Auditory Inspection Time and Intelligence: A New Spatial Localization Task

Denis M. Parker

Glasgow Caledonian University, Glasgow, UK

J.R. Crawford
Edward Stephen

University of Aberdeen, Aberdeen, UK

Sixty-seven university undergraduates were tested on a new auditory inspection time (AIT) task based on the ability of human subjects to locate the left or right position of a stereophonic source and on a modified version of the Raven's Progressive Matrices (MRPM). The AIT task required subjects to locate the spatial position of a dichotic phantom lateralized by a phase difference to their left or right side that was backward masked by a high intensity source located in the midline of the head. Unlike previously described AIT tasks all subjects were able to perform the task. The log of AIT threshold was significantly correlated with the MRPM ($r = -0.37$). In a further sample ($n = 41$) an r of -0.61 was observed; this latter sample was retested, with a delay of between 3 and 10 days and an r of -0.57 was obtained on this second occasion. In comparison, the correlation between visual inspection time (VIT) and MRPM in this sample was -0.41 and -0.21 on retest. Test–retest reliability for AIT was high (0.82) and comparable to that obtained in the same sample for VIT (0.76). It is concluded that this new AIT task appears to be accessible to a wider range of subjects than the successive frequency discrimination tasks employed by others.

INTRODUCTION

Inspection time (IT) is a measure of the presentation time required for an observer to reliably perform a sensory discrimination task (Vickers, Nettelbeck, & Willson, 1972). Although it has its roots in reports stemming from investigations carried out more than a century ago (see Deary, 1986), the finding of a significant correlation between the time required to discriminate two backward masked visual patterns and a standardized measure of Performance IQ (Nettelbeck & Lally, 1976; Lally & Nettelbeck, 1977), reawakened the

---

interest in the relationship of basic perceptual/cognitive processes to measures of intelligence. Subsequent studies have modified estimates of the magnitude of the association; Nettelbeck and Lally had included, in a small subject sample, individuals with a wide range of intellectual competence, including some with intellectual disabilities, and so the very high -0.8 to -0.9 correlations they reported were inflated estimates of the strength of the association. Nevertheless, the repeated finding of a significant relation between IT measures and psychometric intelligence (e.g. Crawford, Deary, Allan, & Gustafsson, 1998) has confirmed the reliability of the association, even though the best estimate of its strength would be a correlation in the region of -0.5 (Nettelbeck, 1987; Kranzler & Jensen, 1989).

The fact that such a basic task can predict approximately 25% of the variance of psychometric IQ is a finding of considerable theoretical interest, and it is natural to ask whether this relationship is confined just to visual processing or whether it is a general property of sensory/perceptual systems. Some indication that the relation may be a general one comes from studies of auditory inspection time (AIT). Using a judgment of the sequence of forward and backward masked tone pairs of different frequencies, Brand and Deary (1982) reported a weak, and non significant association, between AIT and both Raven's Progressive Matrices (RPM) and Mill Hill IQ when mentally handicapped subjects were removed from the sample. A subsequent study by Irwin (1984) found significant correlations (Kendall's tau) between a very similar AIT task to that of Brand and Deary (1982) and both RPM and the Mill Hill Vocabulary scale. Nettelbeck, Edwards and Vreugdenhil (1986) used a development of the Brand and Deary (1982) task and found significant correlations between a range of cognitive tests and AIT in a group of 29 adult subjects. Deary, Head and Egan (1989), using the masking variant of the AIT task developed by Nettelbeck et al. (1986) found significant relationships between AIT and verbal and non verbal IQ in both adults and children. In this group of four AIT studies, spanning seventeen years, the correlations have ranged from -0.23 to -0.45; there would appear to be a prima facie case for concluding that speed of processing auditory information, like speed of visual information processing, is related to measures of intelligence. However, matters are not as straightforward as they might at first appear.

The AIT tasks discussed above all rely on the subject's ability to perform a frequency discrimination, since the subject must specify the order in which the two differently pitched tones were heard. It has been apparent since the first paper by Brand and Deary (1982) that many subjects find it difficult to perform the basic discrimination at any tested exposure duration; in their study over 20% of subjects could not meet the training criterion. Irwin (1984) reported that the ability to perform this type of pitch order task was correlated with the Seashore pitch test (Kendall's tau, -0.51). Nettelbeck et al. (1986) reported that some subjects found the sequential pitch discrimination inordinately difficult, while Deary et al. (1989) found that only 35 of 59 students (59%) and 53 of 125 children (42%) were able to perform the AIT task satisfactorily. It is clearly unsatisfactory if such large proportions of samples cannot do the basic task; one is bound to have reservations about the generality of the relationships if they are based on a subset of any sample tested. The question that occurred to the authors was whether an auditory test could be developed that would not depend on pitch discrimination and so be accessible to the majority of any sample tested.

