Person. individ. Diff. Vol. 9, No. 2, pp. 243–255, 1988
Printed in Great Britain

0191-8869/88 $3.00 + 0.00
Pergamon Press plc

PSYCHOMETRIC *g* AND MENTAL PROCESSING SPEED ON A SEMANTIC VERIFICATION TEST

ARTHUR R. JENSEN, GERALD E. LARSON[2] and STEVEN M. PAUL[2]
School of Education, University of California, Berkeley, CA 94720, [1]Personnel Systems Department, Navy Personnel Research and Development Center, San Diego, CA 92152 and [2]School of Nursing, University of California, San Francisco, CA 94143, U.S.A.

(*Received 27 April 1987*)

Summary—The Semantic Verification Test (SVT) consists of brief, simple statements, which are either true or false, about various arrangements of the letters ABC (e.g. B after A). Mean reaction time (RT) for confirming or disconfirming the various statements varies according to their complexity. In independent studies of university students and Navy recruits, RT and other response latency parameters (intra-individual variability and movement time) from SVT performance show significant correlations of about −0.40 with nonspeeded tests of psychometric *g*. The mean RTs of adults to the various SVT item types are highly related to the mean error rates on these item types when the SVT is taken by elementary school children as a nonspeeded paper-and-pencil test. RT is correlated with the general cognitive ability factor (*g*) and not with the test-taking speed factor that is found in speeded paper-and-pencil tests. The degree of correlation between RT and psychometric *g* does not show any regular relationship to differences in the SVT item-type's complexity or difficulty as indicated by mean RT.

INTRODUCTION

Sentence verification tasks of various types have become a well known paradigm for the study of information processing both in experimental cognitive psychology and in the chronometric study of individual differences. The task essentially consists of a statement, or declarative sentence, followed by a pictorial representation which either conforms or does not conform to the prior statement. The now classic example, originally used by Clark and Chase (1972) to study the effects of various linguistic transformations on the speed of information processing, is of the form, "Star is above plus," or "Plus is not above star," followed by either $\overset{*}{+}$ or $\underset{*}{+}$, to which the subject must respond either "true" or "false" (or "yes" or "no"), depending on whether the picture confirms or disconfirms the statement. Differences in response latencies as a function of differences in sentence structure have been explained in terms of a quantitative scale of the sentence's linguistic complexity (Carpenter and Just, 1975).

The use of the sentence verification paradigm for the study of individual differences was introduced by Baddeley (1968), in the form of a brief speeded paper-and-pencil test that was found to be correlated 0.59 with scores of enlisted men on the British Army Verbal Intelligence Test. Since then, various sentence verification paradigms have figured in studies of individual differences in speed of information processing, and have shown mostly moderate-sized correlations with nonspeeded psychometric tests of mental abilities (Hunt and MacLeod, 1978; Jenkinson, 1983; Macleod, Hunt and Mathews, 1978; Paul, 1984; Tversky, 1975; Vernon and Kantor, 1986; Vernon, Nador and Kantor, 1985).

The first version of this paradigm that originated in this laboratory was an attempt to make the statements as simple as possible and to have them depend only on highly overlearned verbal codes. Because the simplified statements do not consist of declarative sentences with a subject and predicate, the test is termed a *semantic verification test* (SVT) instead of "sentence verification". It consists of all possible permutations of the statements "A before B," "A after B," "A not before B," "A not after B," with each statement followed by either AB or BA, making for a total of 16 items. The entire set is presented (in random order) 2 or 3 times, depending on the desired reliability and the available testing time. A 44-item version of this test, used in a study by Vernon *et al.* (1985),

The views in this article are those of the authors, are not official, and do not necessarily reflect the views of the Navy Department.

had an odd–even split-half reliability for subjects' overall mean RT of 0.97 in a group of college students. In the same study and in another study (Vernon and Kantor, 1986), reaction times of college students to the AB form of the SVT showed correlations with Full Scale IQ (Multidimensional Aptitude Battery administered under both timed and untimed conditions) ranging from −0.36 to −0.48, with a mean *r* of −0.40.

In hopes of possibly increasing the correlation between latency parameters of the SVT and psychometric *g*, a more complex form of the SVT was devised. It made use of statements about various possible permutations of the order of the three letters ABC, which permits a greater variety of statement types (e.g. "B between A and C") while maintaining the simplicity of highly overlearned verbal knowledge required of the task). (The ABC form of the SVT is fully described in the Method section). The greater number of item types that differ in difficulty (as indicated by mean RTs) was also intended to permit a more adequate test of the hypothesis that the correlation of RT with *g* is an increasing function of the degree of cognitive complexity (and hence mean response latency) of the RT task. Jenkinson (1983) found Pearsonian correlations between RTs on a Clark and Chase (1972) type of sentence verification test and the nonspeeded Raven Standard Progressive Matrices and Mill Hill Vocabulary tests (in a group of 6th-grade children) ranging between −0.28 and −0.34 (with a mean *r* of −0.32) for four degrees of sentence complexity, but there was no regular relationship whatsoever between the size of the correlations and sentence complexity as independently scaled in terms of the linguistic model of Carpenter and Just (1975).

The ABC form of the SVT was first used in Jensen's laboratory in a doctoral dissertation by Steven Paul (1984). Some of the results of Paul's study are reported here, but it should be pointed out that his research used a number of other RT paradigms in addition to the SVT and includes many other analyses addressed to other issues that are not reported in the present article. Exactly the same SVT test that was used by Paul with university students was also given (by means of a slightly different display apparatus) to U.S. Navy recruits under the direction of Gerald Larson. The results of these two studies of the SVT are presented here.

