Available online at www.sciencedirect.com

SCIENCE DIRECT®

Personality and Individual Differences 38 (2005) 1835–1845

PERSONALITY AND INDIVIDUAL DIFFERENCES

www.elsevier.com/locate/paid

The relationship between three auditory inspection time tasks and general intelligence

Cathal McCrory, Colin Cooper *

School of Psychology, The Queen's University, Belfast BT7 1NN, UK

Received 25 March 2004; received in revised form 20 September 2004; accepted 15 November 2004
Available online 2 February 2005

Abstract

Three inspection-time tasks measuring the amount of time required to discriminate differences in pitch, loudness and phase were administered alongside the Multidimensional Ability Battery and Raven's Advanced Progressive Matrices to 75 undergraduate students. The auditory tasks were administered adaptively, and thresholds were estimated by fitting a logistic function to each set of data. After correcting for restriction of range the three thresholds intercorrelated significantly, and correlated between −0.33 and −0.68 with scores on the ability tests. A composite auditory inspection time score correlated between −0.35 and −0.42 with the ability measures (−0.50 to −0.54 after correction). Although strategy-use enhanced performance on the pitch inspection-time task, strategy use was unrelated to intelligence and did not mediate the correlation between inspection time and intelligence.

1. Introduction

The reductionist approach to the study of human intelligence relates general ability to variations in nervous system physiology, such as the speed or efficiency of neural transmission (e.g., Deary, 2000). In the continued absence of direct and uncontaminated measures of conduction velocity, many researchers have instead used intermediate level constructs as surrogates for direct

* Corresponding author. Tel.: +44 2890 974545.
E-mail address: c.cooper@qub.ac.uk (C. Cooper).

0191-8869/$ - see front matter
doi:10.1016/j.paid.2004.11.012

neuronal measurement. These micro-level tasks are presumed to be so basic as to eliminate the influence of strategic and educational contaminants, so that the differences that emerge between persons can (only) be ascribed to differences in neural mechanisms.

Prominent amongst these tasks is the inspection time (IT) paradigm, which is posited to provide an index of speed of sensory registration, and has been found to be significantly inversely related to measures of psychometric ability (Grudnick & Kranzler, 2001). Although IT has primarily been assessed using visual tasks, Brand and Deary (1982) developed an auditory inspection time (AIT) task that they hoped would be functionally equivalent to the visual inspection time (VIT) procedure. It required participants to discern the relative ordering (high-low or low-high) of two sequentially presented, backwardly masked tones that differed markedly in terms of pitch. The amount of time that the tones had to be presented in order for the participants to have a particular probability of detecting them was established using standard psychophysical methods. This threshold correlated −0.70 ($p < 0.01$) and −0.66 ($p < 0.02$) with Ravens Progressive Matrices and the Mill Hill vocabulary scale, with the cross-modal AIT–VIT association approaching unity. However, the study used a small sample size ($n = 13$) including some with learning difficulties. When these were excluded from the analyses, the correlations were non-significant.

The most frequently used measure of auditory inspection time (Deary, Head, & Egan, 1989) involves two major methodological improvements over this early study. First, as it has been found that some 50% of participants cannot hear the pitch difference even at a very long duration (Irwin, 1984), pre-test screening is used to remove such individuals from the sample: a procedure which Deary et al. (1989) liken to removing those with poor visual acuity from a test involving visual perception. Second, a new and more effective masking procedure is used. This was developed by Nettelbeck, Edwards, and Vreugdenhil (1986) to limit post-stimulus processing of information. Several studies (e.g., Deary, 1994; Langsford, Mackenzie, & Maher, 1994; Olsson, Bjorkman, Haag, & Juslin, 1998) have used this technique and often report significant negative correlations between AIT and measures of general intelligence. However, performance on this task also correlates well with untimed measures of pitch perception, and correlations between AIT and tests of general ability:

(a) Are sometimes not found. Langsford et al. (1994) found AIT-ability associations which were near zero and modest AIT–VIT relationships—circa 10%—which do little to convince sceptics that AIT and VIT index the same underlying construct.
(b) Sometimes vanish when pitch perception is partialled out (Irwin, 1984; Olsson et al., 1998) and sometimes remain significant (Deary, 1994; Deary et al., 1989).

