# The Suppressed Relationship Between IQ and the Reaction Time Slope Parameter of the Hick Function

ARTHUR R. JENSEN
*University of California, Berkeley*

The two parameters of the Hick paradigm, the intercept (*a*) and the slope (*b*), of reaction time (RT) as a function of the information load scaled in bits (i.e., the binary logarithm of the number of stimulus-response alternatives) differ in their (negative) correlation with IQ, *a* generally having a larger correlation than *b*. The typically low and often nonsignificant correlation between *b* and IQ appears to contradict the theory that rate of information processing is a component of general mental ability (*g*) as approximated by IQ or other highly g-loaded tests. The *a* of the Hick function largely reflects individual differences in the sensory-motor lags in task performance, while the *b* supposedly reflects individual differences solely in the rate of information processing. Hence *b* theoretically should be more highly correlated with *g* or its proxy, IQ, than is *a*. But in fact, the opposite is commonly found. The weakness of the $b \times$ IQ correlation, as compared with the correlation between IQ and *a* (and with other variables derived from the Hick paradigm) is mainly attributable to statistical artifacts that suppress the *b* parameter's correlation with any other variables, e.g., *a* and IQ. When the $b \times$ IQ correlation is estimated under conditions that reduce the statistical suppression of this relationship, the correlation is appreciably increased and is consistent with prediction from information processing theory.

The frequent failure of the slope parameter (*b*) of the Hick function to support its theoretically predicted relationship to IQ is a bum rap. Indeed, the predicted correlation is often low and, with the sample sizes typically used in RT studies, often statistically nonsignificant. (See Jensen, 1987, for a comprehensive review of research on the Hick paradigm.) This fact, however, is not a failure of the theoretical prediction, but rather the effect of inherent statistical artifacts that suppress the theoretically expected correlation. The problem of the slope parameter of RT in response to stimulus conditions that differ in information processing load, or complexity, applies not only to the Hick paradigm, but is relevant to most other elementary cognitive tasks, such as the S. Sternberg mem-

Direct all correspondence to: Arthur R. Jensen, School of Education, University of California, Berkeley, CA 94720-1670.

INTELLIGENCE 26(1): 43-52
ISSN: 0160-2896

ory scan paradigm and the Posner physical versus name identification paradigm. David Lohman (1994) first explicated these problems in an editorial in this journal, particularly with reference to studies of individual differences in spatial ability based on the speed of mental rotation of figures in the Shepard and Metzler paradigm. He also pointed out the difficulty the slope problem creates for cognitive theorists attempting a cognitive components analysis to explain individual differences in the complex abilities measured by conventional psychometric tests in terms of relatively elementary information processes in which individual differences are measured by chronometric techniques. In the context of analysis of variance, individual differences in overall mean RT, along with task difficulty, account for so much of the total variance when RT is measured at different levels of task difficulty (or complexity), as in the Hick paradigm, that comparatively little reliable variance is left over for the linear interaction of individuals × levels of task difficulty (e.g., the slope of RT as a function of BITs of information in the Hick paradigm).

Hick's law states that the RT to a given reaction stimulus that is selected from among $n$ alternative stimuli can be expressed by the regression of RT on the binary logarithm of $n$. (In information theory, the binary logarithm of the number of choices, i.e., $\log_2 n$, or BIT, is the standard unit of information.) Thus, for any given value of $n$, Hick's law states: RT $= a + b(\log_2 n)$, or RT $= a + b$(BITs), where $a$ is the intercept of the Hick function, $b$ is its slope (or regression coefficient), and BITs is $\log_2 n$. Hence for $n = 1, 2, 4, 8$ S-R alternatives, the corresponding BITs are 0, 1, 2, 3.

Hick's formulation provides a test of the theory that individual differences in the speed of information processing is a causal component of individual differences in intelligence as measured by nonspeeded psychometric tests, or IQ. The $a$ of the Hick function is claimed to represent mostly that part of the total RT that is required for the signal to be sensorily transduced and transmitted to the brain and for the efferent nerves to activate the muscles. The $b$ of the Hick function is claimed to represent that part of the total RT required by the decision process, which is a function of the amount of information (in BITs) that has to be processed. Therefore, one should expect $b$ to have a more substantial correlation with IQ than is reported in most studies.

