REACTION TIME AND INTELLIGENCE

Arthur R. Jensen

University of California

Berkeley, California, U.S.A.

Abstract

Measurements of various parameters derived from different reaction time (RT) paradigms are found to be correlated with psychometric measurements of general mental ability. Such RT-derived measurements, when combined in a multiple regression equation, predict some 50 percent or more of the variance in IQ or *g*. This relationship of IQ or *g* to RT parameters indicates that our standard IQ tests tap fundamental processes involved in individual differences in specific knowledge, acquired skills, or cultural background.

This article reviews the main currents in research on the relationship of reaction time (RT) to general intelligence and other psychometric mental abilities.

The first conclusion we can draw with confidence is that RT parameters in a variety of paradigms are significantly related to scores on standard tests of intelligence and other psychometric abilities. As I have noted elsewhere (Jensen, 1979), the study of RT as a measure of mental ability got off to a bad start in the early history of psychology, for a number of reasons, largely due to psychometric naivete and inadequate statistical methods. Modern investigators have been more successful in finding substantial and replicable relationships between RT and IQ.

Correlation coefficients between RT and IQ are not as impressive or as consistent as are mean differences in RT between different criterion groups selected on the basis of IQ or other psychometric indices of ability. Correlations between RT and IQ can be generally characterized as fairly low. But in the entire literature on RT and IQ there are virtually no correlations on the "wrong" side of zero. Most *r*s fall in the range from 0 to -.50, with a mode in the -.30's. A correlation of -.50 is about maximum. It is theoretically important to understand the causes of this apparent low correlation ceiling. But there is no doubt that the present evidence overwhelmingly rejects the null hypothesis. This is true of simple RT as well as choice RT (also termed discriminative or disjunctive RT). Both simple and choice RT are negatively correlated with IQ.

Mean differences in RT (or in various parameters of RT) between criterion groups selected for differences in ability as measured by psychometric tests or scholastic performance always give more clearly impressive evidence of a relationship between RT and general ability than the correlation coefficient. The mean RT difference between criterion groups is often of at least the same magnitude as the mean IQ difference between the groups, when the mean differences in RT and IQ are both expressed in standard deviation or σ units. We have found that borderline retarded young adults, with a mean IQ of about 70, differ from university students about 6σ on Raven's Matrices. These groups differ about 7 σ(σ of the university students) in mean RT. University students compared with academically less highly selected students of the same age in a two-year vocational college differ about 1 σ in scholastic aptitude scores; in mean RT they differ 1.2 σ in terms of the vocational college σ and 1.9σ in terms of the university σ .

From the standpoint of psychometrics, I think the most important conclusion from all the RT research is that it proves beyond reasonable doubt that our present standard tests of IQ measure, in part, some basic intrinsic aspect of mental ability and not merely individual differences in acquired specific knowledge, scholastic skills, and cultural background. The RT parameters derived from typical procedures cannot possibly measure knowledge, intellectual skills, or cultural background in any accepted meaning of these terms. Yet these RT parameters show significant correlations with scores on standard tests of mental ability and scholastic achievement and show considerable mean differences between criterion groups selected on such measures.

Three Basic RT Paradigms

There are three distinct and basic paradigms in RT research. Each paradigm measures different facets of information processing speed, and each has shown a relationshp to psychometric variables. I shall refer to these paradigms by the names of the three psychologists who initiated them.

The *Hick* paradigm measures the linear increase in RT to visual or auditory stimuli as a function of the amount of information (measured as bits=$\log_2$ of the number of stimulus alternatives) conveyed by the reaction stimulus, but involves no need to access either short-term or long-term memory (STM or LTM). The classical experiment contrasting simple and two-choice RT is the simplest example of the Hick paradigm, involving 0 and 1 bit of information, respectively.

The *Sternberg* (1966) paradigm presents the subject with a small set of digits (or letters) followed immediately by a single "probe" digit to which the subject responds "yes" or "no" as to whether the probe was or was not included in the set. The *S*'s RT or decision time in pressing the "yes" or "no" key involves speed of scanning STM, and RT increases as a linear function of the number of *items* in the set, unlike the Hick phenomenon, in which RT increases as a linear function of the *logarithm* (to the base 2) of the number of stimulus alternatives.

