Performance on a sustained attention task as a function of strategy: A cross-sectional investigation using the Mackworth clock-test

LEONARD M. GIAMBRA, REGINALD E. QUILTER,
PAMELA B. PHILLIPS, and BARBARA S. HISCOCK
Gerontology Research Center, National Institute on Aging, Baltimore, Maryland

In an attempt to ascertain strategies used during a Mackworth clock-test vigilance task, we had 128 men, 18 to 89 years old, answer a series of open-ended questions immediately following the task. Each of the posttask response protocols was assigned to one of six strategies by two judges. Age was not related to the strategy used, but strategy was found to be significantly related to target detection accuracy. Strategies based upon active mental involvement in the vigilance task produced the greatest level of target detection, whereas a strategy based only upon self-arousal manipulation through limited physical movement produced the poorest target detection. Strategy had no effect on the time course of performance.

Vigilance performance is influenced by the conditions of the specific vigilance task and the physiological state of the individual (Davies & Parasuraman, 1982). Theories of vigilance have attempted to explain vigilance in terms of arousal, expectancy, habituation, observing responses, inhibition, and signal detection theory (Davies & Parasuraman, 1982; Warm, 1977). The conscious, deliberate actions of the vigilant actor have received virtually no consideration. Occasionally, anecdotal accounts of spontaneous postsession verbalizations have been reported. Conscious or effortful processes in vigilance have been dubbed "control processes" by Fisk and Schneider (1981), and have been shown by them to exert a powerful influence on performance. Thus, one may be able to learn a great deal about control processes for vigilance by asking the vigilant actor to describe his/her strategy.

Bakan (1963) apparently conducted the only study in which vigilant actor reports of conscious control processes were related to task performance. After taking 100 young men through a 48-min auditory cognitive vigilance task, Bakan asked them to respond "yes" or "no" to 78 questions dealing with their task performance and mental state. Of these questions, 19 may be interpreted as being related to attempts to regulate their performance during the task. Of these, 8 were found to be significantly related to accuracy of performance: (1) "After hearing a signal and writing it down I would take a short break and stop listening for a while." (2) "After a while I got to feeling hopeless and I sort of gave up." (3) "After a while I didn't care much about getting the signals." (4) "I

didn't try as hard as I could have, since doing well would not do me any good." (5) "I got to points where I stopped listening, and when I stopped, I stopped for a long time." (6) "Sometimes I was tempted to write down any signal at all, just to be doing something." (7) "At times I felt like giving up and just sitting there till it was over." (8) "My concentration was continuous; that is, I never relaxed my listening, not even for short periods of time." For all but question 8, subjects who responded "no" had higher numbers of detected targets. Overall, the correlations between responses to these questions and performance on the task suggest that men who made deliberate efforts to maintain involvement in the task were better performers than those who withdrew attention or who were passively involved.

Although Bakan's (1963) approach provides information on the vigilant actor's conscious control processes or strategies, the yes/no response format to questions dealing with control processes is limited and requires each actor to accept or reject the use of a strategy without qualification. In the present study we allowed participants to respond to a series of open-ended questions regarding their control processes or strategies. Furthermore, additional questions were posed when the responses were vague or unclear or required further elaboration. Strategies and control processes gleaned from verbal reports were then compared with regard to their effect on accuracy and response speed. A lifespan sample was utilized to investigate the possibility of age differences in control processes or strategies in vigilance.

METHOD

Subjects

The cross-sectional sample of 128 men was partitioned into five age groups: 18-39 ($n=28$), 40-49 ($n=9$), 50-59 ($n=18$), 60-69 ($n=26$), and 70-89 ($n=38$); 9 additional men attempted but did not complete

the clock-test. All were participants in the Baltimore Longitudinal Study of Aging (Shock et al., 1984), and were mostly college-educated whites living in the Baltimore-Washington area. The men were tested in the period from 1980 to 1984. This was part of a larger study of age effects on vigilance, reported by Giambra and Quilter (1988).

Apparatus

The clock was a metal box with a plain white face 30.5 cm in diameter and a single black pointer 15.25 cm in length that moved in 100 discrete steps of 3.6° around the face. Each step occurred within a 1-sec interval. During the 62-min test period, at long and irregular intervals, the pointer traveled twice the usual distance in 1 sec, subtending an arc of 7.2°. These "double jumps," 23 in number, were the targets. The clock was set upright on a shelf 1.22 m in front of the subject and was illuminated by a 40-W incandescent lamp placed above and behind his head. For several subjects, who experienced difficulty in seeing the clock pointer, the 1.22-m distance was reduced by as much as 31 cm so that the pointer could be clearly seen.

Procedure

Mackworth clock-test. Before the test session, each subject completed a drug information form listing all drugs, over-the-counter and prescription, taken on a regular and irregular basis. Subjects who 12 h prior to the test session had taken drugs, including alcohol, that would interfere with performance were excluded. The clock-test was preceded by instruction, a demonstration, and practice. During the entire session, the subject sat comfortably in a hospital bed that was adjusted to support his legs in a horizontal position. During the clock-test, this area was surrounded by a white curtain that isolated the subject and the clock from the rest of the room. As nearly as possible, all extraneous noises were excluded from the room. The only sound emitted by the clock was the steady, faint drone of its motor. This sound mingled with the noise of a chart-drive motor.