In the most commonly used version of the visual inspection time (VIT) task the subject is asked to judge whether the longer or shorter line is on the left or right of a masked display; the subject makes a spatial decision. Furthermore, the task is usually easy to perform when the display is presented for relatively long durations, and the vast majority of subjects are able to carry out the task; it is only when the display interval becomes relatively short that subjects' performance becomes impaired. With the AIT tasks discussed above the subject makes a judgment about the temporal order of the relative pitches of two sounds with different fundamental frequencies. Furthermore, a proportion of subjects find the basic discrimination of the tones which comprise these tasks extremely difficult and either have to be excluded from the analysis or given an arbitrary score. The formal discrepancies between the tasks, spatial vs. temporal and differences in difficulty of the basic tasks, makes it debatable whether one is dealing with the same type of processing constraint in visual and auditory versions.

In order to develop an auditory task that appears to be of equal ease and formally equivalent to the commonly used VIT task, a new auditory discrimination task, and its relationship to estimated IQ, was investigated. The task is based on the ability of observers to locate the position of a sound source to their left or right. Under natural conditions, sound sources positioned outside the midline axis of the head present stimuli differing in amplitude and temporal characteristics (time of arrival and phase) to the observer's left and right ears. Sound localization is an immediate easy task to perform that is present even in infancy (Wertheimer, 1961). The basic characteristics of the sound that allow it to be spatially located have been known since the last century (Thompson, 1877) and over the years studies have refined our knowledge of the process (Howard & Templeton, 1966). If, by using artificial sound sources and headphones, a pure tone is delivered, so that only a phase difference remains between the waveforms at the two ears (amplitude and onset time differences having been removed), the observer will hear the tone lateralized towards their left or right. This lateralized sound source is heard within the head towards the left or right ear, and because it lacks the characteristics of real external sound, it is referred to as a "dichotic phantom." In the AIT task described below, the duration of a backward masked tone, differing solely in phase between the two ears, was varied and the threshold duration for reliable location discrimination determined.

In the first study of the present paper we describe the methods and procedures used for this new AIT task and obtain provisional data to test the hypothesis that it will be significantly related to psychometric intelligence (Raven's Matrices). As Deary (1993) and others have argued that inspection time is more strongly related to fluid intelligence (g_f) than crystallized intelligence (g_c) we also test this hypothesis for AIT by including a Vocabulary test as a measure of g_c.

The second study attempts to replicate the results in a further sample. In addition, because concern about the reliability of existing inspection time tasks has been noted (e.g., Kranzler & Jensen, 1989), the test–retest reliability of the AIT task is examined and is compared with the test–retest reliability of a conventional VIT task.

Finally, this second study offers the opportunity to compare directly the correlations of AIT and VIT with psychometric intelligence in the same sample. Using the new AIT task this comparison can be made with reasonable confidence because the vast majority of the healthy population would be expected to be capable of making the basic discriminations necessary for *both* tasks.

<center>Study 1</center>

<center>**Subjects**</center>

Sixty-seven undergraduate subjects, 44 female and 23 male were recruited for the study. They were aged between 17 and 26 (M = 20.9, SD = 1.5) and were given no screening tests other than being asked if they were aware of having any hearing problems, particularly any relative deafness in one ear. Only two subjects were excluded on the basis of the answers they gave.

<center>**Methods**</center>

Auditory Inspection Time

The hardware used to produce the waveforms for the AIT test consisted of an IBM compatible PC fitted with an 8253/8255 Timer/Input–Output Board. The PC was connected to a Cambridge Electronic Design CED1401, a microprocessor controlled multi-channel digital-to-analogue converter with RAM. Three of the analogue outputs were fed to a relay switching unit operated by the digital output from the 8253/8255 board; the relays were of the reed type with an operating time of 0.5 ms. Analogue channels 0 and 1 carried the sine wave stimulus waveform and the third channel carried the square wave masking waveform. The outputs from the relay unit were fed to a headphone amplifier which had a low pass filter to remove high frequency audio noise. A program written in turbo Pascal computed two sets of values for the stimulus waveforms, one set for a single cycle of a sine wave over 256 steps, the other, again for single cycle of a sine wave, but with the phase angle shifted by 40°. The CED RAM was then loaded with these values along with that of a cycle of a square wave. The CED was programmed to deliver a specified number of cycles, during which the relays were energized to deliver the appropriate output to the headphone amplifier. The duration of the target tone could vary from a maximum of 300 ms to a minimum of 1 ms (0.45 Hz). In this way two 450 Hz sine waves, one with a 40° phase shift relative to the other, were produced. The two waveforms were switched by the relay unit from left to right channel, as determined by the output of the PEST algorithm (Taylor & Creelman, 1967) adapted from a BBC BASIC program developed elsewhere. The sine waves were followed by a mask consisting of a 450 Hz square wave but with no phase difference between the ears. The sound levels were adjusted with the help of a Bruel and Kjaer Sound Level Meter and an artificial ear so that the target tones delivered 59 dB at each earphone while the square wave mask delivered 73 dB.