METHOD

Subjects

Two distinct studies are reported, one based on 50 students in the University of California. Berkeley (Paul, 1984), the other based on 105 Navy recruits at the U.S. Naval Training Station in San Diego, California. Members of the two groups were approximately the same age (20 ± 3 years). The 50 university students (group U) were paid volunteers obtained through an advertisement in the campus newspaper. The 105 Navy recruits (group N) were tested while awaiting interviews regarding assignment to specialized training programs. Informed consent was obtained from all subjects.

The university students were given the Sentence Verification Test [SVT] (as well as two other reaction-time tasks) twice, on separate days, to determine the SVT's test–retest reliability. Only the means of the first SVT test are reported in this article, however, to permit comparisons of group U with group N, which could be given the SVT on only one occasion. All correlations reported for group U are based on the averaged data from the two administrations of the SVT, in order to have the advantage of the higher reliability afforded by the composite data.

Psychometric tests

Group U was given the Raven Advanced Progressive Matrices (APM), a high-level nonverbal test of reasoning. Subjects took the test individually, with instructions to take all the time they needed to attempt to solve every one of the 36 multiple-choice items. As an incentive to take their time, subjects were paid according to the amount of time they took. This procedure thus fully meets the criterion of a power test, in contrast to a speed test.

Group N was given the Armed Services Vocational Aptitude Battery (ASVAB), which consists of ten diverse subtests, as follows:

General Science	A 25-item test of knowledge of the physical (13 items) and biological (12 items) sciences—11 min

Arithmetic Reasoning	A 30-item test of ability to solve arithmetic word problems—36 min
Word Knowledge	A 35-item test of knowledge of vocabulary, using words embedded in sentences (11 items) and synonyms (24 items)—11 min
Paragraph Comprehension	A 15-min test of reading comprehension—13 min
Numerical Operations	A 50-item speeded test of ability to add, subtract, multiply, and divide one- and two-digit numbers—3 min
Coding Speed	An 84-item speeded test of ability to recognize numbers associated with words from a table—7 min
Auto and Shop Information	A 25-item test of knowledge of automobiles, shop practices, and use of tools—11 min
Mathematics Knowledge	A 25-item test of knowledge of algebra, geometry, fractions, decimals, and exponents—24 min
Mechanical Comprehension	A 25-item test of knowledge of mechanical and physical principles—19 min
Electronics Information	A 20-item test of knowledge of electronics, radio, and electrical principles and information—9 min

The ASVAB was given as a group test in one session lasting approx $2\frac{1}{2}$ hr.

Semantic Verification Test

The Semantic Verification Test (SVT) consists of 7 positively worded and 7 negatively worded types of statements about various permutations of the order of just the letters ABC. Each statement is followed by a permutation of ABC; the subject's task is to indicate whether the order of the letters does ('Yes') or does not ('No') agree with the prior statement. Because the statements are not really sentences with a subject and predicate, but do convey a definite meaning, they are termed a *semantic* (rather than *sentence*) verification test. The statement types (termed SVT *conditions*) are as follows:

Condition	*Positive*	*Negative*
1	__ first	__ not first
2	__ last	__ not last
3	__ between __ & __	__ not between __ & __
4	__ before __ & __	__ not before __ & __
5	__ after __ & __	__ not after __ & __
6	__ before __	__ not before __
7	__ after __	__ not after __

The capital letters A, B, or C appear in the blank spaces. There are 6 possible permutations of each of the 14 statements; hence the total test consists of $6 \times 14 = 84$ statements. The permutation of ABC following each statement is in agreement with it in half the cases and in disagreement the other half. The order of presentation of the statement types and of agreement/disagreement (i.e. "Yes" or "No" responses) are completely randomized; the random order is programmed for computer administration and is the same for all subjects.

Just before giving the SVT as a reaction-time test, all subjects were given verbal instructions of the task requirements, followed by an untimed paper-and-pencil version of the SVT, which consisted of 28 statements, including two examples of each of the 14 statement types, followed by permutations of ABC, of which half were "true" and half were "false". This pretest was immediately scored by the examiner before going on to the reaction-time form of the SVT, and any incorrect response was pointed out to the subject. All the university students and the vast majority of the Navy recruits demonstrated error-free performance on the paper-and-pencil SVT.

The reaction-time form of the SVT consisted of 6 practice trials and 84 test trials. The entire sequence of stimulus presentation and the recording of the subject's reaction time (RT) and movement time (MT) were under the control of a microcomputer. However, the 6 practice trials and 84 test trials were subject-paced. The subject's response console consisted of three microswitch pushbuttons, each 2.5 cm in diameter, arranged in the form of an equilateral triangle (the apex pointing toward the subject) with 10 cm between the centers of the buttons. Directly above the top left and right pushbuttons are the labels YES and NO, respectively. The button at the lower apex of the "triangle" is known as the "home" button. The sequence of events is as follows: A trial begins when the subject presses the "home" button and holds it down (with the index finger of the preferred hand). After 2 sec, a single SVT statement appears on a stimulus display screen and remains in view for 3 sec, followed by a blank interval of 1–4 sec. The blank interval is continuous between 1 and 4 sec and its duration is randomized from trial-to-trial. Following the blank interval, the reaction stimulus (RS) appears, i.e. some permutation of ABC. The subject responds as quickly as he can by releasing the "home" button and pressing either the YES button (if the permutation of ABC agrees with the prior statement) or the NO button (if they disagree). Upon pressing the YES or NO button, the display screen goes blank; the next trial is initiated by the subject's depressing the "home" button.