During the instruction and demonstration period, the subject was told that he would be working at a "watch-keeping" task and that it was a demanding one requiring his full attention and cooperation. After the subject was comfortably settled in front of the clock, the experimenter started the mechanism and, together with the subject, observed the ordinary movement of the pointer. The experimenter told the subject that every now and then, at long and irregular intervals, the pointer moved twice the usual distance and that this was termed a "double jump." After the subject learned to recognize the difference between the two kinds of movements, 5 to 10 double jumps were presented at different positions around the face of the clock, interspersed with ordinary movements. The experimenter watched the clock with the subject and prompted him whenever necessary to make sure that he understood the task.

The practice period was a short sample of the actual test, and was designed to give the subject some idea of what to expect from the clocktest. The subject was given a response key—a microswitch, contained in a comfortable hand-held grip, which was actuated by a thumbpress—and was told to hold it in his right hand. The subject was informed that double jumps would appear, and that his task was to press the response key as quickly as possible when he recognized them. During this time the subject was alone. Four double jumps were presented, and at each successful keypress the response time (RT) was voice relayed to the subject, who was encouraged to respond faster. If the subject failed to press the response key after the presentation of a double jump, the experimenter called out that one had just occurred. False presses were also reported to the subject.

A rest period of 5 min or more followed the practice period. The subject was reminded that the clock-test entailed watching the clock pointer for 1 h, during which an unspecified number of double jumps would occur in a random order; he was told to respond to the double jump "as fast as possible," but even when late always to press the response button, since we were also interested in whether or not he detected the double jumps. False presses, it was pointed out, would not affect his score on the test. The subject was not permitted the use of a timepiece, and was told that the experimenter would neither communicate with him nor answer any query unless some emergency arose.

The 23 double jumps presented did not, in reality, occur at random. They appeared at precisely the following times, in minutes, into the clock-

test: 2, 4, 7, 11, 13, 15, 19, 21, 23, 26, 30, 32, 34, 37, 41, 43, 45, 49, 51, 53, 56, 60, 62. The subjects' postsession comments indicated that they were unaware of this, and many expressed the belief that the double jumps had indeed appeared in random order.

Strategy assignment. Immediately following the test, the experimenter explained to the subject that she would like to ask him some questions about the test and that his answers and comments would be taped. If the subject had no objections, the experimenter asked 13 questions, with the following pertaining to strategy: (1) "What was going on in your mind throughout the session?" (2) "How did you keep yourself alert?" (3) "How did you make sure you were reacting as quickly as possible to a double jump?" (4) "How did you separate a single jump from a double jump?" If any of the subject's answers were not clear to the experimenter, she asked the subject to elaborate or explain further.

The interviews were transcribed and the strategy employed was rated as belonging to one of six types. The six rating values, from highest to lowest, and their attendant strategies were as follows: (6) watched the clock and counted the single jumps in an attempt to determine the existence of a pattern in the appearance of the target; (5) counted single jumps, but in a nonsystematic manner with no attempt at determining a target pattern; (4) watched the clock with no counting; (3) reported only an attempt at concentration and at a nonspecific anticipation of the target; (2) attempted physical movements in order to maintain alertness; (1) reported being hypnotized or mesmerized by the clock movement.

An initial group of 62 protocols was rated independently by two judges, with an interjudge correlation of .80. Disagreements in absolute values of ratings were reviewed by the two judges and resolved by discussion. The remaining 66 protocols were then judged independently by the two judges, with an interjudge correlation of .89. Disagreements in absolute value ratings were again resolved by discussion. The protocols of 3 subjects could not be rated. The protocol of 1 subject was rated 1, and he was omitted from the subsequent analyses.

RESULTS

Three performance measures were directly obtained from the completed 62-min Mackworth clock-test. The first measure, hits, was simply the number of targets detected out of 23. The second measure, false presses, was the total number of times that the subject mistook a single jump for a double jump or target and pressed the response key. This measure was not analyzed, however, since most of the subjects produced very few or no false presses, and a few produced a great number of false presses; thus mean values were unreliable. This also precluded the use of signal detection measures. When A' was calculated it was found to correlate .95 with hits; thus it was essentially redundant with hits. The third measure, mean RT, was the mean RT for all the targets detected. The RT was the interval from the appearance of a double jump to the subject's press of the response button, with intervals longer than 6 sec excluded.

The subsequently reported analysis of variance (ANOVA) calculations were based upon the data described above and upon transformations of these measures suggested by Winer (1971) to accommodate potential violations of ANOVA assumptions. The transformations used were square root of hits and natural logarithm of 1.0 added to mean RT (sec). Unless otherwise stated, the ANOVA results reported are for these transformed measures and any untransformed data outcomes that did not correspond to the transformed data outcomes with regard to statistical significance at least at the .05 level.

