

Relationships between measures of intelligence and choice reaction time

GLEN SMITH and GORDON STANLEY
University of Melbourne, Parkville, Victoria 3052, Australia

This paper draws attention to some of the problems associated with studies of reaction time and intelligence. The experiment reported here used choice reaction time without a home key and was concerned with the relationship between rate of reduction of uncertainty (RRU) and a verbal measure of intelligence (ML) developed for Australian adolescents. Results on a sample of 45 adolescent girls produced a correlation of .20 between RRU and ML. That this measure of intelligence, which has no primary spatial factor, seems to correlate less with RRU than do progressive matrices (SPM) lends support to the view that a common factor other than general intelligence contributes to the correlation between SPM and reaction time.

There is common-sense appeal in the idea that a person's rate of information processing may be one basic aspect of intelligence. A person who can handle an increasing amount of information with a smaller increase in response time is arguably more able to cope in most situations and is more intelligent than one who is greatly slowed by every extra bit of information confronting him. This is not to say that one who responds quickly per se is more intelligent than a slow responder, as a quick response may come from too little thought on the problem. In this paper, we are concerned more with the extra time needed to make a more complex choice compared with a simple one.

A useful measure of the rate of information processing must therefore involve both speed and accuracy. Such a measure, rate of reduction of uncertainty (RRU), is available and has long been used in reaction time paradigms in psychology. It comes from the mathematical definition of information or uncertainty presented by Shannon and Weaver (1949).

When a person is asked to make a different response to each of a set of stimuli, it is known that his reaction time (RT) is linearly proportional to his reduction of uncertainty (Hick, 1952; Hyman, 1953). That is, when RT is plotted against the reduction in uncertainty, it produces a straight line. This has been found repeatedly on a wide range of RT tasks (Welford, 1976). Roth (1964) investigated the relationship between intelligence and the slope of the linear increase in RT with increase in stimulus uncertainty in a choice RT task. His subjects faced a semicircular row of eight lights, each with a response key in front of it. They were told to rest one forefinger on a ninth key, the "home key," equidistant from the others. As soon as one of the lights came on, they were to press the key adjacent to the light as quickly as possible. The uncertainty was governed by the

number of alternative stimuli. Each subject had blocks of trials on various information loads, from one light or "simple reaction" with 0 bits to eight equiprobable lights with 3 bits. The RT was the time from the stimulus onset to the corresponding keypress. Roth found that the simple RT did not correlate with intelligence as measured by the Amthuser-Intelligenz-Struktur test. However, the angle between the linear RT function and the uncertainty axis, which is a measure of RRU, yielded a Pearson product-moment correlation of .39.

More recently, Jensen and colleagues (Jensen & Munro, 1979; Jensen, Note 1) have reported more experimental evidence linking RRU to intelligence. They used RT apparatus similar to that used by Roth (1964): an arc of eight lights and buttons centered on a home key. Intelligence was measured using Raven's Standard Progressive Matrices (SPM), which is a nonverbal test with a high loading on *g*, the general ability factor, and significant loading on *k*, spatial ability. Subjects were one college class of 39 girls about 15 years old. Using choice RT apparatus with a separate home key as they did, it is possible to split the RT measure used by Roth into two components: the time from the stimulus onset to the subject's lifting from the home key (RT) and the time from leaving the home key to pressing the response key (movement time, MT). Jensen and Munro measured RT and MT separately and found that RT increased linearly with number of bits of stimulus uncertainty, whereas MT varied less and was less related to uncertainty. The correlation between the slope of the regression of RT on uncertainty (their measure of RRU) and the Raven's SPM score was $r = -.30$ ($p = .06$). Correcting for attenuation gave a correlation of $-.36$ ($p < .05$). They also found that MT correlated with SPM, with $r = -.43$ ($p < .01$), whereas MT and RT correlated $r = .37$, which they interpreted as suggesting that RT and MT arise from different aspects of intelligence. Including other RT information such as the intercept of the regression and the standard deviation of the reaction

This study was supported by ARGC Grant A7915372I to the authors.

times together with MT measures, giving 10 measures in all, it was possible to produce a multiple correlation of about .50 with SPM. This compares favorably with correlations in the range of .35-.50 between SPM and other standard intelligence tests.

Jensen (Note 1) has reported considerable further research with the same general choice RT paradigm, measuring RT and MT parameters on a wide range of abilities, on several groups totaling about 800 subjects, from university students to retarded groups with mean IQs of 39. Between groups, the differences on a number of RT and MT measures predicted from their IQs were found to be significant. Within every group the correlations were positive and significant, mostly between .2 and .5. Jensen concluded that such RT-derived measures, when combined in a multiple-regression equation, predict some 50% or more of the variance in IQ or *g*.