On each trial, the computer recorded three things: reaction time, or RT (i.e. the interval, recorded in milliseconds, between the onset of the RS [e.g. ABC] and the subject's releasing the "home" button); movement time, or MT (i.e. the interval between release of the "home" button and pressing either the YES or NO button); and whether the response was correct or incorrect.

Five summary measures were obtained for each subject on each of the 14 SVT conditions and over all conditions: (1) median RT (the median RT on the 6 trials of a given SVT condition); (2) median MT; (3) intraindividual variability in RT, symbolized $RT\sigma_i$ (the standard deviation of the subject's RTs on the 6 trials of a given SVT condition); (4) intraindividual variability of MT, symbolized $MT\sigma_i$; and (5) number of incorrect responses, or errors.

Only one feature of the SVT procedure differed between groups U and N, namely, the stimulus display screen. In group U this was an alphanumeric 20-character display screen, 1.5 cm × 119 cm, with 0.7 cm red letters on a dark gray background; the display screen, tilted at a 45° angle, was located at the upper end of the 3-button response console, 8 cm above the response buttons. Unfortunately, for technical reasons that need not be explained here, this display screen could not be used with group N. Although the response console itself is the same for both groups, the display screen for group N was a computer monitor (IBM-PC color display), set directly behind the response console. The 0.5 cm letters appear white against a dark gray background. Although the great clarity of both types of stimulus display makes it unlikely that they would not yield highly comparable results, we do not know for certain. Therefore, no strong interpretation can be given to the mean absolute differences in the latency data between groups U and N. The main points of theoretical interest are the relationships and correlations that exist *within* each group.

RESULTS

Mean effects

The means of the Navy (N) and University (U) groups on the seven positive and seven negative statements of the SVT are shown for the median RT and MT and the intraindividual variability over trials ($RT\sigma_i$ and $MT\sigma_i$) in Figs 1 to 4. The relative sizes of the main effects displayed in these figures are shown in Table 1. An index of similarity between groups N and U in the shapes of the "profiles" of means over the seven SVT conditions is provided by the Pearson r, as shown in Table 2. Several features of these data warrant comment.

The high degree of similarity between the N and U groups' profiles, especially for median RT and MT, and to a somewhat lesser degree for $RT\sigma_i$, indicates a marked consistency in the relative difficulty of the different SVT conditions, both within and between the positive and negative types of statements. The RTs of the negative statements average some 200–250 msec longer than the corresponding positive statements; the MTs average some 40 msec longer for the negative than for the positive statements. Intraindividual variability (σ_i) in the negative conditions averages about 160 msec. longer than the positive conditions for $RT\sigma_i$ (in both groups N and U), and for $MT\sigma_i$ the negative conditions exceed the positive by about 246 msec and 88 msec in groups N and U, respectively.

RT is a much more sensitive indicator of differences in item difficulty than is MT. In marked contrast to the RT profile, the MT profiles are comparatively flat across the different SVT conditions. RT shows approx 100 times greater variance across all 14 SVT conditions than MT. Apparently, the largest part of the information processing is reflected by RT. $RT\sigma_i$ also reflects

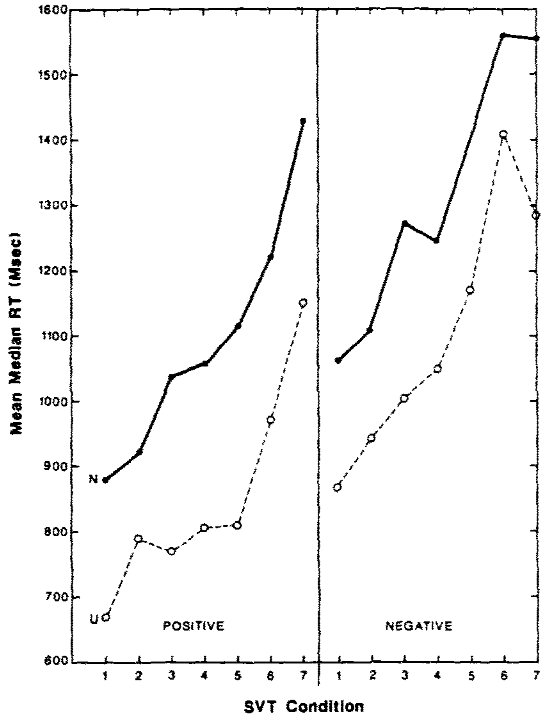

Fig. 1. Mean median RT (in msec) of Navy (N) recruits and University (U) students on the various conditions of the SVT.

differences in item difficulty (and *ipso facto* in information processing) to a considerable degree, especially in the negative conditions. $MT\sigma_i$ is peculiar, in that it not only shows by far the largest difference (1.8σ) between groups N and U, but is the only variable that shows a marked groups × conditions interaction. In group U, $MT\sigma_i$ scarcely reflects any differences in SVT conditions, while in group N the differences in $MT\sigma_i$ across conditions are highly correlated with item difficulty as reflected by median RT. This would seem to suggest that in group U information

Table 1. Effect size[a] for various contrasts and parameters of the SVT data shown in Figs 1 to 4

Contrast	Parameter			
	Med. RT	Med. MT	$RT\sigma_i$	$MT\sigma_i$
Groups (N vs U)	0.67σ	0.79σ	0.25σ	1.81σ
Condition (Pos. vs Neg.)	0.66σ	0.32σ	0.39σ	0.43σ
Groups × Conditions Interaction[b]	0.09σ	0.17σ	0.01σ	0.35σ

[a]Effect size is expressed in σ units, where σ is the average within-groups standard deviation.
[b]The interaction in this case is the mean difference between groups N − U in the negative (n) condition and N − U in the positive (p) condition, i.e. $(N - U)_n - (N - U)_p$.