Age Effects

A one-way ANOVA was performed in which age group was the independent variable and strategy rating was the dependent variable; the effect of age group was not significant [$F(4,114) = 1.24$, $MSe = 1.96$, $p > .05$]. The previous ANOVA presumes that the strategy ratings were on an equal interval scale, and the lack of an age effect may have been due to the ratings' being on a nominal scale. Therefore, because age was on a ratio scale, an alternative ANOVA was carried out in which strategy rating was treated as the independent variable and age was the dependent variable; the effect of strategy rating was not significant [$F(4,110) = 0.72$, $MSe = 392.31$, $p > .05$]. It was concluded that age and strategy rating were independent; in all subsequent analyses no age distinctions were made.

Strategy Effects

Because there was only a single subject with a strategy rating of 1, that category was omitted in the subsequent ANOVAs. Strategy rating significantly affected hits [$F(4,110) = 4.79$, $p < .01$, $MSe = 0.3317$], but not mean RT [$F(4,110) = 1.09$, $p > .30$, $MSe = 0.0092$]. The smallest number of hits was obtained by strategy rating 2 ($M = 10.9$), whereas the greatest number of hits was obtained by rating 6 ($M = 17.4$); ratings 3, 4, and 5 had similar numbers of hits ($Ms = 15.4$, 16.2, and 15.5, respectively). Clearly, strategies of rating 2 were the least effective in terms of detecting targets; less clearly, strategies of rating 6 were the most successful. A Scheffé post hoc test found that the hits for rating 2 were significantly different ($p < .05$) from all other rating types, which, in turn, were not significantly different from each other.

Strategy Rating × Time Into Task Interaction

To determine the influence of strategy upon the time course of vigilance, a strategy rating × time into task ANOVA was carried out, using 23 time periods. The time periods were those associated with the appearance of each target. The strategy rating × time into task interaction was not significant ($p > .05$) for hits [$F(88,2464) = 0.89$, $MSe = 0.1788$] or mean RT [$F(88,1625) = 0.71$, $MSe = 0.177$].

DISCUSSION

Strategy was associated with detection accuracy. Generally, a strategy of active, conscious, and directed involvement in the vigilance task was superior to haphazard or passive involvement or concentrating on self-arousal through physical movements. A strategy of manipulation of physical movements to keep alertness levels up was associated with a substantially lower detection accuracy. Davies and Parasuraman (1982) reported several experiments in which physical exercise during vigilance improved performance. In the present experiment physical movement was confined to actions carried out while seated. This difference may account for the ineffectiveness of the physical movement strategy in the present experiment. More important, direct involvement in the vigilance task appeared to be the better strategy. Among the direct involvement strategies, a strategy including very active and persistent efforts to understand the vigilance task or the parameters of the target occurrences was associated with the highest levels of detection accuracy—although not significantly so, using the Scheffé post hoc test. This is consistent with Bakan's (1963) work. These results suggest that for tasks requiring a vigilant actor, such as industrial inspection tasks, attempts to maintain alertness based only on limited physical movements do little to improve detection likelihood. Furthermore, an active, aggressive involvement in the task is most likely to yield the greatest target detection accuracy.

Finally, young, middle-aged, and elderly men appeared to be equally disposed to employ each of the six strategies. The failure to obtain a relationship between age and strategy usage may reflect a true age equivalence or may reflect the crudeness with which the strategies were separated. Any decision on this must await the use of more precise methods and replication across a variety of vigilance tasks.

REFERENCES

Bakan, P. (1963). An analysis of retrospective reports following an auditory vigilance task. In D. N. Buckner & J. J. McGrath (Eds.), *Vigilance: A symposium* (pp. 88-101). New York: McGraw-Hill.

Davies, D. R., & Parasuraman, R. (1982). *The psychology of vigilance*. London: Academic Press.

Fisk, A. D., & Schneider, W. (1981). Control and automatic processing during tasks requiring sustained attention: A new approach to vigilance. *Human Factors*, 23, 737-750.

Giambra, L. M., & Quilter, R. E. (1988). Sustained attention in adulthood: A unique, large sample, longitudinal and multicohort analysis using the Mackworth Clock-Test. *Psychology & Aging*, 3, 75-83.

Shock, N. W., Greulich, R. C., Andres, R., Arenberg, D., Costa, P. T., Jr., Lakatta, E. G., & Tobin, J. D. (1984). *Normal human aging: The Baltimore longitudinal study of aging* (NIH Publication No. 84-2450). Bethesda, MD: The National Institutes of Health.

Warm, J. S. (1977). Psychological processes in sustained attention. In R. R. Mackie (Ed.), *Vigilance: Theory, operational performance, and physiological correlates* (pp. 623-644). New York: Plenum Press.

Winer, B. J. (1971). *Statistical principles in experimental design* (2nd ed.). New York: McGraw-Hill.