The first study (Roth, 1964) of the relationship of the Hick parameter $b$ to IQ reported a significant correlation of −.39; that is, the participants with higher IQ showed a lesser increase in the time taken to process increasing amounts of information than did the participants with lower IQ. In this study, the correlation between $a$ and IQ was nonsignificant. But the correlation between $a$ and $b$ was −.41. But this negative correlation seemed theoretically inconsistent and paradoxical. Why should individual differences in speediness in the processing component of RT (reflected by $b$) be *negatively* correlated with individual differences in the sensory-motor component(reflected by $a$)? Many subsequent studies modeled after Roth's experiment (with some improved procedural modifications) have also generally found negative correlations between $a$ and $b$. But they have also usually found weaker correlations of IQ with $b$ than with $a$ or any other variables that can be derived from the Hick paradigm, such as the overall mean RT ($\bar{X}_{RT}$) and the intraindividual variability in RT across trials as (measured by the standard deviation of the participant's RT over a given number of trials), symbolized as $s_i$.

## THE CORRELATION BETWEEN INTERCEPT (A) AND SLOPE B, ($R_{AB}$)

In seven independent studies totalling 537 participants (university undergraduates), the average value of $r_{ab}$ was −.151; corrected for attenuation and restriction of IQ range, $r_{ab} = -.277$ (Jensen, 1987, p. 143). All other possible correlations between the Hick variables ($\bar{X}_{RT}$, $a$, $b$, $s_i$) except $r_{ab}$ were positive correlations and were much larger. Why the negative value of $r_{ab}$?

The answer lies in the concatenation of four facts: (1) $b$ has the lowest reliability of any of these variables; (2) $a$ and $b$ both share exactly the same errors of measurement, (3) even though $a$ and $b$ share the same measurement error, because the total variance of $b$ is much smaller than that of $a$, $b$ has much lower reliability than $a$, and (4) the shared errors of measurement go in opposite directions for $a$ and $b$, i.e., they are *negatively* correlated, to a degree depending on the number of elements (e.g., levels of task complexity) that enter into the regression of RT on BITs. The correlation between the errors of measurement for $a$ and $b$ is given by Marascuilo and Levin (1983, p. 161):

$$r_{ab} = \frac{-\bar{X}}{\sqrt{\frac{\Sigma X^2}{N}}},$$

Where $X$ is each of the values of the independent variable (in this case the number of BITs at each level of complexity, e.g., 0, 1, 2, 3 BITs, corresponding to $n$ = 1, 2, 4, and 8 S-R alternatives in the Hick task). Therefore, with these four levels of complexity that are used in all of my studies of the Hick paradigm, the errors of measurement in $a$ and $b$ are correlated −.80. Given the comparatively low reliability coefficient of $a$ and especially of $b$, the large (i.e., −.80) negative correlation between their errors of measurement entirely accounts for the near-zero and negative correlations between $a$ and $b$ typically found in Hick studies. As will be demonstrated later on, the negatively correlated errors of measurement in $a$ and $b$ causes each parameter to act as a *suppressor* variable with respect to the other parameter's correlation with any external variable, such as IQ.

### Reliability Coefficients of *a* and *b*

The Spearman-Brown boosted split-half reliability (odd-even trials) of $a$, based on 490 participants, is .95. The split-half reliability coefficient for $b$ is .81.

The more important measure of the stability of an individual differences variable is afforded by the test-retest reliability coefficient. The test-retest reliability (with 2 to 3 days retest interval) of $a$, based on 272 participants, is .72. The test-retest reliability of $b$ is .39.

The reliability coefficients for $a$ and $b$ can be compared with those for $\bar{X}_{RT}$: Split-half = .95; test-retest = .84.

It is also important to note that the variances both of $\bar{X}_{RT}$ and of $a$ are much greater (by at least a factor of 10) than the variance of $b$. Also, the typical correlation between $a$ and $\bar{X}_{RT}$ is about + 0.9 compared with the typical correlation between $b$ and $\bar{X}_{RT}$ of about +0.3 (Jensen, 1987, p. 143). In light of these conditions, therefore, it becomes apparent from the formulas for $a$ and $b$ why $a$ is by far the more reliable and robust variable in its correlation with IQ (or with any other variable external to the Hick data). Letting T stand for a participant's median RT at each of the 4 levels of task complexity (each level measured in BITs, abbreviated as B), $\sigma_T$ is the standard deviation of the median RTs across the 4 levels of complexity, and $\sigma_B$ is the standard deviation of the number of BITs at each level of complexity (in this case, BITs = 0,1,2,3), then:

*Intercept*: $a_T = \bar{X}_T - b_{TB}\bar{X}_B$
*Slope*: $b_{TB} = r_{TB}(\sigma_T/\sigma_B)$.