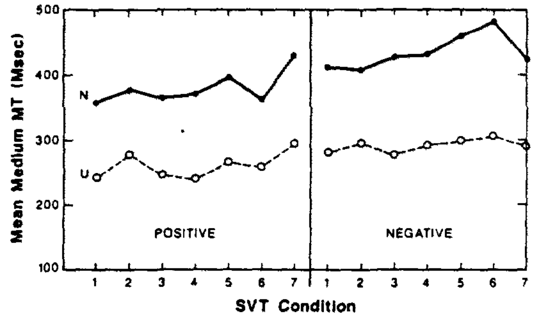

Fig. 2. Mean median MT (in msec) of N and U groups on the SVT.

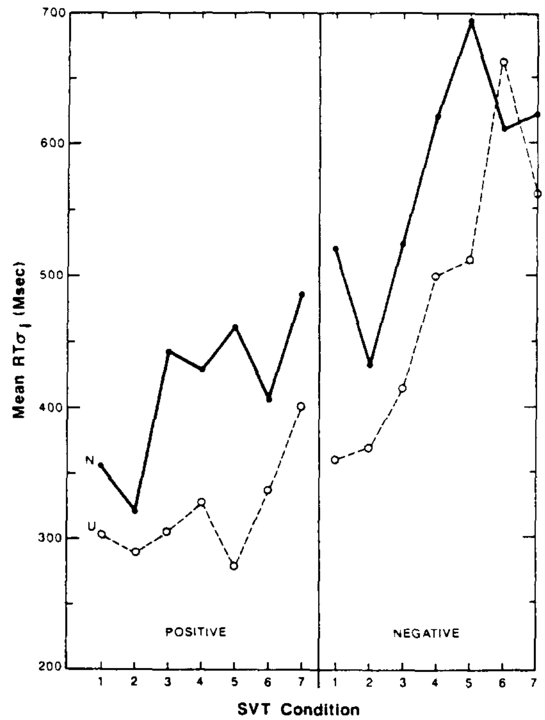

Fig. 3. Mean intraindividual variability ($RT\sigma_i$) of RT of the N and U groups on the SVT.

processing is confined almost entirely to the time interval measured by RT, while in group N some part of the information processing also occurs during the MT interval. This finding is suggestive of Sternberg's (1977) similar finding that in analogical reasoning tasks more intelligent subjects spend a relatively larger proportion of their total processing time in the encoding phase and relatively less time in the choice and response execution phase. In the present data, however, this is manifested mainly in $MT\sigma_i$ and only slightly in median MT, as can be seen by comparing Figs 2 and 4. It is noteworthy, and especially surprising in view of prior expectations, that, over all conditions, groups N and U differ more, in terms of σ units, on median MT (0.79σ) than on median

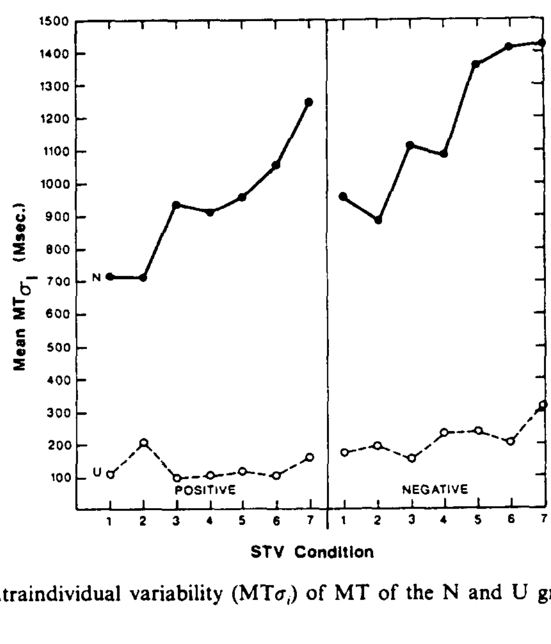

Fig. 4. Mean intraindividual variability ($MT\sigma_i$) of MT of the N and U groups on the SVT.

Table 2. Index of profile similarity (r) between Navy and University groups in Figs 1 to 4

	Profile similarity (r)[a]		
SVT Parameter	Positive	Negative	Both
1. Mean Median RT	+0.96	+0.97	+0.97
2. Mean Median MT	+0.85	+0.91	+0.89
3. $RT\sigma_i$	+0.53	+0.71	+0.83
4. $MT\sigma_i$	−0.13	+0.60	+0.32

[a]Positive and Negative profiles each based on group means for 7 SVT conditions: the Both profiles are based on all 14 (positive *and* negative) SVT conditions.

RT (0.67σ), and on $MT\sigma_i$ (1.81σ) than on $RT\sigma_i$ (0.25σ). It will be important to note how these variables are correlated with psychometric intelligence *within* each group.