Several other auditory tasks have produced significant negative correlations with tests of general ability. Raz, Willerman, and Yama (1987) gradually reduced the pitch difference between two 20 ms unmasked tones. Threshold measures were then obtained from each person, and Cattell Culture Fair IQ scores correlated −0.47 and −0.54 with this measure of pitch discrimination ability (PDA) for the log transformations of the 1 ms and 9 ms ramps, respectively. This pattern of results was confirmed in a follow up study (ibid.). Raz et al. concluded that the task should perhaps be more appropriately viewed as a measure of stimulus fidelity rather than conduction velocity, as they did not believe that their task imposed any processing speed limitations—an issue with which Deary (1994) disagreed. However, it may be misleading to regard this as a true inspection time task, as it resembles the psychophysical tasks used by Spearman (1904) amongst others to determine whether intelligent individuals can detect fine differences between stimuli rather than whether they can detect a gross difference quickly.

Another task determined the duration for which pairs of tones need to be presented in order for a substantial difference in *loudness* to be detected (Olsson et al., 1998). This is a worthwhile contribution because, unlike pitch perception, there is little evidence to suggest that loudness discrimination ability (LDA) correlates with intelligence. Moreover, the authors believed that the use of loudness might circumvent the problem of participant screening and exclusion which seems to be necessary for pitch discrimination experiments. Scores on Raven's matrices were found to be more strongly related to the loudness inspection time task ($r = 0.36$; $p < 0.01$) than to the pitch inspection time task ($r = 0.21$; n.s.) whilst a separate unspeeded test of pitch discrimination shared significant associations with all three measures, but most notably with the pitch inspection time task. When partial correlations were computed controlling for the effects of pitch discrimination, the correlation between the pitch inspection time task and the Raven's score dropped to zero, but that between loudness inspection time and the Raven's score was negligibly affected (0.30; $p < 0.05$). This suggests that the link between pitch inspection time and *g* may be tenuous and dependent for its existence on variance shared with pitch discrimination ability. Unfortunately, as the authors did not screen participants, the inclusion of participants with poor pitch discrimination may have reduced the chance of finding a substantial correlation between pitch inspection time and general ability.

A third task determines how long a pair of simultaneous tones must be presented in order for a phase difference (as reflected in the apparent location of the tone) to be detected (Parker, Crawford, & Stephen, 1999). This has one clear advantage over the others in that the stimuli are presented simultaneously (one to each ear) as in the visual inspection time tasks, rather than sequentially. Parker and colleagues reported correlations ranging from −0.37 ($p < 0.01$) to −0.61 ($p < 0.001$) between performance on the task and a modified version of the Raven's matrices. After correcting for unreliability in the experimental measures and attenuation of range, the correlations rose to −0.41 and −0.67, respectively. In evaluating the effectiveness of the enterprise, Parker et al. (1999) claimed that the moderate cross-modal correlation ($r = -0.45$; $p < 0.01$) between their composite AIT and VIT measures attested to the validity of this new task for capturing the same temporal limitation in information processing as the original VIT procedure.

The overall conclusion from these auditory inspection time tasks is somewhat damaging for those (e.g., Jensen, 1998) who suggest that mental speed is the very core of *g*. The correlation between pitch-based inspection time tasks and *g* vary from near zero (Langsford et al., 1994; Olsson et al., 1998) to a relatively modest 0.4 (Deary, 1994) and may be mediated somewhat (or entirely) by pitch perception ability. Whilst two other promising auditory tasks have been developed by Olsson et al. (1998) and Parker et al. (1999), their results do not appear to have been replicated. Nor does there appear to have been any attempt to administer all three tasks in an attempt to discover the extent to which they measure common variance. There are thus three main unresolved issues.

(a) Do all three AIT tasks measure the same underlying construct?
(b) If so, is this essentially identical to visual inspection time?