The *Posner* (1969) paradigm contrasts discriminative ("same" versus "different") RTs to pairs of stimuli which are the same or different either physically or semantically. For example, the letters *AA* are physically the same, whereas *Aa* are physically different but semantically the same. When *Ss* are instructed to respond "same" or "different" to the physical stimulus, RTs are faster than when *Ss* must respond to the semantic meaning. The physical discrimination is essentially the same as classical discriminative RT, but RT in the semantic discrimination involves access to semantic codes in LTM, which takes considerably more time than physical discriminative RT. The difference between semantic and physical RT thus measures access time to highly overlearned semantic codes in LTM. Interestingly, Hunt (1976) and his co-workers have found that this measurement is especially related to verbal ability as measured by the Scholastic Aptitude Test (SAT-V) in university students.

Typical Findings

Posner Paradigm. Figure 1 shows the results of a study by Hunt (1976) using the Posner paradigm with groups of university students scoring high or low on the SAT-Verbal. *AA* represents the physical identity choice (same-different) RT task; *Aa* represents the semantic identity task. University students require on the average about 75 milliseconds more time to respond to *Aa* than to *AA* types, which is the time taken by semantic encoding of the stimulus. Two features of Figure 1 are particularly interesting in relation to findings from the Sternberg and Hick paradigms: (1) the high and low groups on SAT-V show a mean difference in RTs even on the physical, nonsemantic identity task, which is essentially just a form of classical two-choice discriminative RT;

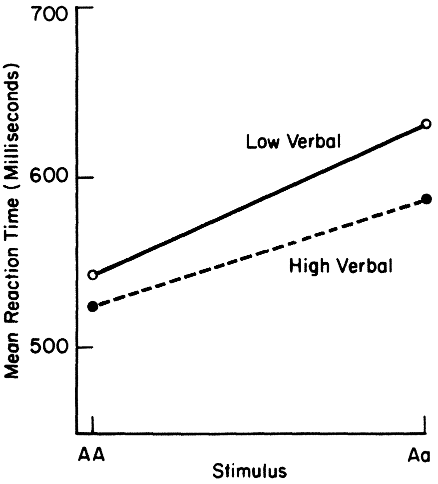

Figure 1. Time required to recognize physical or semantic identity of letter pairs by university students who score in the upper (high) or lower (low) quartile on the SAT-Verbal. (After Hunt, 1976, Table 1, p. 244.)

and (2) the mean RTs are all greater than 500 milliseconds, which is appreciably slower than the RTs of university students in the Hick paradigm, even for RT to three bits (i.e., eight stimulus alternatives) of information, which has a mean RT of 350 to 400 msec. Because the times needed for physical discrimination between extremely familiar stimuli and for accessing simple, highly overlearned semantic codes in LTM are in excess of the RTs to three bits of information in the Hick paradigm, it suggests that performance in our Hick paradigm does not depend on discriminating anything as difficult as familiar letters or accessing anything in LTM. The average RT difference between AA and Aa (i.e., semantic encoding time) of 75 msec for Hunt's university students is exactly the same as the difference in RT between 0 and 3 bits of information in our Hick paradigm with university students.

Sternberg Paradigm. Figure 2 shows Sternberg STM-scan RTs for groups of fifth and sixth grade children with moderate and high IQs, from a study by McCauley et al. (1976). The intercepts and slopes of the moderate and high IQ groups both differ significantly. Stanford University students given a comparable Sternberg task (Chiang and Atkinson, 1976) show much lower intercepts (about 400 msec) but show about the same slope (i.e., a scan rate of 42 msec per digit in target set) as the high IQ children (with a scan rate of 40 msec per digit), whose IQs (with a mean of 126) are probably close to the IQs of the Stanford students. The moderate IQ group has a significantly greater slope (i.e., slower STM scanning rate) of 58 msec per digit. IQ

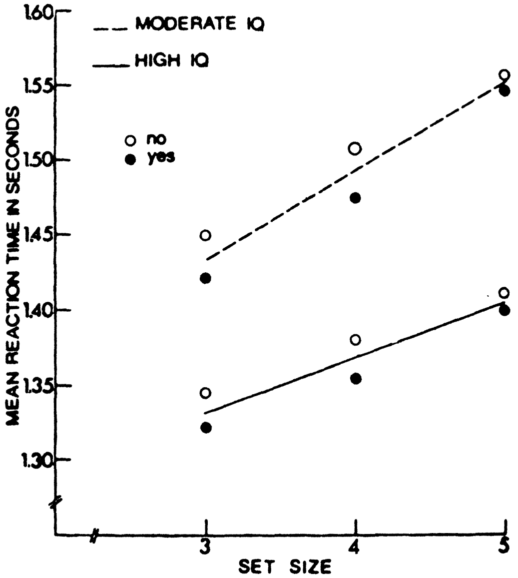

Figure 2. Mean RTs for correct "yes" and "no" (i.e., presence or absence of probe digit in target set) for moderate IQ (95 or below, X=88) and high IQ (115 or above, X=126) fifth and sixth grade children. The equations for the two lines are: moderate IQ RT = 1265 + 58s, and high IQ RT = 1210 + 40s, where RT is in milliseconds and s = number of digits in the target set. (From McCauley et al., 1976.)