In summary then, the basic task required subjects to indicate the apparent location, left or right, of a low frequency target tone which was delivered over headphones to produce a dichotic phantom located towards the left or the right ear. The sine wave pairs could be switched between left and right headphone earpieces, thus producing an apparent shift in the location of the tone. The tones were backward masked, using a zero inter-stimulus interval, by the square wave delivered in-phase to both ears for 500 ms, which produced a dichotic phantom located midway between the two ears. The 85% accuracy level was estimated automatically using the PEST algorithm (Taylor & Creelman, 1967). Before the main AIT task commenced subjects were given 20 practice trials at the longest 300 ms duration, with immediate feedback concerning the accuracy of their judgment, in order to

familiarize them with the features of the task. Subjects responded by pressing the left or right button of a computer mouse, as appropriate.

Modified Raven's Matrices Test

A modified version of the Raven's Standard and Advanced Matrices were developed by Page and Crawford (unpublished) because they observed that the Standard Matrices had high negative skew and ceiling effects in a sample of 295 subjects drawn to match the UK adult census figures for age, sex and social class. The aim was to produce a set of 60 items that could be used with groups that included individuals whose ability ranged from the intellectually normal to the intellectually superior. The selection was carried out on sets of 12, rather than individual items, because each item is related to the one that precedes it in the set, and, while the Advanced Matrices are not explicitly divided into sets of 12 the 36 items were so split and labeled A, B, C for the purposes of Page and Crawford's study. The sets which comprised the Modified Raven's Progressive Matrices (MRPM) were sets B, C and E from the Standard series and sets B and C from the Advanced series. The test was administered as an untimed test to each participating subject following the procedure recommended for the Standard RPM (Raven, 1958).

WAIS-R Vocabulary Subtest

The Vocabulary subtest from the Wechsler Adult Intelligence Scale — Revised (Wechsler, 1981) was completed by each participating subject.

Results

All 67 subjects completed the AIT task ($M = 55.2$ ms, SD $= 63.9$, Median $= 34$ ms, Min. $= 5$ ms, Max. $= 298$ ms), the modified Raven's ($M = 45.4$, SD $= 6.9$) task and the WAIS-R Vocabulary subtest ($M = 54.89$, SD $= 8.07$). As anticipated, the distribution of AIT was positively skewed. A log transformation was therefore applied to AIT scores prior to computing correlations involving AIT. A Kolmogorov–Smirnov test revealed that the transformed scores did not differ significantly from a normal distribution.

The Pearson product moment correlation between AIT and the Raven's score was -0.37; $p < 0.01$, 95% confidence limits $= -0.145, -0.562$. A regression analysis, in which power functions of AIT were used to predict Raven's performance, provided no evidence of any non-linear component in this relationship.

AIT and Vocabulary show a weak and non-significant relationship ($r = -0.16$, ns), as do Vocabulary and the Raven's ($r = 0.21$, ns). Using Williams' (1959) test for non-independent correlations, the correlation between Raven's and AIT was not significantly higher than that between Vocabulary and AIT($t = 1.44$, $df = 64$, $p = 0.08$).

Study 2

Subjects

Forty-one undergraduate subjects (24 female and 17 male) were recruited for the study. They were aged between 17 and 43 ($M = 22.2$, SD $= 5.18$).

Methods

The AIT task and MRPM were administered as in study 1. In addition, visual inspection time was measured using the method described by Petrie and Deary (1989). All participants were retested on AIT and VIT with a delay ranging between 3 and 10 days. As expected VIT and AIT scores were positively skewed. A log transformation was applied to the AIT and VIT scores from both test periods prior to conducting further analyses; Kolmogorov–Smirnov tests revealed that none of the four distributions of transformed scores differed significantly from a normal distribution.

