Visual vs. Auditory Memory in an Aviation Task: a Potential Performance Theory Analysis

Gayle Hunt
Stephen Rice
David Trafimow
Jeremy Schwark
Joshua Sandry

See next page for additional authors

Follow this and additional works at: https://corescholar.libraries.wright.edu/isap_2011

Part of the Other Psychiatry and Psychology Commons

Repository Citation

This Article is brought to you for free and open access by the International Symposium on Aviation Psychology at CORE Scholar. It has been accepted for inclusion in International Symposium on Aviation Psychology - 2011 by an authorized administrator of CORE Scholar. For more information, please contact corescholar@www.libraries.wright.edu, library-corescholar@wright.edu.
VISUAL VS. AUDITORY MEMORY IN AN AVIATION TASK:
A POTENTIAL PERFORMANCE THEORY ANALYSIS

Gayle Hunt¹
Stephen Rice¹
David Trafimow¹
Jeremy Schwark¹
Joshua Sandry¹
Lisa Busche¹
Kasha Geels¹
¹New Mexico State University
Las Cruces, New Mexico

Information can be relayed to pilots by visual presentation or auditory presentation; both methods are frequently used. To date, there is quite a bit of conflicting literature regarding which type of communication is most effective for recalling information. The current study tested memory differences between digits presented visually or audibly in number strings. Results showed that performance in the visual presentation condition was superior to the auditory presentation condition. However, when a PPT analysis was conducted on the data, it was revealed that performance in the visual condition was only superior because participants were more consistent in that condition compared to the auditory condition. In other words, had participants been perfectly consistent in both modalities, there would have been no differences in performance. We conclude that inconsistency, and not different strategies, was responsible for the increased performance of visual memory over auditory memory.

The current study focuses on visual and auditory digit memory, an area that is well-studied in the basic cognitive memory literature but one that does not always scale up to the real-world in ways that are useful for applied research. In the basic literature, when comparing visual to auditory presentations for memory recall, it is critical to avoid any confounds that might give one presentation an inherent advantage over another. Thus, if auditory digits are presented serially, visual digits also must be presented serially, and the rate of digit presentation must be controlled (e.g. Jensen, 1971).

However, in the real world, it often is not the case that visual digits can be, or should be, presented serially in such a fashion. For example, a pilot reading digits should have access to the entire string of digits at once glance. An air traffic controller reading aircraft call signs has, or should have, immediate access to the entire string of digits. Cases like these can be imagined in many types of applied aviation settings.

Of course, comparing a visual presentation, where all digits are presented simultaneously, to an auditory presentation, where all digits are presented serially, results in an inherent advantage to the visual presentation if the two presentations are compared for overall memory performance. In the visual presentation, participants can review previous digits whereas in the auditory presentation, backtracking is impossible. In this situation, it seems likely that simultaneous visual presentation promotes superior information processing strategies compared to serial auditory presentation, which in turn provides a reason for observed performance to be better in the former case than in the latter one. Alternatively, it is possible that the two types of presentations allow for equally good information processing strategies but that the information processing strategies promoted by simultaneous visual presentation are used more consistently than those promoted by serial auditory presentation. Either way, observed performance in the simultaneous visual processing condition should exceed that in the serial auditory condition.

Because the goodness of strategy hypothesis and the consistency hypothesis make the same prediction with respect to observed performance, it is necessary to find some means to distinguish them from each other. Consequently, we focus on the issue of parsing strategy and consistency in digit-recall memory tasks. But before we examine this issue, it is necessary first to briefly review the literature pertaining to visual versus auditory memory.

Visual Versus Auditory Memory
Much work has been done comparing memory of information presented in a visual versus auditory manner. There are conditions where visual presentations cause better performance than auditory presentations (e.g., Hawkins, 1897; Henmon, 1912; Worchester, 1925) but there also are conditions where the reverse is so (e.g., Binet, 1894; Dornbush, 1968; Koch, 1930).

Day and Beach (1950) reviewed the entire body of research prior to 1950 and concluded that visual memory is generally superior to auditory memory; however, studies also show that visual memory degrades more quickly than auditory memory with age (McGhie, Chapman & Lawson, 1965) and that children produce superior performance with auditory presentations compared to visual presentations (Abbot, 1909). These effects may be differentially affected by a number of factors (Beaman, 2002; Jensen, 1971; Norman, 1966; Saults & Cowan, 2007; Sherman & Turvey, 1969).

Several researchers have proposed variations on an argument relating presentation format to learning histories (e.g., Dornbush, 1968; Johnson & Miles, 2009; Saults & Cowan, 2007). To understand this argument, consider that in normal living, people tend to be exposed to auditory stimuli serially and to visual stimuli all at once. Although there doubtless are exceptions, it seems reasonable to suppose that most people have learned to process auditory stimuli in accordance with serial presentations and visual stimuli in accordance with simultaneous presentations. Consequently, serial presentation of stimuli confers an advantage for auditory versus visual presentation, which provides one explanation for the cases where participants perform better in auditory than visual presentation conditions.