There are two disadvantages in the paradigm used by Rothe (1964) and by Jensen and Munro (1979): their disregard of incorrect responses and the confounding of RT and MT. First, the uncertainty of the stimulus, and not that processed, is a valid measure only when performance is near perfect or comparable for all subjects. Although the stimulus situation may involve, say 3 bits of uncertainty, the full amount is processed only if performance is 100% correct. At the other extreme, 0 bits are processed if the responses are random. Within an individual, responses will be faster the more errors he makes, in a speed-accuracy tradeoff. Dividing the increase in RT by the increase in stimulus uncertainty will therefore be a misleading estimate of RRU and will not allow the comparison of subjects with different error rates, as we would want for a general measure. University students usually have low error rates on average, but the mildly mentally retarded have higher rates (Brewer & Nettelbeck, 1979). This can be overcome by using a measure of uncertainty reduction that makes allowances for errors (Hyman, 1953). Error-free performance also introduces problems: How can we be sure that subjects are being as fast as they can? Perhaps if they went faster, they would still be error free and have a higher RRU. This may be a "nonissue," in that the possibly underestimated RRUs may in fact be a better correlate of intelligence, but the point should be kept in mind.

Second, the separation of MT and RT is not clear-cut, except operationally. Intelligent and motivated subjects can follow the instructions well enough to leave their finger on the home key until they have decided which response to make, and only then move to make the response, allowing an accurate measure of RT and MT. Less able subjects have a tendency to lift their finger as soon as a stimulus light appears, then decide and make their response, which does not allow RT and MT to be measured properly. There is some suggestion of this effect in Jensen's (Note 1) report. He notes that the ratio of mean simple RT to mean MT is related to the average intelligence level of the group, with university students having much slower RTs than MTs, but with a

group of mean IQ of 39 having slower MTs than RTs, and other groups fitting almost linearly in between. If MT includes some processing time for the lower intelligence groups, their MTs would be inflated, producing the results noted. On the other hand, a large MT may arise just from poor perceptual motor skills, such as shown in a ballistic finger move between keys or more purely in a Fitts (1954) tapping task, which may be worth investigating in its own right.

Some doubt remains as to what precisely the MT measure involves. Clearly, it would be better to use a slight modification of the apparatus and procedure that permitted a clearer estimate of RT. Another common choice RT task enables this to be done. For this, the response keys are set out so that the subject can rest one finger directly on each key, and the subject presses the finger and key corresponding to the stimulus, with no extraneous movement. Welford (1971) has shown that differences between finger abilities in response speed have little effect with such a layout.

The simplest way of showing the validity of RRU as a measure of intelligence is by correlating it with a good measure of *g*. Raven's SPM is such a measure, but it may have one drawback. It has loadings on a spatial ability factor, *k*, and it seems likely that so does choice RT to stimuli defined by spatial position. Correlations between SPM and RRU will reflect all their common dimensions, in particular both *g* and *k*. Inferring the degree of relationship between *g* and RRU from the SPM-RRU correlation alone may give an overestimate.

The experiment reported here used choice RT without a home key and was concerned with the relationship between RRU and a verbal measure of *g* developed for and validated on an Australian sample.

METHOD

Subjects

The subjects were 45 girls at a metropolitan private school, ranging in age from 14 years 2 months to 16 years 3 months (average: 15 years 2 months).

Apparatus

The stimulus display, response, and RTs were controlled and measured by an LSI-11 computer. The display was presented on a vide screen (Intercolor 8001) and consisted initially of the resting display, a central vertical line down the screen, with a linear array of eight dashes, four on either side of the line. The length of the dashes was 5 mm, and their intercenter distance was 20 mm. Subjects viewed the screen from a distance of 50 cm, so the full display subtended a visual angle of 4.2 deg. The subject rested the four fingers (excluding thumbs) of each hand on eight keys arranged in two arcs of four, which the subject could span comfortably. Depressing a key activated a microswitch detectable by the computer. At the onset of each trial, one of the dashes was changed into a solid rectangle, 1 cm high by 5 mm wide, giving the stimulus display.

Procedure

Each subject was instructed to press the key that corresponded in position to the given stimulus as soon as it appeared. The stimulus stayed on until a response, correct or not, was made, and then the screen reverted to the resting display. The RT was

the time between the onset of the stimulus and the first key-press. The computer recorded the stimulus given, the response made, and the RT for each trial. The next stimulus was given 1.5 sec after the previous response key was released. Each subject was given 10 eight-choice practice trials, followed by 130 trials of two, four, and eight choices. The subset of stimuli used in the two- (and four-) choice conditions was central two (and four). The order of presentation was counterbalanced across subjects in a Latin square design. Within the two-, four-, and eight-choice conditions, the sequence of stimuli was random.