(c) Do the correlations between inspection time tasks and tests of general ability arise because of the *g* component of the general ability tests, or narrower abilities, such as pitch perception or perceptual speed, or strategies?

The present study seeks to address the first and third of these issues through administering the three auditory inspection time tasks alongside measures of general ability. It can consequently answer a number of important questions. It can show which task demonstrates the strongest associations with psychometric intelligence. Secondly, it shows the extent to which the different tasks share common variance. This is important given the fairly modest correlation ($r = 0.36$; $p < 0.001$) between the pitch and loudness inspection time tasks reported by Olsson et al. (1998). Finally, Langsford et al. (1994) suggested (but did not test) the idea that the pitch inspection time task devised by Deary may be correlated with *g* because intelligent participants devise and use strategies to solve the task. Given the problems associated with strategy use in visual inspection time tasks (e.g., Simpson & Deary, 1997), it seems prudent to test this empirically.

2. Method

2.1. Participants

An opportunistic sample of 75 participants (22 male, 53 female) drawn predominantly from the university population volunteered to take part. Mean age for the sample was 20.14 years (SD = 3.4) and no participant reported any hearing problems.

2.2. Cognitive ability measures

The multidimensional aptitude battery II (MAB; Jackson, 1998) is a group administered and highly reliable alternative to the WAIS. It has previously been used in the UK by Bates and Eysenck (1993) and yields scores on verbal and performance IQ as well as an overall IQ estimate. Raven's Advanced Progressive Matrices (RAPM; Raven, 1965) is considered a marker variable for Gf and gives a good approximation to *g* (Kline, 2000) and was also used. The number-checking scale from the Ekstrom, French, and Harman (1976) kit was also included as a measure of perceptual speed.

2.3. Auditory stimuli

2.3.1. Auditory inspection time-pitch (AIT-P)

The AIT-P task was derived from the work of Deary et al. (1989). Each trial consisted of a cue-tone (832 Hz) lasting 500 ms which was designed to promote attention, 1000 ms of silence, then a pair of sinusoidal tones of different frequencies (880 Hz/784 Hz or 784 Hz/880 Hz) and variable duration, with zero inter-stimulus interval, followed by rapidly alternating 10 ms bursts of both stimulus tones. Intensity remained constant throughout at 60 dB, measured using an artificial ear. The amplitude envelopes were ramped to reduce click artefacts. Participants were asked to decide whether the high-pitched (880 Hz) tone or the lower-pitched (784 Hz) tone was presented first.

2.3.2. Auditory inspection time-loudness (AIT-L)

The AIT-L task which was taken from the work of Olsson et al. (1998) was identical except that stimuli were of a constant frequency (832 Hz) but differed in loudness (61.5 and 55.5 dB). Participants determined the presentation order (loud-faint, or faint-loud) of pairs of tones.

2.3.3. Auditory inspection time-spatial (AIT-S)

The AIT-S task was designed to replicate the work of Parker et al. (1999) and required participants to make a spatial judgement concerning the apparent location of a target tone delivered to the left or right hand side of the mid-line axis of the head. Each trial consisted of a cue tone lasting 500 ms, 1000 ms of noise, a stimulus tone pair—one tone being lateralised by a phase shift of 40° relative to the other—followed by 500 ms of square wave masking noise at 81 dB. The frequency of all tones was held constant at 450 Hz while intensity level for the target tones was 58 dB.

All stimuli were generated in real time using MATLAB 6.1 and delivered to headphones via a SoundBlaster compatible sound card on an IBM compatible PC. Intensity levels were measured using a Bruel and Kjaer 4152 artificial ear with a 4160 1 in. condenser microphone, 2610 measuring amplifier, 2619 pre-amplifier, and a 4230 1 kHz, 94 dB class 1 calibrator. To ensure that the resulting stimuli were 'clean' of any transient noise, the sine waves were also captured and examined visually using an auditory graphing package. The empirical literature was used as a guide to inform choice of stimulus step size. On the basis of this consultation it was decided that 10 ms steps (range 200–10 ms) would be adequate for the pitch and loudness tasks, while 5 ms steps (range 200–5 ms) would be required for the spatial task. Pilot testing confirmed the adequacy of these values for each of the auditory procedures.