would appear to be more crucial than mental age for short-term memory scan rate. This has interesting implications for scanning and rehearsal of information in STM to consolidate it into LTM. In terms of such a model, and in view of the observed differences in scan rates as a function of IQ, it should seem little wonder that high IQ persons in general know more about nearly everything than persons with low IQs. Snow, Marshalek, and Lohman (1976) were able to "predict" the intercepts and slopes of the Sternberg memory scan paradigm for individual Stanford students with multiple *R*'s of .88 and .70, respectively, using scores on several psychometric tests (in addition to sex). The intercept and slope parameters of the Sternberg scan, on the other hand, predicted each of four factor scores derived from a large battery of psychometric tests with *R*'s between .33 and .56. SAT-Verbal and SAT-Quantitative scores were predicted with *R*'s of .54 and .21, respectively. Remember, we are dealing here with the quite restricted range of ability in Stanford University students.

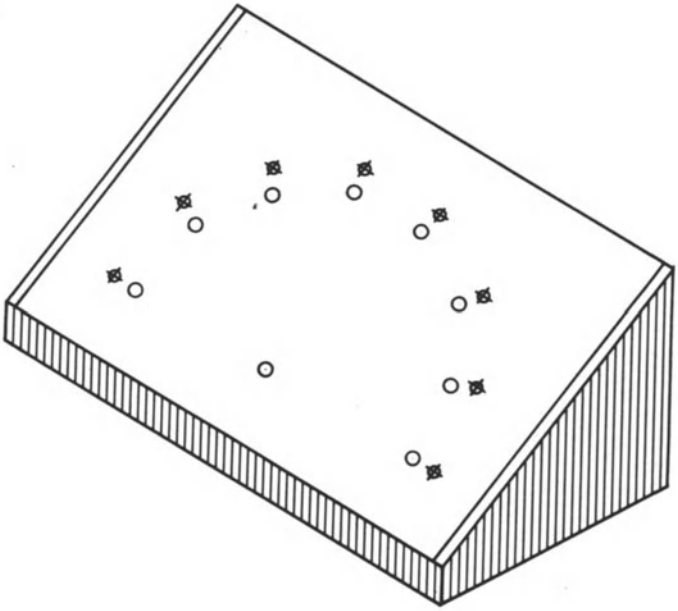

Figure 3. Subject's console of the reaction time-movement time apparatus. Push buttons indicated by circles, green jeweled lights by circled crosses. The "home" button is in the lower center.

Hick Paradigm. This is an elaboration of simple and choice RT. Hick (1952) discovered that RT increases linearly as a function of $\log_2$ of the number of choices or stimulus alternatives -- a phenomenon now known as Hick's Law. I have been doing studies of this paradigm, using an apparatus show in Figure 3. (It is described in more detail by Jensen and Munro, 1979.) The S places his index finger on the "home" button, a "beep" warning signal is sounded for 1 second, and after a random interval of 1 to 4 seconds one of the green lights goes on. The S must turn off the light as fast as possible by touching the button adjacent to it. The time between the light's going on and removal of the S's finger from the home button is the RT. The interval from release of the home button to turning out the light is the movement time (MT). Templates can be placed over the console to expose any number of light/button alternatives from 1 to 8. We have most often used 1, 2, 4, and 8 alternatives, corresponding to 0, 1, 2, and 3 bits of information. Following instructions and several practice trials, Ss are usually given 15 trials on each number of alternatives (60 trials in all) in a single session lasting about 20 minutes.

To insure that RT is in fact related to intelligence, I have sought correlations between RT parameters and IQ in criterion groups selected from every available level of the IQ distribution, ranging from the severely retarded (with IQs of 15 to 50), to the mildly retarded and borderline (IQs 50 to 80 or so), to average and bright school children and average young adults, and to university students with IQs above the 95th percentile of population norms. We have now tested nine such groups totalling about 800 persons. Without exception, groups differing in mean IQ also differ very significantly in the expected direction in a number of RT (and also MT) parameters. Also, within every group we have tested, the RT parameters are significantly correlated with IQ, with all correlations in the theoretically expected direction, mostly ranging between about .20 and .50. Many of these findings have been described elsewhere (Jensen, 1979; Jensen and Munro, 1979).