Results

The test–retest reliability coefficient for AIT was 0.82, $p < 0.001$. Test–retest reliability for VIT was 0.76, $p < 0.001$. Comparison of these two coefficients using a test of $\rho_{12} = \rho_{34}$ developed by (Steiger, 1980, p. 249), revealed that they did not differ significantly from each other ($z_2 = 0.757$, $p = 0.45$).

The correlation between AIT (time 1) and Raven's was significant; $r = -0.61$ $p < 0.001$, 95% confidence limits $= -0.38$, -0.77. A similar magnitude of correlation was observed between Raven's and AIT at time 2 ($r = -0.57$, $p < 0.001$, 95% CLs $= -0.32$, -0.75). The correlations between VIT and Raven's were for time 1, $r = -0.41$ ($p < 0.01$); and for time 2, $r = -0.25$ (ns).

The log of the average inspection time over both test occasions was computed for both AIT and VIT (to yield more reliable measures of inspection time and to limit the number of comparisons conducted). Williams' test revealed that the correlation between the composite AIT score and Raven's ($r = -0.62$) was not significantly higher than the correlation between the VIT composite and Raven's ($r = -0.36$); $t = 1.83$, $df = 38$, $p = 0.07$. Finally, the correlation between the composite VIT and AIT measures was $r = 0.45$, $p < 0.01$.

DISCUSSION

The results indicate that this new auditory spatial location task was easy enough for all participants to complete successfully. Only four subjects in sample 1 scored over 200 ms (215, 242, 281 and 298 ms) with only one scoring close to the cut off point of 300 ms. In interpreting these results no problems are raised by excluding subjects who cannot actually do the task. They also indicate that this new task, which bears some formal similarity to conventional VIT tasks, in that both demand a left–right spatial judgment of target location, is significantly related to a measure of intelligence ($r = -0.37$ in the initial sample; -0.61 and -0.57 for time 1 and 2 respectively in the replication sample). As previously suggested (Deary, 1993), the correlation with the measure of fluid intelligence (Raven's) was found to be significant, while that with crystallized intelligence was not, although the difference between these correlations was not significant.

The magnitude of the correlations reported in the present study are at the high end of the range reported in previous studies using the pitch order type of test. Indeed the

correlations in study 2 are higher than any previously reported. Nettelbeck et al. (1986) found AIT correlated with a range of 12 cognitive ability tests at values between -0.33 to -0.38, in Irwin's (1984) study values of -0.23 (RPM) and -0.32 (Mill Hill Vocabulary) were obtained, while in Deary et al.'s (1989) investigation rs of -0.26 and -0.27 were found for the AIT-Alice Heim 6 Verbal (Heim, Watts, & Simmonds, 1983) in children and adults respectively, and -0.36 and -0.45 for the AIT–Non Verbal subscale of the same test, again in children and adults respectively.

The test–retest reliability of this new AIT task was impressively high (0.82) and at least as high as it's existing visual analogue (0.76 in our own results). The current procedure for measuring AIT takes ca. 10 min to complete (although there is considerable variability from subject to subject). Thus, it would be possible to lengthen the AIT task to increase the reliability of this measure even further without running into problems with subject fatigue or boredom.

Using the correction formula for attenuation (e.g., see Nunnally & Bernstein, 1994) and the test–retest reliability coefficient obtained in study 2 to correct for measurement error in AIT, the correlation between AIT and Raven's is estimated at -0.41 in study 1 and -0.67 and 0.63 in study 2.

Although the initial sample of subjects that completed the AIT test in the present report was larger ($n = 67$) than in any of the other AIT investigations of adults cited, we regard this investigation as a preliminary one and the results should be seen in their proper context. On the matter of sample size it is certainly the case that Deary (1995) and Deary et al. (1989) reported large AIT studies with data on 108 children. However, the latter study reported that only 53 subjects were able to complete the successive pitch version of the AIT task that was used, so the present report may be the largest with complete data that has been reported since all participants were able to complete the tasks, even though the spread of AIT values was wide and the distribution of untransformed scores was positively skewed.

In previous studies, e.g., Deary et al. (1989), it has been suggested that poor pitch discrimination may underlie the poor performance of some subjects on the successive pitch version of the AIT. This cannot be the explanation here since pitch discrimination was not required, but there is a possible explanation about why such a basic sensory skill should prove problematic for a proportion of any sample undertaking auditory tasks. It is possible that the difference that has been noted in the relative difficulty of AIT and VIT lies in the fact that, in the population at large, there is a general awareness of visual shortcomings and optical correction is routinely available in visual tasks (subjects wear spectacles if they need to). The same is probably not true for the auditory system where people usually reach the stage of seeking help only when speech recognition becomes a problem; subtle impairments of pitch discrimination and localization abilities may perhaps pass unnoticed. Without a careful audiometric screening of the sample it may be impossible to offer a convincing explanation for any difference in task difficulty between visual and AIT tasks.