2.4. Psychophysical technique

Little is known about the effects of practice on IT, but it is clearly dangerous to use descending staircase procedures whereby participants work systematically downward from some asymptotic level of performance to some lower bound (e.g., Olsson et al., 1998). This introduces the potential confound that participants of higher ability register shorter IT's because practice at high durations affords an opportunity to develop effective response strategies.

Rather than presenting several blocks of trials each block being of a particular difficulty level, adaptive staircase procedures (e.g., Barrett, Petrides, & Eysenck, 1998; Taylor & Creelman, 1967; Wetherill & Levitt, 1965) use information garnered on previous trials to 'home in' on the threshold rather quickly. Several correct responses at a particular duration lead to a reduction in exposure time; an incorrect response causes an increase. A mathematical model guides the estimation procedure which iterates until a set number of trials have been administered or the threshold has been accurately determined.

2.5. Procedure

All 75 participants were individually administered the multidimensional ability battery using standard instructions. The 12 items in Book I of the Raven's Advanced Progressive Matrices were administered as a 10-min speeded test, while the 36 item set was given as a power test with no time restrictions.

The three auditory tasks were administered by computer. The task interface consisted of two on-screen buttons, one for each of the response alternatives, and all responses were made via mouse clicks. Participants were encouraged to rest between trials if they wished. During practice and prior to the experimental session, participants were repeatedly cautioned that it was accuracy of response that was important—not speed—and that they could, therefore, deliberate over their responses for as long as they thought necessary. Participants were asked to guess if unsure of the correct response for any trial.

The existence of practice effects in the IT literature is a robust finding, and while the reasons for this are poorly understood, there has been some discussion concerning their potential impact. Thus while Chaiken and Young (1993) found that practice strengthens the IT-ability association. Bors and MacLeod (1996) presented evidence to the contrary. It seemed prudent, therefore, to counterbalance order of administration so as to mitigate the influence of any transference effects which may exist. All three auditory tasks were undertaken in a sound-attenuated room. Before being allowed to proceed into each main task, participants were presented with 10 practice trials delivered at the highest experimental stimulus duration of 200 ms. Those failing to meet the criterion of 90% correct responding underwent further training at 300 ms. Having demonstrated substantive improvement (10/10 correct at 300 ms), these subjects then retook the trials at 200 ms, otherwise they were thanked and dismissed. Four participants were removed from the pitch task in this manner, while the corresponding number for the loudness and phase tasks was three and two exclusions, respectively.

The stimulus duration for each trial was determined adaptively using the BEST PEST estimation procedure (Liebermann & Pentland, 1982) which uses maximum likelihood estimation to determine the most informative exposure level for each trial, based on past performance. The algorithm terminated after 70 trials as it was felt that this would give a fairly good approximation to threshold when modelled using our logistic function (see below) while not unduly fatiguing the participants.

Upon successful completion of each of the tasks, participants were quizzed to determine whether they had used a response strategy to guide discriminative performance. Following Langsford et al. (1994), this consisted of a simple probe question of the form 'how could you tell' which was designed to facilitate consideration and explication. Written responses were subsequently examined and respondents categorised as users or non-users.

3. Analysis

It is not possible to estimate the threshold by merely totalling correct responses, as the adaptive psychophysical procedures ensure that each person receives a different number of trials at each duration. It is instead necessary to fit each person's psychophysical function to the data from each task. Bates and Eysenck (1993) fitted a third order polynomial function to inspection time data, but gave no theoretical justification as to why the polynomial was to be preferred to other curve fitting techniques: a polynomial function does not necessarily imply that stimuli of zero duration would be expected to show a chance success rate, or that very easy stimuli should be perfectly discriminated. The logistic function (which closely resembles the cumulative normal distribution) shows these two desirable properties. Hence each participant's data were fitted using a logistic function defined by two parameters. These were the threshold (defined as the stimulus value where a participant had a 0.75 chance of making a correct decision, 0.75 being midway between chance and perfect performance for these two-choice tasks) and the slope of the psychometric function (the gradient of the logistic curve at the threshold, which might reasonably be expected to vary from person to person). The average proportion of correct responses at each duration (x_i) was calculated, and the two-parameter logistic function shown below was fitted to the data using weighted least squares.