SVT item difficulty and response latency

In attempting a theoretical explanation of the correlation between individual differences in RT on simple elementary cognitive tasks and scores on unspeeded complex tests of intelligence, Jensen (1982) hypothesized as follows:

> "The more complex the information and the operations required on it, the more time that is required, and consequently the greater the advantage of speediness in all the elemental processes involved. Loss of information due to overload interference and decay of [memory] traces that were inadequately encoded or rehearsed for storage or retrieval from LTM [long-term memory] results in "breakdown" and failure to grasp all the essential relationships among the elements of a complex problem needed for its solution. Speediness of information processing, therefore, should be increasingly related to success in dealing with cognitive tasks to the extent that their information load strains the individual's limited channel capacity. The most discriminating test items thus would be those that "threaten" the information processing system at the threshold of "breakdown." In a series of items of graded complexity, this "breakdown" would occur at different points for various individuals. If individual differences in the speed of the elemental components of information processing could be measured in tasks that are so simple as to rule out "breakdown" failure, . . . it should be possible to predict the individual differences in the point of "breakdown" for more complex tasks." (p. 122)

First, this 'breakdown' hypothesis predicts that there should be a relationship between mean error rates and mean median RTs of the various SVT conditions. The overall mean error rates for groups U and N are 7 and 11%, respectively. The rank-order correlation between the 14 different SVT conditions' error rates and mean median RTs are +0.82 and +0.80 in groups U and N, respectively. Error rates on the negative SVT conditions are almost 3 times greater than on the positive conditions. There is a quite close connection between error rates and all of the latency parameters of the SVT. These can be seen most clearly in the Navy data, with its larger N and higher (hence more reliable) error rates. The correlations between error rates and SVT conditions for group N are shown in Table 3. Because error rates and latencies in msec have such different scale properties, it is deemed advisable to report the rank-order correlations (ρ) in addition to the

Table 3. Pearson correlation (r)[a] and rank-order correlation (ρ) between mean error rates and mean response latency parameters of the various SVT conditions in the Navy group

	Positive		Negative		Both	
Parameter	r	ρ	r	ρ	r	ρ
Median RT	+0.66	+0.45	+0.70	+0.86	+0.76	+0.80
Median MT	+0.86	+0.70	+0.75	+0.82	+0.92	+0.74
$RT\sigma_i$	+0.86	+0.74	+0.89	+0.82	+0.89	+0.92
$MT\sigma_i$	+0.64	+0.36	+0.82	+0.82	+0.83	+0.79

[a]Correlations for positive and negative SVT conditions are each based on 7 pairs of means; for Both, the correlation is based on all 14 pairs of means.

Table 4. Pearson correlation (r)[a] and rank-order correlation (ρ) between mean error rates of children on the nonspeeded SVT and mean response latency parameters of the various SVT conditions in the University and Navy groups

SVT	University						Navy					
	Positive		Negative		Both		Positive		Negative		Both	
Parameter	r	ρ	r	ρ	r	ρ	r	ρ	r	ρ	r	ρ
Median RT	+0.55	+0.64	+0.48	+0.74	+0.69	+0.76	+0.69	+0.75	+0.51	+0.47	+0.72	+0.79
Median MT	+0.57	+0.46	+0.40	+0.36	+0.71	+0.71	+0.84	+0.64	+0.38	+0.59	+0.77	+0.75
RTσ_i	+0.28	+0.11	+0.57	+0.63	+0.69	+0.70	+0.83	+0.89	+0.83	+0.83	+0.89	+0.89
MTσ_i	+0.08	+0.03	+0.67	+0.79	+0.67	+0.63	+0.70	+0.79	+0.38	+0.61	+0.79	+0.81

[a]Correlations for Positive and Negative SVT conditions are each based on 7 pairs of means; for Both, the correlation is based on all 14 pairs of means.

Pearson correlations (r) in this case. It is evident in Table 3 that there is a high degree of correlation between the mean error rates on the various SVT conditions and the mean response latency data on the corresponding conditions. The correlations would probably be even larger if the error rates were higher and hence more reliable.

Second, and more importantly, the 'breakdown' hypothesis predicts that there should be a relationship between mean error rates and mean median RTs of the various SVT conditions even when error rates are based on *nonspeeded* performance on the SVT. When the SVT was given simply as a paper-and-pencil test without time limit, the task was so easy that there were absolutely no errors at all in the University group and an average of less than one error per subject in the Navy group. With such error-free performance on the nonspeeded form of the SVT, it would be impossible to test the 'breakdown' hypothesis. Therefore, the nonspeeded SVT was given to 77 children (ages 8–9 years). This paper-and-pencil test consisted of 28 items, which included two examples each of the 7 positive and 7 negative SVT conditions. Subjects were urged to attempt all the items and to take as much time as needed to mark the correct answers. Even under these lenient conditions, the children's mean error rates were 11.8 and 22.7% for the positive and negative SVT statements, respectively, with 17.24% errors overall. The percentage of erroneous responses on each SVT condition, based on all 77 subjects, was correlated with the mean response latency data for the corresponding SVT conditions in groups U and N. These correlations are shown in Table 4. These correlations are quite in accord with the hypothesis. The fact of generally higher correlations in group N is most probably attributable to its larger number of subjects, with consequently more reliable measures of the SVT parameters. All the correlations in Table 4 are, of course, somewhat attenuated by the less than perfect reliability of the profile of mean error rates based on the children's nonspeeded SVT data, the reliability of which, as determined by the Hoyt (1941) method, is only 0.90. Thus the mean response latencies in the adult groups afford a high degree of prediction of the mean error rates on the SVT items when these are taken as a nonspeeded test by young children.

Correlations of SVT with psychometric tests

University data. The criterion for psychometric g in group U is the raw score (number correct) on the 36-item Raven Advanced Progressive Matrices (APM) test administered individually without time limit. The mean score is 28.68, SD = 5.48. The internal consistency reliability, as determined by the KR-20 formula, is 0.86.

Overall correlations (r) of APM scores with SVT parameters were based on two types of SVT measures: (1) each subject's mean of the given raw parameter measures over all 14 SVT conditions, and (2) the parameter measures converted to standardized z scores within each SVT condition and summed over all 14 conditions. (The purpose of the second method was to give equal weights in terms of means and SDs to each of the SVT conditions represented in the composite score to be correlated with the APM.) The correlations based on these two types of SVT composite scores are shown in Table 5.

