where they synapse. Skinner has assembled a variety of evidence to suggest that these neurons are inhibitory and that they act to gate thalamic input to the cerebral cortex. Stimulation of the ascending reticular formation of the midbrain elicits a blanket inhibition of neurons in the reticular nucleus of the thalamus, thereby opening the gate between thalamus and cortex. Therefore the effect of activity in the ascending reticular activating system is to globally facilitate thalamocortical interactions. In addition, there are numerous topographically organized projections from granular frontal cortex to the reticular nucleus. These projections are excitatory and act to close the gate between thalamus and cortex in the region of their termination. This provides the granular frontal cortex with the capacity to modulate the pattern of cortical activity on a regional basis. The momentary level of thalamocortical interaction in any cortical region reflects the balance of relatively specific excitatory frontal influences and generalized inhibitory input of the ascending reticular activating system upon the inhibitory gating neurons of the thalamic reticular nucleus.

The effect of these modulating influences must be to regulate the ability of the forebrain to develop and maintain orderly dynamic patterns of neural interactions that form the basis of cognitive activity. It is probably important that the excitability of the thalamocortical system is dynamically modulated, so that distant connections may be formed during intense information processing and not otherwise. Viewed in this light, the two great structures of the forebrain, the cerebral cortex and the thalamus, form the basis of the cognitive system of the human brain. These are the neurological structures that subserve the function of processing resources in cognitive theory. The utilization of this system is reflected in the state of generalized facilitation of thalamocortical interaction determined by the reticular activating system, probably in response to descending influences from frontal cortex. Since the pupil reflects in part the influence of the
reticular activating system, task-evoked pupillary responses may be hypothesized to provide a physiological indication of the momentary extent of generalized facilitation of the thalamocortical system.

The connections by which activity in the reticular core might affect the pupillary control systems are several. On the sympathetic side, the reticular system is known to modulate activity in the ciliospinal center of Budge. Similarly, in the parasympathetic pupillary control system the nucleus Edinger-Westphal is known to receive direct modulating input from the reticular formation. Fibers originating in the nucleus cuneiformis of the midbrain reticular formation also project directly to the nucleus Edinger-Westphal (Edwards and DeOlmos, 1976). In addition, pathways have been traced from the nucleus pontis oralis in the pontine reticular formation (Graybiel, 1977).

However, not all central inputs to the pupillary system arise from the reticular core. The nucleus Edinger-Westphal is also known to receive inhibitory afferent input from telencephalic and diencephalic structures. A major pathway from prefrontal cortex has been demonstrated by DeVito and Smith (1964) using the Nauta method (1957) for tracing axonal degeneration following prefrontal lesion. Additional input from frontal cortex to nucleus Edinger-Westphal is mediated by the massa intermedia (Lowenstein and Loewenfeld, 1962). Thus, task-evoked pupillary responses might also reflect thalamo-cortical activity directly.

Electrophysiological evidence has demonstrated the sensitivity of pupillary movements to changes in reticular activity. For example, Bonvallet and Zbrozyna (1963; Zbrozyna and Bonvallet, 1963) simultaneously recorded electrocortical activity, pupillary diameter, and multiple unit activity in the cervical sympathetic nerve and the short ciliary nerves during electrical stimulation of the pontine-mesencephalic reticular core. Stimulation produced cortical desynchronization and an immediate and profound pupillary dilation, which was the result of both parasympathetic inhibition and sympathetic activation. Thus both neuroanatomical and neurophysiological evidence suggests that the pupillary dilations observed during cognitive processing might reflect reticular and frontal influences acting upon the thalamocortical gating system of the thalamic reticular nucleus.

SOME CONCLUSIONS

These lines of evidence lead to a new and interesting view of both the nature of cognitive processes and their neuralphysiological basis. From the viewpoint of cognitive
theory, a solution to the question of one or many processing resources is proposed. The idea of a single, general and relatively unspecified processing resource was first suggested by Moray (1967), amplified by Kahneman (1973) and formalized by Norman and Bobrow (1975, 1976). In all of these theories, two or more tasks could be performed concurrently if their aggregate demand for the general processing resource does not exceed the total capacity available. However, matters in fact appear to be more complicated than this hypothesis would predict. Timesharing performance appears to depend rather heavily on the structure of the tasks, not just upon their total demands. Apparently trivial changes in the structure of one task can have large effects on the efficiency of timesharing performance (Wickens, 1979). This and other difficulties led to the proliferation of relatively formal, multiple resource models, in which a number of unspecified individual processing resources are postulated (Sanders, 1979; Navon and Gopher, 1979). In these models, timesharing is possible so long as the aggregate demand for any resource does not exceed the available capacity of that resource. The problem with these models is that the identity of the individual processing resources are not known and the methods proposed for identifying individual resources appear impractical at best (Navon and Gopher, 1979).

The present formulation offers some help in these matters. It suggests a structure of processing resources that is hierarchically organized. The idea of a general processing resource is retained and identified with the regulation of the gating system of the reticular nucleus of the thalamus by the brainstem reticular formation. Since this system modulates the efficacy of thalamocortical interactions, it may be thought of as expanding and contracting the general information processing capacity of the forebrain in its actions. The idea of specific processing resources is also supported by neurophysiological evidence. Specific processing resources in the sense of Navon and Gopher (1979) are identified with regionally restricted functionally specialized cortical information processing systems. These systems may be defined by neuropsychological investigations of the effects of restricted brain lesions on the structure of information processing skills. Thus, most clinicians would be willing to postulate a language processing system, a memory system, specialized visual, auditory and somatosensory perceptual systems, and a response organization and motor control system (Walsh, 1978). The important point here is that the neuropsychological literature provides guidance for the cognitive theorist in suggesting the identities of probable specific processing resources.

A major feature point of this physiologically based
Model is the hierarchical relation between the general or primary processing resource and the specific or secondary processing resources. It offers a meaningful incorporation of the best features of both single and multiple resources models, along the lines suggested by Sanders (1979). Mental workload may now be defined as the demand for the general processing resources and may be measured by pupillometric and other related physiological methods. The role of specific task structure in timeshared task performance may now be better understood by examining the secondary processing limitations imposed by functionally specific cortical secondary processing resources. By proposing an initial secondary structure of processing resources the cognitive scientist need not proceed without landmarks in analysing structural interference between concurrently performed cognitive tasks. Finally, this hierarchical structural model also has relevance for the study of individual differences in cognitive processing, being quite compatible with two level factor analytic theories of intelligence.

From a neurophysiological point of view, the present model suggests that increased emphasis should be placed upon the study of the structure of thalamocortical interactions and the dynamics of their modulation. This type of information appears likely to contribute to a deeper understanding of the neurophysiology of higher mental processes in the human nervous system.
REFERENCES


Ahern, S.K., & Beatty, J. Physiological evidence that demand for processing capacity varies with intelligence. In M. Friedman, J.P. Dos & N. O'Connor (Eds.), Intelligence and learning. New York: Plenum, in press.


Bradshaw, J.L. Background light intensity and the


Miller, G.A. The magical number seven, plus or minus two: Some limits of our capacity for processing information. Psychological Review, 1956, 63, 81-97.


Moruzzi, G., & Magoun, H.W. Brain stem reticular formation and activation of the EEG. Electroencephalography and Clinical Neurophysiology, 1949, 1, 455-473.


Task-Evoked Pupillary Response


Spielberger, C.D. State-trait anxiety inventory. Palo


END
DATE FILMED
9-80
DTIC