$$\text{expected proportion correct} = \frac{1}{2} + \frac{1}{2}\left(\frac{e^{a(x_i-b)}}{1+e^{a(x_i-b)}}\right)$$

Parameter b represents the person's threshold, and parameter a the slope of their psychometric function. These threshold estimates for each person on the various inspection time tasks were used for all subsequent analyses. The computer program that performed these analyses may be obtained from the second author.

4. Results

These University students performed well above average on the verbal ($M = 115.4$, SD = 7.6), performance ($M = 110.5$, SD = 9.9) and full scale IQ ($M = 114.1$, SD = 8.23) scales of the MAB. Set I of the Raven's Advanced Progressive Matrices showed a mean of 10.88 (SD = 1.24) and set II a mean of 24.3 (SD = 5.8). Inspection of the threshold distributions of the various auditory inspection time tasks revealed that while loudness thresholds demonstrated acceptable symmetry, the pitch and spatial thresholds were positively skewed. A square root transformation was applied to both sets of data, and as the spatial threshold remained significantly skewed after both square root and inverse transformations had been applied, non-parametric correlations were also computed. However, as there were few differences the parametric statistics are reported below. As the IT-ability association represents a well-established empirical finding, the statistical significance of each correlation was evaluated using a one-tailed test.

Table 1 summarises these results. As predicted, inspection time threshold was found to be inversely related to the various measures of psychometric ability, and this pattern was consistent across all three auditory tasks. This signifies a general tendency for those of higher ability to make perceptual discriminations at lower exposure durations. Loudness threshold showed only a weak and non-significant association with the various measures of ability, though that with FSIQ bordered significance ($r = 0.19$; $p = 0.059$). Given the considerable restriction of range evident in our sample, the AIT-ability correlations reported here are likely to be substantially lower than those found in the general population. Correcting the correlations for restrictions of range raises the correlations between full scale IQ and the three auditory thresholds to −0.68, −0.33 and −0.44, respectively, and each of the correlations between loudness thresholds and the cognitive tasks is at least 0.30 ($p < 0.01$). It is no simple matter to estimate the reliability of thresholds established using an adaptive program, and so no attempt was made to correct the thresholds for the less-than-perfect reliability of the threshold estimates and ability tests. A canonical correlation performed between the three thresholds and full scale IQ and the two Raven's matrices scores was 0.53, with all variables showing loadings above 0.4 on the first canonical variate.

Table 1
Two-tailed correlations between auditory inspection time thresholds and cognitive abilities

	AIT-loudness	AIT-spatial	Verbal IQ	Performance IQ	Full scale IQ	Ravens 1	Ravens 2	Perceptual speed
AIT-pitch	0.33** (0.33**)	0.25* (0.24*)	−0.40*** (−0.40***)	−0.39*** (−0.39***)	−0.45*** (0.44***)	−0.42*** (−0.42***)	−0.43*** (−0.42***)	−0.08
AIT-loudness		0.19 (0.16)	−0.17 (−0.15)	−0.17 (−0.15)	−0.19 (−0.16)	−0.24* (−0.23*)	−0.19 (−0.17)	−0.08
AIT-spatial			−0.26* (−0.21*)	−0.22 (−0.14)	−0.26* (−0.20)	−0.17 (−0.11)	−0.29** (−0.25*)	−0.25*
Verbal IQ				0.57***	n/a	0.55***	0.59***	
Performance IQ					n/a	0.54***	0.62***	
Full scale IQ						0.61***	0.69***	
Ravens 1							0.57***	
Mean standardised threshold			−0.37***	−0.35***	−0.40***	−0.38***	−0.42***	

Figures in brackets are partial correlations, controlling for perceptual speed.
* Significant at the 0.05 level.
** Significant at the 0.01 level.
*** Significant at the 0.001 level.