Intelligence was measured by the ML section of the ACER ML-MQ (revised) test. The ML is a 34-item multiple-choice questionnaire with a 15-min time limit, loading .6 on *g*, with a small linguistic loading. It correlates .65 with SPM scores in the 15-year-old group for which it was designed.

Although it is designed as a group test, it was administered individually in this instance by an interactive computer program. The computer displayed each question and its alternative answers and waited for a keyboard response specifying the subject's choice. When the subject had answered or decided to omit a question, another buttonpress brought up the next question. After one pass through the questions, any omitted questions were displayed again. At that stage, the subjects had the option either of stopping or of going through the questionnaire again to check their answers, as many times as they wished or until the 15-min limit was reached. The computer recorded the response made, the time between the fully displayed stimulus and the response (called the response time, to distinguish it from RT), the number of passes through the questionnaire, and the total time on the test.

RESULTS

The reduction in uncertainty (H_t) and the means and standard deviations of RT were calculated for each individual and degree of choice. For all subjects, the linear component of the regression of reduction in uncertainty on mean RT was significant at the .05 level, and the residual was significant only for 5 of the 45. Error rates ranged from 0% to 9.4% (median = 1.5%). The ML scores ranged from 7 to 29 out of 34, with a mean of 18.4 and standard deviation of 5.9. The norms for the ML are a mean of 14.9 and standard deviation of 6.4, so this sample was of above average ability.

The correlation between the RRU and ML scores was .20 ($p > .18$), which was not significant in this sample and is at the low end of the range of correlations Jensen (Note 1) reported for SPM and RT measures. That this intelligence measure, which has no spatial factor, seems to correlate less with RRU than SPM does lends support to the suggestion that a common factor other than *g* contributes to the correlation between SPM and RT measures.

No other RT measure correlated significantly with the ML score. The measures tried included the intercept from the regression and sometimes thought of as analogous to simple RT ($r = .05$), total number of errors ($r = -.11$), mean or standard deviation of RT for each degree of choice (all less than $r = .15$), and total RT, the sum of the two-, four-, and eight-choice mean RTs ($r = .11$).

DISCUSSION

The failure to find a significant positive relationship between intelligence test performance and RT may be due to a number of factors in combination or isolation. First, the positive correlation obtained by Jensen (Note 1) between his RT measure and SPM may have been due to a common loading on a specific spatial/performance ability factor. Second, it is possible that the ability involved in the less adequate measure of RRU used by Jensen reflects intelligence better than the measure of RRU used here.

To obtain some data relevant to this question, choice RT and SPM data were obtained from a group of 16 learning-disabled children attending a summer school. Their ages ranged from 8 years 1 month to 12 years 8 months. The SPM scores were converted to age-norm referenced scores. These ranged from 85 to 135+, with mean and median = 111 and standard deviation = 13.2. These scores were correlated with the RRU scores calculated as before. The Pearson product-moment correlation was .41, twice that between ML and RRU and consistent with the results reported by Jensen (Note 1). At least the low correlation obtained in the main experiment cannot be attributed to the absence of a home key and attendant procedural changes, so further studies can be carried out using the procedure advocated above as giving a cleaner measure of RT, untrammelled by MT. In view of the small sample, this new correlation was not significant ($p > .08$), but the results are consistent with the idea that correlations between SPM and RRU may overestimate the loading of RRU measures on *g*. Certainly, a larger factor-analytic study administering a battery of intelligence tests together with a choice RT task to measure RRU seems warranted before RRU can be fully accepted as an adequate intelligence measure.

REFERENCE NOTE

1. Jensen, A. R. *Reaction time and intelligence*. Paper presented at the NATO Conference on Intelligence and Learning, York University, York, England, July 19, 1979.

REFERENCES

- BREWER, N., & NETTELBECK, T. Discrimination, translation, or response organization: A clarification of factors underlying slower responding among mentally retarded persons. *American Journal of Mental Deficiency*, 1979, **84**, 195-199.
- FITTS, P. M. The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 1954, **47**, 381-391.
- HICK, W. E. On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, 1952, **4**, 11-26.
- HYMAN, R. Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, 1953, **45**, 188-196.
- JENSEN, A. R., & MUNRO, E. Reaction time, movement time, and intelligence. *Intelligence*, 1979, **3**, 121-126.
- ROTH, E. Die Geschwindigkeit der Verarbeitung von Information und ihr Zusammenhang mit Intelligenz. *Zeitschrift für Experimentelle und Angewandte Psychologie*, 1964, **11**, 616-622.
- SHANNON, C. E., & WEAVER, W. *The mathematical theory of communication*. Urbana: University of Illinois Press, 1949.
- WELFORD, A. T. What is the basis of choice reaction time? *Ergonomics*, 1971, **14**, 679-693.
- WELFORD, A. T. *Skilled performance: Perceptual and motor skills*. Glenview, Ill: Scott, Foresman, 1976.

(Received for publication April 15, 1980.)