Taking this argument seriously implies the possibility of a subtle confound in experiments designed to test visual versus auditory presentation conditions. That is, if participants encounter stimuli serially, with a resulting advantage in the auditory presentation condition relative to the visual presentation condition, it could be due to a better matching of presentation condition with previous learning history rather than due to the superiority of auditory presentations to visual ones. On the other hand, if the participants in the visual presentation condition encounter the stimuli simultaneously, whereas the participants in the auditory presentation condition encounter the stimuli serially, an obtained effect might be due to the simultaneous-serial difference rather than due to the visual-auditory difference. Either way there is a potential confound.

Given that there exists a confound no matter how the experiment is designed, theoretical and methodological considerations cannot, by themselves, determine the confound with which we should choose to live. Rather, we believe that applied considerations also should figure into the decision. And it is obvious that in most applied settings, visual information is presented simultaneously and auditory information is presented serially, in accordance with our past learning histories. Therefore, we choose to resolve the learning history confound and live with the simultaneous-serial confound—a decision that we emphasize is strongly influenced by the fact that our goal is applied rather than basic.

As we pointed out earlier, because the effects of visual versus auditory presentations on digit-recall are so dependent on a multitude of factors, it is unlikely that any one perspective is going to explain all of the data, at least not in the near term. Consequently, it is not our intention to choose a “winner.” However, choosing a winner is not the only possible contribution. If it could be determined (a) whether simultaneous visual presentation causes better performance than serial auditory presentation and (b) whether the effect is due to random or nonrandom factors, there would be important implications for applications. Until recently, there has been no way to parse random and nonrandom factors; however, a recent theory—termed potential performance theory (PPT; Trafimow & Rice, 2008, 2009)—does just that.

Potential Performance Theory

Potential Performance Theory is a mathematical theory of task performance that allows researchers to look beyond observed performance and calculate potential performance in the absence of any random factors. According to this theory, one of the primary factors that must be taken into account is consistency. ‘Consistency’ is a term that can be characterized in many different ways, depending on the area to which it is being applied (Brunswick, 1952; Kelley & Michela, 1980; Orvis, Cunningham, & Kelley, 1975). PPT uses the term consistency to reference the correlation coefficient calculated between two identical blocks of trials in a study. For example, suppose a person is given a pair of two ten-digit numbers and asked to identify whether the pair of numbers is the same or different. The
person repeats this task 100 times, each time being presented with a unique set of numbers. Following this task is a short break and a second block of 100 trials. The second block is identical to the first, so the person is giving an answer to each pair of numbers two times. This allows the researcher to calculate a correlation coefficient across the two blocks. This correlation coefficient is also known as the consistency coefficient, which is essentially an inverse measure of randomness. More randomness necessarily implies a lower consistency coefficient while less randomness implies a higher consistency coefficient.

Potential performance, as defined by PPT, is the score a person could achieve in the absence of randomness. When a person’s accuracy is greater than chance, more randomness will cause observed performance to decrease. In the absence of randomness, a person’s answers to each block of trials should be identical, producing a consistency coefficient of exactly 1.00 and an observed score equal to the potential score. As randomness increases, this coefficient will decrease and cause the observed score to fall below the potential score. Once observed scores are measured and consistency coefficients are calculated, a researcher can use the formulas given by Trafimow and Rice (2008, 2009) to calculate potential scores for each participant. (For lack of space, those formulas are not presented here.)

By virtue of having both observed scores and the results of PPT computations, it is possible to have the three main variables of importance for each participant. These are the observed score or observed proportion of successes, the potential score, and the consistency. In turn, as the experiment to be presented demonstrates, these variables make it possible to disentangle whether differences between simultaneous visual presentations and serial auditory presentations are due to nonrandom factors, random factors, or both.

Current Study

In the current study, participants viewed a 10-digit string of random numbers visually or they were exposed to the same 10-digit string of numbers auditorily, via headphones. Visual presentation where all digits are presented concurrently provides an inherent advantage over an auditory presentation where all digits are presented serially. Thus, our first hypothesis is that visual presentation of digits will result in superior observed performance compared to auditory presentation of digits. However, what is not definitively known is why this particular paradigm advantage results in superior performance. One might assume that this advantage results in a superior type of memory processing strategy (nonrandom factor) which would in turn indicate better potential performance. However, it is possible that this advantage also triggers more consistency (or less randomness). Thus, our second hypothesis is that visual presentation of digits will result in more consistent performance compared to auditory presentation of digits. It is entirely plausible that the visual presentation will result in both superior potential performance and superior consistency, and thus we have three additional competing hypotheses: A: Visual presentation of digits will result in superior potential performance and higher consistency compared to auditory presentation of digits; B: Visual presentation of digits will only result in higher consistency compared to auditory presentation of digits; and C: Visual presentation of digits will only result in superior potential performance compared to auditory presentation of digits.

Method

Participants

Twenty (13 females) undergraduate students from a large southwestern university participated in the experiment for partial course credit. The mean age was 20.3 (SD = 2.70). All participants were tested for normal or corrected-to-normal vision and color-blindness.