Table 2
Inspection time thresholds for strategy users and non-users

	Strategy users			Non-users			t	p
	n	Mean threshold	SD	n	Mean threshold	SD		
AIT-pitch	41	91.0	47.6	30	126.0	56.0	2.81	<0.05
AIT-loudness	36	122.1	37.0	37	126.5	47.5	0.4	n.s.
AIT-spatial	26	25.0	35.4	47	48.5	54.6	1.97	n.s.

Table 2 shows the relationship between threshold and strategy use. Although strategy users were characterized by lower mean inspection time thresholds across all three auditory tasks (significantly so in the case of the pitch task) t-tests showed that strategy use was unrelated to any measure of IQ (all p's > 0.05). When the IT-ability correlations were computed for each group, the associations were somewhat (though never significantly) larger for strategy users in the case of the loudness and spatial tasks.

5. Discussion

There can be little doubt that the three auditory thresholds are related to cognitive ability, and that some of these relationships are substantial. We believe that these results support the neural processing theory of intelligence for three reasons. First, the auditory tasks were quite different in nature. Hence the correlations between the three tasks are unlikely to be caused by common higher-level processes, such as pitch discrimination. Second, the forced-choice design of the experiment means that variations in performance from person to person cannot be attributed to variations in the willingness to guess when uncertain of the correct answer. Finally, the adaptive algorithm presented each individual with rather few extremely hard or easy trials, which should help to avert boredom or frustration, and the stimulus durations that are close to an individual's threshold are not bunched together at the end of the experiment.

It is also clear from Table 1 that even within this homogenous sample of intelligent young adults, there is some overlap between the various tasks although the uncorrected correlation between the loudness and spatial inspection time tasks fell just short of statistical significance ($r = 0.19$; $p = 0.056$). A subsequent principal components analysis of the three auditory tests revealed that each of the tasks had a large and significant loading on the first unrotated factor (0.62–0.78), which accounted for some 50% of the total variance. A composite measure calculated by summing the three standardised auditory thresholds correlated −0.40 ($p < 0.001$) with FSIQ before correction for range restriction (0.53 after correction).

The number of participants who had to be rejected from the pitch discrimination task was lower than usual. This may be because the present study used a sinusoidal waveform in order to reduce the number of harmonics, whereas Deary et al. (1989) used a square wave.

The finding that some visual inspection time tasks correlate higher with tests of perceptual speed than with the more *g*-loaded primary abilities (e.g., Cooper, Kline, & Maclaurin-Jones, 1986) raises the possibility that the correlation between *g* and IT is mediated by (visual) perceptual speed. The partial correlations shown in Table 1 show that this is not the case for auditory inspection time. It is also unlikely that individual differences in strategy-use can account for the correlations between IT and intelligence. Whilst it seems that the use of appropriate strategies can enhance performance at the pitch discrimination task, there is no hint that strategy use can affect performance on the loudness task, and strategy use was not significantly related to performance on the phase-detection task. Furthermore, participants who developed strategies on a particular task were no more or less intelligent than those who did not. In sympathy with existing research in the visual domain, this would suggest paradoxically, that if strategy use exerts any influence, it may well serve to obscure the magnitude of the association by contributing error variance to the estimate of IT.

Whilst previous research had found that performance on each individual auditory task was significantly associated with general intelligence, only one of these findings had been replicated. In addition, it was unclear whether there was appreciable correlation between the three auditory tasks, and whether the correlations between a task and *g* reflected task-specific variance, or some broader ability—perhaps reflecting neural conduction velocity. As the canonical variate showed substantial loadings from all IT tasks, and the average standardised threshold showed a substantial correlation with general intelligence, it seems that these three IT tasks do indeed measure a common characteristic which is related to *g*. To further strengthen this conclusion, a general ability factor extracted from the battery of MAB and Raven's scores correlated 0.45 ($p < 0.001$) with the latent auditory IT factor.