It is frequently hypothesized that the degree to which RT measures are correlated with IQ, or psychometric g, is directly related to the complexity of the RT task. To examine this hypothesis in the case of the SVT, the correlations of Raven scores with median RT and RTσ_i were plotted for the various SVT conditions, as shown in Fig. 5. Surprisingly, there appears to be no systematic relationship between SVT conditions and their correlations with the Raven APM. The expected

Table 5. Correlation (r) between Raven Advanced Progressive Matrices and SVT parameters in the University group ($N = 50$)

SVT Parameter	SVT Raw Scores[a]	z-Score Composite[b]
Median RT	−0.454**	−0.445**
Median MT	−0.166	−0.174
$RT\sigma_i$	−0.433**	−0.496**
$MT\sigma_i$	−0.304*	−0.230

*$P < 0.05$, one-tailed test.
**$P < 0.01$, one-tailed test.
[a]Raw score is the subject's mean over all 14 SVT conditions.
[b]z-score composite is the sum of the SVT scores after they were standardized within each of the 14 SVT conditions.

Table 6. Correlation (r) of ASVAB principal component scores with SVT parameters in Navy group ($N = 105$)

SVT Parameter	Unrotated: I General ability	Unrotated: II Test speed	Varimax Rotated: 1 Cognitive ability	Varimax Rotated: 2 Speediness
Median RT	−0.263*	−0.128	−0.256*	−0.141
Median MT	−0.069	−0.103	−0.064	−0.106
$RT\sigma_i$	−0.389*	−0.033	−0.387*	−0.052
$MT\sigma_i$	−0.333*	−0.052	−0.330*	−0.069
Errors	−0.233*	−0.045	−0.231*	−0.057

*$P < 0.01$, one-tailed test.

higher correlation for the relatively difficult *negative* SVT condition than for the less demanding *positive* condition does not appear, for either median RT or $RT\sigma_i$.

Navy data. Scores on the ten subtests of the Armed Services Vocational Aptitude Battery (ASVAB) were made more manageable by subjecting them to a principal components analysis. There are only two components with eigenvalues greater than 1. Similar analyses performed on the total national standardization sample of over 12,000 subjects also yield no more than two components (or two factors) in the ASVAB. The first unrotated principal component, which accounts for 42% of the total ASVAB variance, is a good representation of the general factor of the ASVAB. The second unrotated component, which accounts for 18% of the total variance, has high loadings only on the most speeded subtests (Numerical Operations [0.82] and Coding Speed [0.77]), and represents a speed-of-work factor in test taking. The varimax rotated components are quite similar to the unrotated components, the first rotated component being general cognitive ability, the second rotated component being a test-speed factor.

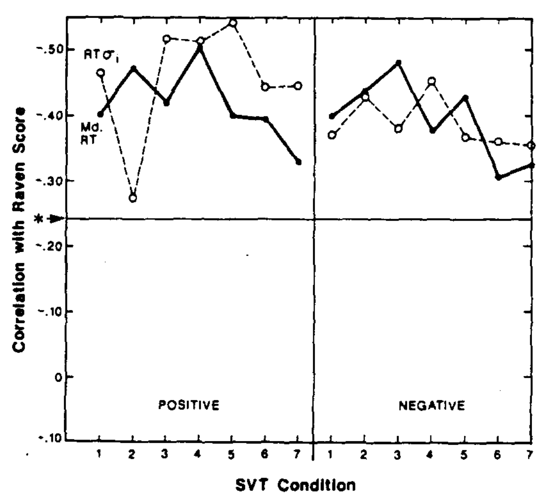

Fig. 5. Correlation of Raven Advanced Progressive Matrices with SVT median RT and $RT\sigma_i$ as a function of SVT condition, in University group. The *-arrow indicates the value of r that is significant at $P < 0.05$, one-tailed test.

Table 7. Stepwise multiple correlation (R) in predicting ASVAB rotated principal component scores from SVT parameters

Independent variables	Component 1[a] R	Component 1[a] p	Independent variables	Component 2[b] R	Component 2[b] p
1. $RT\sigma_i$	−0.387	0.0001	1. Median RT	−0.141	0.154
2. 1 + Errors	−0.418	0.0001	2. 1 + $MT\sigma_i$	−0.175	0.212
3. 2 + Median MT	−0.419	0.0003	3. 2 + Errors	−0.188	0.311
4. Median RT	−0.420	0.0007	4. 3 + Median MT	−0.193	0.438
5. 4 + $MT\sigma_i$	−0.421	0.0019	5. 4 + $RT\sigma_i$	−0.205	0.518

[a]General cognitive ability.
[b]Speediness in test-taking.

Table 8. Correlation between the profile of ASVAB subtests' g factor loadings (uncorrected and corrected for attenuation) and the profile of ASVAB subtests' correlations with SVT parameters (uncorrected and corrected for attenuation)

SVT Parameter	Profile correlation: Uncorrected	Profile correlation: Disattenuated[a]
Median RT	−0.508	−0.508
Median MT	+0.143	−0.134
$RT\sigma_i$	−0.846	−0.843
$MT\sigma_i$	−0.847	−0.845
Errors	−0.554	−0.532

[a]The correlations reported here are not themselves disattenuated, but are based on (a) disattenuated g loadings and (b) disattenuated correlations between ASVAB subtests and SVT parameters. The above correlations are r_{ab}.