We describe an individual's RT performance in the Hick paradigm in terms of three parameters: the slope of the linear regression of RT on bits, the intercept of the regression line, and the intraindividual variability over trials, which is indexed by the root mean square of the variances among trials within bits. (We have also used the slope of the regression of the standard deviation among trials, as a function of bits.) Individual differences in all of the RT parameters are positively intercorrelated. Other investigators, too, have found a positive correlation between intercepts and slopes in the Sternberg paradigm (Dugas and Kellas, 1974; Snow et. al, 1976; Oswald, 1971). Moreover, all these parameters are negatively correlated with g. At first I expected that intercepts, which represent simple RT, and hence involve little or no information processing, would not be correlated with IQ. I was wrong; intercepts are negatively correlated with IQ, although within fairly homogeneous criterion groups the correlations are often too small to be significant and are almost invariably smaller than the correlations of slope and intraindividual variability with IQ. Figure 4 shows the intercepts and slopes of RT data from seven criterion groups. None of the regression lines except that of the severely retarded group shows a significant nonlinear trend.

Intraindividual Variability. Surprisingly little attention was ever given to intraindividual variability in RT in the older literature. Yet it is this aspect of individual differences in RT that seems to be the most profoundly related to intelligence level, as has been frequently noted by investigators of RT in the mentally retarded (Berkson and Baumeister, 1967; Baumeister and Kellas, 1968a, 1968b, 1968c; Liebert and Baumeister, 1973; Wade, Newell, and Wallace, 1978; Vernon, 1979). The negative correlation

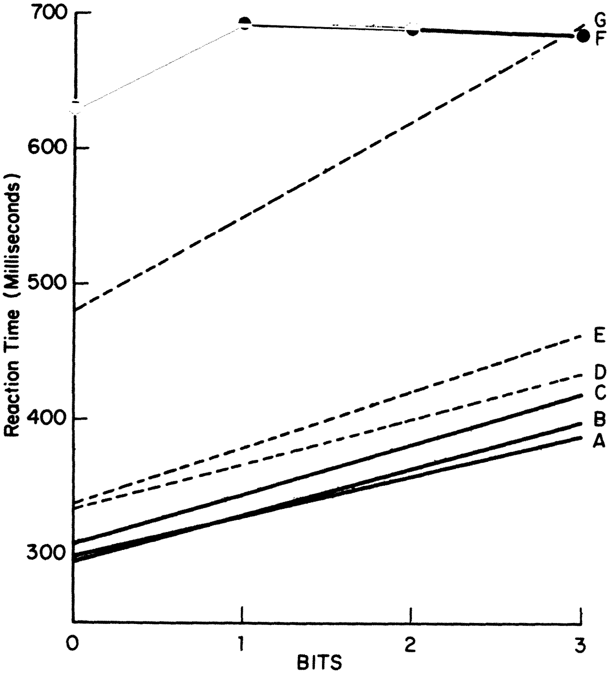

Figure 4. RT as a function of bits, illustrating Hick's law and differences in intercepts and slopes, for diverse groups varying in age and intelligence: A - university students, B - ninth grade girls, C - 6th graders in a high SES-high IQ school, D and E - white and black, respectively, male vocational college freshmen with approximately equal scholastic aptitude scores, F - severely mentally retarded young adults (mean IQ 39), G - mildly retarded and borderline young adults (mean IQ 70). (From Jensen, 1979.)

between intraindividual variability in RT and IQ is found within every level of intelligence, from the severely retarded to university students.