Deary (2000) has observed that one of the less impressive features of some research into the components of intelligence is that both the ability tests and the cognitive measures may share some common cognitive processes—for example, reading speed may influence performance on both the ability test and the 'lower level' task. It is, therefore, important to study the relationship between ability tests and structurally different tasks in order to reduce the chance that correlations may arise because of variations in performance on cognitive components which are common to both tasks. The results shown above reveal broadly similar relationships are found between three different auditory tasks a variety of (visual) ability tests—some speeded, some not, some involving language, some not. They provide further evidence for a link between speed of information processing and general intelligence.

References

Barrett, P. T., Petrides, K. V., & Eysenck, H. J. (1998). Estimating inspection time: response probabilities, the BRAT IT algorithm, and IQ correlations. *Personality and Individual Differences, 24*, 405–419.

Bates, T. C., & Eysenck, H. J. (1993). Intelligence, inspection time, and decision time. *Intelligence, 17*, 523–531.

Bors, D. A., & MacLeod, C. M. (1996). Attention, information processing, and IQ. *International Journal of Psychology, 31*, 34–52.

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

Chaiken, S. R., & Young, R. K. (1993). Inspection time and intelligence: attempts to eliminate the apparent movement strategy. *American Journal of Psychology, 106*, 191–210.

Cooper, C., Kline, P., & Maclaurin-Jones, E. (1986). Inspection time and primary mental abilities. *British Journal of Educational Psychology, 56*, 304–308.

Deary, I. J. (1994). Intelligence and auditory discrimination: separating processing speed and fidelity of stimulus representation. *Intelligence, 18*, 189–213.

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain*. Oxford: Oxford University Press.

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

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). *Manual for the kit of factor-referenced cognitive tests*. Princeton, NJ: Educational Testing Service.

Grudnick, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence, 29*, 523–535.

Irwin, R. J. (1984). Inspection time and its relation to intelligence. *Intelligence, 8*, 47–65.

Jackson, D. (1998). The Multidimensional Aptitude Battery II (MAB). Sigma Assessment Systems.

Jensen, A. R. (1998). *The g* factor: the science of mental ability. Westport, CT: Praegar.

Kline, P. (2000). *The handbook of psychological testing* (3rd ed.). London: Routledge.

Langsford, P. B., Mackenzie, B. D., & Maher, D. P. (1994). Auditory inspection time, sustained attention, and the fundamentality of mental speed. *Personality and Individual Differences, 16*, 487–497.

Liebermann, H. R., & Pentland, A. P. (1982). Microcomputer-based estimation of psychophysical thresholds: the best PEST. *Behavior Research Methods and Instrumentation, 14*(1), 21–25.

Nettelbeck, T., Edwards, C., & Vreugdenhil, A. (1986). Inspection time and IQ: evidence for a mental speed-ability association. *Personality and Individual Differences, 7*, 633–641.

Olsson, H., Bjorkman, C., Haag, K., & Juslin, P. (1998). Auditory inspection time: on the importance of selecting the appropriate sensory continuum. *Personality and Individual Differences, 25*, 627–634.

Parker, D. M., Crawford, J. R., & Stephen, E. (1999). Auditory inspection time and intelligence: a new spatial localisation task. *Intelligence, 27*, 131–139.

Raven, J. C. (1965). *Advanced progressive matrices*. London: H.K. Lewis.

Raz, N., Willerman, L., & Yama, M. (1987). On sense and senses: intelligence and auditory information processing. *Personality and Individual Differences, 8*, 201–210.

Simpson, C. R., & Deary, I. J. (1997). Strategy use and feedback in inspection time. *Personality and Individual Differences, 23*, 787–797.

Spearman, C. (1904). 'General intelligence' objectively determined and measured. *American Journal of Psychology, 15*, 201–293.

Taylor, M. M., & Creelman, C. D. (1967). PEST: efficient estimates on probability functions. *Journal of the Acoustical Society of America, 41*, 782–787.

Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical Statistical Psychology, 18*, 1–10.