Component scores (analogous to factor scores) were obtained and correlated with the SVT parameters, as shown in Table 6. The most noteworthy feature here is the absence of any significant correlation of the SVT variables with the speed factor of the ASVAB. This finding contradicts the common interpretation of the correlation between response latency measures on elementary cognitive tasks (ECTs) and psychometric test scores as being the result of a speed of test-taking factor that speeded psychometric tests (such as those that mark the speed factor in the ASVAB) have in common with the speed of performing ECTs. Clearly, the latency variables of the SVT are significantly correlated with the *cognitive* component of the ASVAB, but not at all with its *speediness* component.

To determine the maximum multiple correlation with which each of the rotated component scores of the ASVAB (as the dependent variable) could be predicted by an optimum combination of the SVT parameters (as the independent variables), multiple regression analysis was performed. It is summarized in Table 7. The purely cognitive component shows a highly significant R at every level of the optimal stepwise regression, while the test speediness component is not significantly predictable from the SVT variables. This constitutes a strong refutation of the test-speed explanation of the correlation between reaction time variables on simple tasks and scores on complex psychometric tests, and it further supports the arguments made by Vernon *et al.*, 1985 and Vernon and Kantor, 1986.

In fact, it appears that it is the general factor, g, of the ASVAB that mostly determines the magnitudes of the various subtests' correlations with the SVT. A principal axes factor analysis was performed on the ten ASVAB subtests to obtain what is probably the best possible estimate of Spearman's g from this battery, as the first unrotated principal factor. It turns out that the degree of correlation (r) of the ASVAB subtests with the SVT variables is directly related to the subtests' g loadings; in fact, no other factors or principal components, unrotated or rotated, shows as high a relationship to the ASVAB subtests × SVT correlations. Table 8 shows the correlations (r_{ab}) between (a) the profile of g factor loadings (both uncorrected and corrected for attenuation, based on the ASVAB subtest reliabilities in the national standardization sample [Bock and Mislevy, 1981]) and (b) the profile of correlations of each of the ASVAB subtests with the given SVT parameter (uncorrected and disattenuated). The highest correlations, of about −0.85, are found for the intraindividual variability in SVT latencies, both $RT\sigma_i$ and $MT\sigma_i$. Tables 5, 6, and 8 indicate that median MT has the last correction of all the SVT parameters with psychometric g.

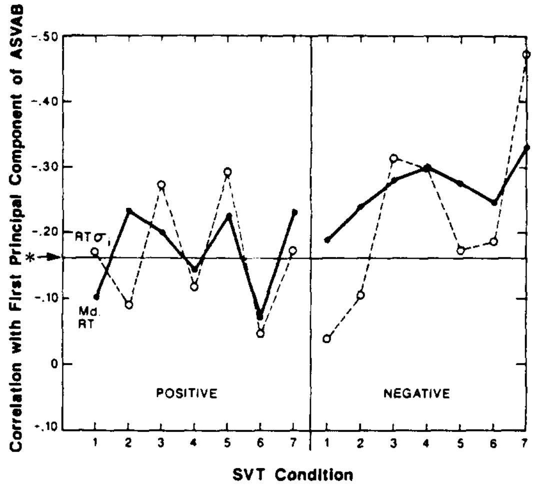

Fig. 6. Correlation of first principal component of ASVAB with median RT and $RT\sigma_i$ of the SVT, as a function of SVT condition, in the Navy group ($N = 105$). The *-arrow indicates the value of r that is significant at $P < 0.05$, one-tailed test.

Figure 6, based on the Navy data, is analogous to Fig. 5, based on the University data. Again, there seems to be no very regular relationship between SVT conditions and their correlation with psychometric *g*, here represented by the first principal component scores of the ASVAB. Unlike the University group, however, the Navy group shows somewhat consistently higher correlations of the negative than of the positive SVT conditions with psychometric *g*. Overall, however, the correlations are generally higher in the University group, for no clear apparent reason—possibly because of differences in reliability or differences between the Raven and the ASVAB, it is impossible to say from the present evidence.

Speed–accuracy trade-off

One rather common explanation of the correlation between RT and psychometric *g* is that the brighter subjects adopt a performance strategy on the RT task that maximizes speed of response by sacrificing accuracy or correctness of response, known as 'speed/accuracy trade-off'. If this explanation is valid, then two empirical consequences of a speed/accuracy trade-off should be evident: (1) There should be a *negative* correlation between individual differences in median RT and error rate, and (2) there should be a *positive* correlation between individual differences in psychometric *g* and error rate. As far as we have been able to determine, neither of these conditions has ever been reported in any kind of RT data. Moreover, they have been contradicted in every set of RT data ever collected in this laboratory—data on over 2000 subjects to date. The present study is no exception. In the University sample, the correlation between median RT and error rate on the SVT is +0.13, which is nonsignificant ($P = 0.35$) and of opposite sign to that predicted by the speed/accuracy trade-off hypothesis. In the Navy sample, the error rate is correlated −0.07 (n.s.) with median RT. Also, error rate is significantly correlated −0.23 ($P < 0.02$) with the general factor of the ASVAB, but the direction of the correlation is *opposite* to the prediction of the trade-off hypothesis. Hence, the trade-off hypothesis does not in the least account for the observed correlations between RT and unspeeded psychometric measures of intelligence, or *g*.