I have looked more closely at this phenomenon in our data by rank ordering each S's RTs from the shortest to the longest in 15 trials. (The 15th rank is eliminated to get rid of possible outliers.) Figure 5 shows the means of the ranked RTs of 46 mildly retarded (IQ 70) and 50 bright normal (IQ 120) young adults each given 15 trials on simple (0 bit) RT. Note that even on the fastest trial (rank 1) the retarded and normal Ss differ by 111 msec. In fact, the normal Ss' slowest RT (rank 14) is 32 msec

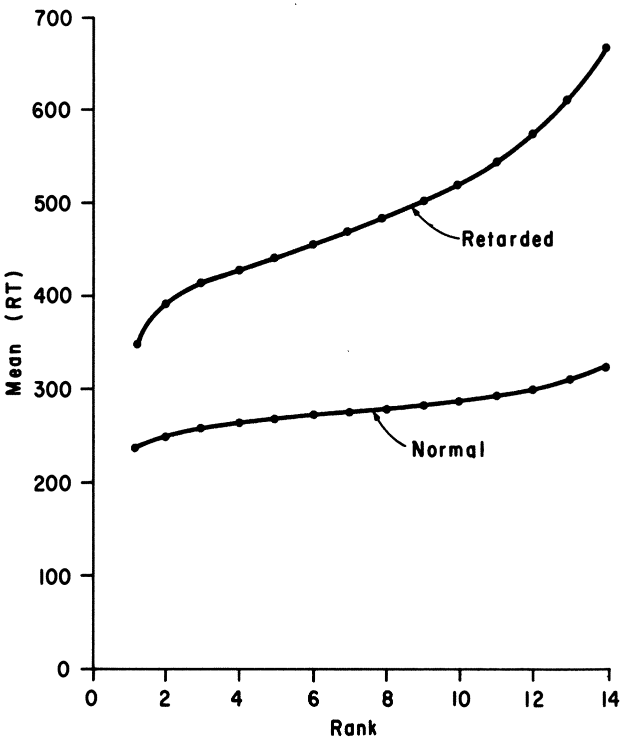

Figure 5. Mean simple RT plotted after ranking RTs on 15 trials from the fastest to the slowest trial (omitting the 15th rank) for retarded and normal Ss.

shorter that the retardates' fastest RT. In case anyone might think these are trivial differences, let us look at then in terms of standard deviation or σ units, i.e. (normal RT minus retarded RT)/σ, as shown for simple RT in Figure 6 for σ differences based on both normal and retarded σ units. The fastest simple RT of retardates and normals differs 1.2σ in terms of the retardates' σ units and 4.8 σ in terms of the normals' σ units.

The fact that even the fastest RTs of the retarded Ss are slower than the RTs of normals, even for simple RT, suggests that the difference is at some very basic, one might almost say neural, level and not at any very complex level of information processing. Possibly even simpler responses might show reliable speed differences related to general intelligence.

Combining RTs in the Hick, Sternberg, and Posner Paradigms.

If RT and the derived parameters in the three different

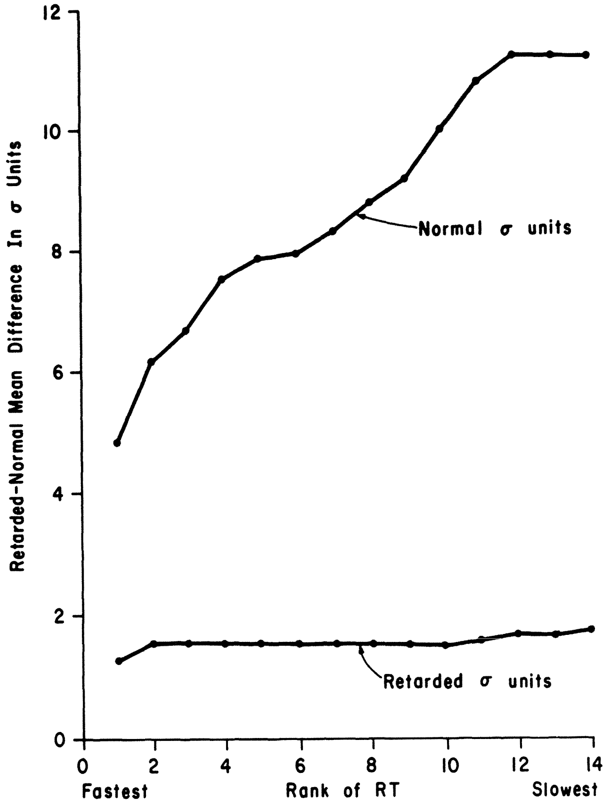

Figure 6. Difference in simple RT between retarded and normal Ss, expressed in both normal and retardate σ units, with RTs for 15 trials ranked from fastest to slowest.