It should also be noted that both samples were drawn from an undergraduate population, the MRPM test has only been used on samples in studies carried out at Aberdeen University, and we do not know how well younger or older subjects may adapt to this spatial AIT test. However, overall we believe the results are encouraging and that the version of the "Localization" AIT we have developed may prove to be

more accessible to a wider range of subjects than the pitch order versions previously used. Finally, the results from study 2 suggest that the relationship between AIT and psychometric intelligence ($r = -0.62$) is at least as strong as that observed for its more established visual analogue ($r = -0.36$). The present study is one of the few to directly compare the two in the same sample, and is the first to do so without the interpretive problems that stem from having to limit analysis because a significant proportion of the sample cannot complete one of the tasks. The fact that, under these circumstances, two measures of IT in different sensory modalities are significantly related to a measure of intelligence lends support to the view that some fundamental time limited aspect of processing underlies the relationships.

REFERENCES

Brand, C. R., & Deary, I. (1982). Intelligence and "inspection time." In H. J. Eysenck (Ed.), *A model for intelligence*. New York: Springer.

Crawford, J. R., Deary, I. J., Allan, K. M., & Gustafsson, J. (1998). Evaluating competing models of the relationship between inspection time and psychometric intelligence. *Intelligence, 26*, 27–42.

Deary, I. J. (1986). Inspection time: Discovery of rediscovery? *Personality and Individual Differences, 7*, 625–631.

Deary, I. J. (1993). Inspection time and WAIS-R IQ subtypes: A confirmatory factor analysis study. *Intelligence, 17*, 223–236.

Deary, I. J. (1995). Auditory inspection time and intelligence: What is the direction of causation? *Developmental Psychology, 31*, 237–250.

Deary, I. J., Head, B., & Egan, V. (1989). Auditory inspection time, intelligence and pitch discrimination. *Intelligence, 13*, 135–147.

Heim, A. W., Watts, K. P., & Simmonds, V. (1983). *AH6 group tests of high level intelligence manual*. Windsor, England: NFER-Nelson.

Howard, I. P., & Templeton, W. B. (1966). *Human spatial orientation*. London: Wiley.

Irwin, R. J. (1984). Inspection time and it's relation to intelligence. *Intelligence, 8*, 47–85.

Kranzler, J. H., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence, 13*, 329–347.

Lally, M., & Nettelbeck, T. (1977). Intelligence, reaction time and inspection time. *American Journal of Mental Deficiency, 82*, 273–281.

Nettelbeck, T. (1987). Inspection time and intelligence. In P. A. Vernon (Ed.), *Speed of information processing and intelligence*. Norwood, NJ: Ablex.

Nettelbeck, T., & Lally, M. (1976). Inspection time and measured intelligence. *British Journal of Psychology, 67*, 17–22.

Nettelbeck, T., Edwards, C., & Vreugdenhil, A. (1986). Inspection time and IQ: Evidence for a mental speed–ability association. *Personality and Individual Differences, 7*, 633–642.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*, 3rd edn. New York: McGraw-Hill.

Petrie, R. X. A., & Deary, I. J. (1989). Smoking and human information processing. *Psychopharmacology, 99*, 393–396.

Raven, J. C. (1958). *Standard progressive matrices*. London: Lewis.

Steiger, J. H. (1980). Tests for comparing elements of a correlation matrix. *Psychological Bulletin, 87*, 245–251.

Taylor, M. M., & Creelman, C. D. (1967). PEST: Efficient estimates in probability functions. *Journal of the Acoustical Society of America, 4*, 782–787.

Thompson, S. P. (1877). The phenomena of binaural audition I. *Philosophical Magazine, 4*, 274–276.

Vickers, T., Nettelbeck, T., & Willson, R. J. (1972). Perceptual indices of performance: The measurement of "inspection time" and "noise" in the visual system. *Perception, 1*, 263–295.

Wechsler, D. (1981). *Manual for the Wechsler adult intelligence scale — Revised*. New York: The Psychological Corporation.

Wertheimer, M. (1961). Psychomotor coordination of auditory and visual space at birth. *Science, 134*, 1692–1693.

Williams, E. J. (1959). The comparison of regression variables. *Journals of the Royal Statistical Society (Series B), 21*, 396–399.