DISCUSSION AND CONCLUSIONS

These studies are a further demonstration of a significant and substantial correlation, of about −0.40 to −0.50, between response latency measures on a very simple task, the Semantic Verification Test (SVT), and scores on complex psychometric tests of intellectual ability (Raven Advanced Progressive Matrices and the Armed Services Vocational Aptitude Battery). The Raven was administered without time limit. The SVT is so simple that the mean response latency to the items is only about one second for young adults, and all such subjects (university students and Navy recruits) are capable of virtually error-free performance on the SVT when it is administered as a nonspeeded paper-and-pencil test. Hence the speed of processing very simple cognitive tasks, in which the informational content *per se* is so minimal as to not be a source of individual differences under nonspeeded conditions, is capable of predicting scores on unspeeded complex tests of reasoning and general information.

It was shown that this correlation is not attributable to a test-speed factor of the kind that contributes to individual differences in tests of clerical speed and accuracy. A speed factor in the ASVAB, with high loadings on the simple but highly speeded tests of Coding Speed and Numerical Operations, showed no significant correlation with the SVT latency measures. The latter were significantly correlated with only the general cognitive ability factor, or *g*, extracted from the ten ASVAB subtests. Moreover, the degree of correlation of the SVT RT and $RT\sigma_i$ measures with the ASVAB subtests is directly related to the subtests' loadings on the *g* factor. It appears that it is *g*, rather than any other factors, that is the basis of the correlation between scores on a variety of conventional psychometric tests and speed of information processing in elementary cognitive tasks.

The mean response latencies of adults to SVT items predict the item difficulties (percent errors) of 8- and 9-year old children when the SVT is given to them as a nonspeeded paper-and-pencil test. While adults have virtually zero error rates on these simple SVT items under nonspeeded conditions, whatever differences in the information processing demands of the items result in mean differences in their item latencies when the SVT is given as a reaction time test also cause parallel differences in error rates when the items are taken as a paper-and-pencil test by children for whom the items present some cognitive difficulty even under nonspeeded conditions. This finding supports the hypothesis that successful problem solution, even when all of the information necessary for attaining the correct solution is available in the subject's repertoire, is constrained by the interaction of the complexity of the processing required by the problem and the individual's speed of information processing. When the time required for solution exceeds the decay time of the input information (or the products of the mental processing operations effected on it) in the individual's working memory, the resulting 'breakdown' decreases the probability of successful performance. This disadvantage can be overcome, in the case of complex problems for which the solution time exceeds the decay time of the information in working memory, by the subject's adopting a complex strategy for solution, such as rehearsing information in short-term memory (STM) to get it into long-term memory (LTM), searching LTM for relevant information, dividing the problem into a number of subproblems for solution, making notes, and the like. The employment of such strategies is relatively test-specific, and individual variations in the adoption of strategies for the solution of complex problems probably diminishes the correlation between complex performance measures and reaction time measures of performance on elementary cognitive tasks, which are so simple as to offer relatively little scope for individual differences in the adoption of various strategies. Accordingly, from this viewpoint, we should expect that RT on elementary cognitive tasks would have smaller true-score correlations with single psychometric tests composed of highly homogeneous item types than with *g* factor scores derived from a number of diverse psychometric tests, because the test-specific variance in complex psychometric tests is minimized in their general factor.

REFERENCES

Baddeley A. D. (1968) A 3 min reasoning test based on grammatical transformations. *Psychonom. Sci.* **10**, 341–342.

Bock R. D. and Mislevy R. J. (1981) *The profile of American youth: Data quality analysis of the Armed Services Vocational Aptitude Battery.* National Opinion Research Center, University of Chicago.

Carpenter P. A. and Just M. A. (1975) Sentence comprehension: A psycholinguistic processing model of verification. *Psychol. Rev.* **82**, 45–73.

Clark H. H. and Chase W. G. (1972) On the process of comparing sentences against pictures. *Cogn. Psychol.* **3**, 472–517.

Hoyt C. (1941) Test reliability obtained by analysis of variance. *Psychometrika* **6**, 153–160.

Hunt E. and MacLeod C. (1978) The sentence-verification paradigm: A case study of two conflicting approaches to individual differences. *Intelligence* **2**, 129–144.

Jenkinson J. C. (1983) Is speed of information processing related to fluid or crystallized intelligence? *Intelligence* **7**, 91–106.

Jensen A. R. (1982) Reaction time and psychometric *g*. In (Ed. by Eysenck H. J.). *A Model for Intelligence*. Springer-Verlag, Heidelberg.

MacLeod C. M., Hunt E. and Mathews N. N. (1978) Individual differences in the verification of sentence-picture relationships. *J. Verb. Learn. Verb. Behav.* **17**, 493–507.

Paul S. M. (1984) Speed of information processing: the semantic verification test and general mental ability. Unpublished doctoral dissertation, University of California, Berkeley.

Sternberg R. J. (1977) *Intelligence, Information Processing and Analogical Reasoning: The Computational Analysis of Human Abilities.* Erlbaum, Hillsdale, N.J.

Sternberg R. J. (1986a) Haste makes waste versus a stitch in time? A reply to Vernon, Nador, and Kantor. *Intelligence* **10,** 265–270.

Sternberg R. J. (1986b) Time capsules are not a panacea: A reply to Vernon. *Intelligence* **10,** 277–279.

Tversky B. (1975) Pictorial encoding of sentences in sentence-picture comparison. *Q. Jl exp. Psychol.* **27,** 405–410.

Vernon P. A. (1986) He who doesn't believe in speed should beware of hasty judgments: A reply to Sternberg. *Intelligence* **10,** 271–275.

Vernon P. A. and Kantor L. (1986) Reaction time correlations with intelligence test scores obtained under either timed or untimed conditions. *Intelligence* **10,** 315–330.

Vernon P. A., Nador S. and Kantor L. (1985) Reaction times and speed-of-processing: Their relationship to timed and untimed measures of intelligence. *Intelligence* **9,** 357–374.