paradigms reflect different processes, involving stimulus encoding, scanning of STM, and retrieval of semantic codes in LTM, all of which are probably involved in arriving at the correct answers to the relatively complex items used in ordinary intelligence tests, we should expect that an optimally weighted combination of RT measurements derived from all three paradigms should show a much more substantial correlation with mental test scores than measurements derived from any one RT paradigm. This is exaclty what Keating and Bobbitt (1978) found. Three RT-derived measures were obtained on each S: (1) choice RT minus simple RT (Hick paradigm), (2) semantic minus physical same/difference RT to letter pairs (Posner paradigm), and (3) slope of RT on set size with sets of 1, 3, or 5 digits (Sternberg paradigm). The multiple R of these three measurements with Raven scores of 60 school children in grades 3, 7, and 11 was .59, .57, and .60, in the three grades, respectively. I imagine that still higher correlations would be obtained if intraindividual variability were taken into account and if the correlations were corrected for attenuation using the between days test-retest stability coefficients. The average intercorrelation among the three paradigm measures was only .27, indicating that they are tapping different processes as well as sharing some variance in common.

The burning question is this: Will it be possible to discover a small number of such basic processes, measurable by means of RT, that will yield parameters which, in an optimally weighted combination, will "account for" practically all of the true *g* variance in psychometric tests of mental ability? Might not differently weighted combinations of a few process measurements based on RT also account for the variance in the so-called group factors involved in verbal, quantitative, and spatial abilities? This is what we must try to find out. Whatever the outcome may be, the effort will be amply rewarded by the gain in our theoretical understanding of the nature of mental abilities, to say nothing of the potential for practical applications should it turn out that most of the variance in complex mental abilities now measured by psychometric tests can be accounted for in terms of a number of RT parameters in a few fundamental paradigms.

References

Baumeister, A. A., and Kellas, G. Reaction time and mental retardation. In N. R. Ellis (Ed.), *International Review of Research in Mental Retardation*, Vol. 3. New York: Academic Press, 1968. (a)

Baumeister, A. A., and Kellas, G. Distribution of reaction times of retardates and normals. *American Journal of Mental Deficiency*, 1968, *72*, 715-718. (b)

Baumeister, A. A., and Kellas, G. Intrasubject response variability in relation to intelligence. *Journal of Abnormal Psychology*, 1968, *73*, 421-423. (c)

Berkson, G., and Baumeister, A. A. Reaction time variability of mental defectives and normals. *American Journal of Mental Deficiency*, 1967, *72*, 262-266.

Chiang, A. and Atkinson, R. C. Individual differences and interrelationships among a select set of cognitive skills. *Memory and Cognition*, 1976, *4*, 661-672.

Dugas, J., and Kellas, G. Encoding and retrieval processes in normal children and retarded adolescents. *Journal of Experimental Child Psychology*, 1974, *17*, 177-185.

Hick, W. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 1952, *4*, 11-46.

Hunt, E. *Varieties of Cognitive Power*. In L. B. Resnick (Ed.) *The Nature of Intelligence*. Hillsdale, N.J.: Erlbaum, 1976. P. 237-259.

Jensen, A. R. g: Outmoded theory or unconquered frontier? Creative Science and Technology, 1979, 2, 16-29.

Jensen, A. R., and Munro, E. Reaction time, movement time, and intelligence. Intelligence, 1979, 3, 121-126.

Keating, D. P., and Bobbitt, B. Individual and developmental differences in cognitive processing components of mental ability. Child Development, 1978, 49, 155-169.

McCauley, C., and Dugas, J., and Kellas, G., and DeVellis, R. F. Effects of serial rehearsal training on memory search. Journal of Educational Psychology, 1976, 68, 474-481.

Oswald, W. D. Über Zusammenhange zwischen Informationgescwindigkeit, Alter und Intelligenzstruktur beim Kartensortieren. Psychologische Rundschau, 1971, 22, 197-202.

Posner, M. I. Abstraction and the process of recognition. In G. H. Bower and J. T. Spence (Eds.) The Psychology of Learning and Motivation (Vol. 3). New York: Academic Press, 1969, 43-100.

Snow, R. E., Marshalek, B., and Lohman, D. F. Correlation of selected cognitive abilities and cognitive processing parameters: An explanatory study. Technical Report No. 3., Aptitude Research Project, School of Education, Stanford University, December, 1976.

Sternberg, S. High speed scanning in human memory. Science, 1966, 153, 652-654.

Vernon, P. A. Reaction time and intelligence in the mentally retarded. Unpublished paper, 1979.

Wade, M. G., Newell, K. M., and Wallace, S. A. Decision time and movement time as a function of response complexity in retarded persons. American Journal of Mental Deficiency, 1978, 83, 135-144.