Materials and Stimuli

The experimental display was presented via E-prime 1.0. Participants were exposed to 100 visually-presented trials, and 100 auditorially-presented trials. Each set of 100 trials was divided into 2 blocks of 50 trials each. Of these, 25 trials were matched and 25 trials were unmatched. A matched trial consisted of two consecutive display presentations whereby a string of 10 randomly generated numbers were identical across both displays. An unmatched trial consisted of the same format with numbers that were not identical; one number in the string was randomly changed.
Each trial began with a fixation display that was presented for 500 msecs. Following this, a Number display presented the first 10-digit number, followed by a Mask display that remained for 200 msecs. Next, a second Number display presented a 10-digit number that either was identical to, or different from, the first Number display. Whether or not the trial contained a matched or an unmatched display was randomly determined. Following this, a Choice display was presented whereby participants were asked to decide whether or not the two numbers were identical. Participants pressed the J key if they determined that the numbers were identical and the F key if they were not identical. Since PPT methodology requires 2 identical blocks of trials, participants were exposed to 2 blocks of visual stimuli and 2 blocks of auditory stimuli. These blocks were presented randomly and all trials within each block were presented randomly as well. Each block consisted of 50 trials. At the end of each block, participants were able to take a short break. Instructions were given at the beginning of each block to inform participants whether the trials would be presented visually or audibly.

**Procedure**

Participants first signed a consent form and were then seated in a comfortable chair facing the experimental display. Viewing distance was controlled by a chin rest at 21 inches. Participants first read on-screen instructions and then were given the opportunity to ask questions. Once they were comfortable with the instructions, participants pressed a button to begin the experiment. The entire experiment took approximately 70 minutes. Upon completion, participants were debriefed and dismissed. The experiment employed a within-participants design, whereby all participants were exposed to both the visual and auditory conditions.

**Results**

The data were analyzed using PPT theorems (see Trafimow & Rice, 2008, 2009) and a series of three t-tests were conducted. As can be seen from Figure 1, observed performances in the visual condition exceeded those in the auditory condition ($M = .84$ and $M = .71$), $t(19) = 3.59, p < .001, d = 1.65$. In addition, consistencies in the visual condition exceeded those in the auditory condition ($M = .56$ and $M = .26$), $t(19) = 4.63, p < .0001, d = 2.12$. Finally, potential scores did not differ discernibly between the visual and auditory conditions ($M = 1.00$ and $M = .98$), $t(19) = 0.34, p = .37, d = 0.16$.

![Figure 1. Experimental data.](image)

**Discussion**

Our first hypothesis was that performance in the visual presentation would exceed that of the auditory presentation. Not surprisingly, this was the case. Clearly, there was an advantage for participants to view all visual
digits concurrently. Of course, it could also be the case that the visual presentation would have produced superior performance even if the digits had been presented serially (Jensen, 1971); however, we believe it is not wise to speculate here. Our second hypothesis was that visual presentation of the digits would result in more consistent behaviors. The data clearly show this to be the case. Consistency in the visual presentation exceeded that in the auditory presentation by both a statistically and practically significant amount with an impressive effect size.

We pointed out earlier that PPT implies three classes of explanations for observed effects; these are potential performance, consistency, or both. Because the two types of presentations resulted in differences in observed performances and consistencies, but not in potential performances, it is clear that two of the three possibilities can be ruled out. Put more simply, it is consistency, and not potential performance, that is responsible for the observed effect in the current study. Thus, our third hypothesis (B) appears to be correct, and we can reject its competing hypotheses (A and C).

What does this mean in plain English? Consider again that observed performance is caused by a combination of nonrandom effects and random effects. Suppose that there were no random effects. In that case, the PPT analyses indicate that observed performances in the two stimulus presentation conditions would be equal. In other words, the systematic components of task performance are equally effective in the visual and auditory presentation conditions. The reason for the observed effect, then, is because people are more consistent in the visual presentation condition than in the auditory presentation condition. The stimulus presentation manipulation causes people to behave more randomly in the auditory presentation condition than in the visual presentation condition, thereby reducing observed performance more in the former than in the latter condition. To our knowledge, none of the available literature can account for this.

But the data are even more surprising than that. Consider that mean strategy scores equaled or were very close to unity. This means that practically all of the error in observed performances was due to randomness; there was almost no error due to systematic causes in either condition.

The lack of systematic error also suggests some practical applications. Suppose that one wishes to train people to have better observed performances in the current paradigm, whether with visual or auditory presentations. Because the problem is one of consistency, it would be useful to find out exactly what it is people are doing and train them to do that all of the time. Possibly, this could be done simply by making people’s strategies explicit, thereby placing them under conscious control.

In conclusion, the data reveal that different levels of potential performance are not responsible for the superior visual memory performance; rather, participants were simply more consistent in the visual condition than in the auditory condition. We believe that designers of equipment to be used by pilots and air traffic controllers should be aware that different modality presentations can differentially affect how consistently operators interact with visual and auditory displays.

Acknowledgements

The authors wish to thank Eric Johnson, Audrey Rosenblatt, David Flemming, Sandra Deming, Eduardo Rubio, Crystal Sandy, Rachael Currier, Crystal Garcia, and Natasha DeVries for their help in collecting data.

References