Note that $\sigma_B$ and $\bar{X}_B$ are constants, equal to 1.118 and 1.5, respectively, when there are four discrete levels of task complexity ranging from 0 to 3 BITs. Hence these constants contribute nothing to individual differences variance, and as we are not here concerned with absolute numerical values, the above formulas can be simplified by setting each of these constants equal to 1 without losing information essential to this demonstration; thus simplified:

$$a_T = \bar{X}_T - r_{TB}\sigma_T$$
$$b_{TB} = r_{TB}\sigma_T.$$

Also note that $r_{TB}$ is an index of the participant's degree of conformity to Hick's law; it is the standardized linear regression of RT on task complexity as scaled in BITs. The average value of $r_{TB}$ among university and vocational college students is +.93 ($SD$ = .11). So, for a given individual, on average, the value of $r_{TB}$ scarcely differs from the individual's $\sigma_T$. For participants who conform perfectly to Hick's law (i.e., $r_{TB} = 1$), the above formulas simplify to

$$a_T = \bar{X}_T - \sigma_T$$
$$b_{TB} = \sigma_T.$$

There are essentially only two individual differences parameters here, $\bar{X}_T$ and $\sigma_T$. The first, $\bar{X}_T$, which, in terms of the analysis of variance, constitutes the main effect of individual differences, empirically has a large variance and is highly reliable; the second is the participant × BITs interaction, and, typically of the interaction terms in analysis of variance, it constitutes a minor share of the total variance and is considerably less reliable than the main effect. $\bar{X}_T$ captures by far most of the true-score variance in the Hick RT data. A sum (or mean) of a number of correlated values each containing error components is *more* reliable than the average reliability of the single values, and the reliability is *directly* related to the average correlation among the separate values. The value $b_{TB}$ (or $\sigma_T$), on the other hand, is essentially the average of the difference scores (i.e., $RT_{BIT3} - RT_{BIT2}$, $RT_{BIT2} - RT_{BIT1}$, etc.), which are always *less* reliable than the average reliability of the single values entering into the differences, and the reliability is *inversely* related to the average correlation between the separate values.[1]

Simple inspection of the above formulas reveals why $r_{ab}$ is inevitably a negative correlation; for all individuals with the same value of $\bar{X}_T$, those who have larger values of $\sigma_T$

will necessarily have smaller values of $a_T$ and vice versa. Hence directly correcting the *negative* correlation $r_{ab}$ for attenuation is an improper test of the theoretically based hypothesis that $r_{ab}$ is *positive*. The disattenuated value of $-r_{ab}$ (if the obtained value is not zero) would be larger, but of course it would remain a negative value.

A proper test of the hypothesis can be had by obtaining the Hick parameters from aggregated data, which consists of using the mean values of RT for each level of BITs obtained in differing groups of participants. Individual measurement error (i.e., the standard error of measurement, $SE_{meas}$) tends to be averaged out in the mean, decreasing toward zero the larger the sample size. For a sample of size $N$, this error component is $\pm SE_{meas}/\surd N$. When the means for $a$ and for $b$ were obtained for each of 27 diverse samples (Jensen, 1987, Table 3) totalling 1,850 participants (averaging 68 participants per sample), the correlation $r_{ab}$ is +.71, in marked contrast to the average correlation ($r_{ab} = -.15$) when $a$ and $b$ are calculated on the RT data for individuals. Thus, when the above-noted statistical artifact is largely overcome by aggregation, the theoretical expectation that $r_{ab}$ is a substantial positive value is clearly borne out.

## Slope Differences Between Groups that Differ in IQ

The slope parameter of the Hick function, contrary to theoretical prediction, generally shows small and typically nonsignificant negative correlations with IQ in most studies, the average correlation (in 35 studies totalling 1558 participants) being about –.12 (data from Jensen, 1987). A quite different picture appears when we look at comparisons between groups that, on average, differ significantly in IQ, although the groups contain no individuals who could be classified as mentally subnormal. Figure 1, for example, shows Hick RT data for two age-matched groups of seventh graders; one group was labelled "gifted" (IQ above 130), the other group, "average," was from regular classes in a white middle-class neighborhood). The groups differed 1.9σ in mean IQ (Cohn, Carlson, & Jensen, 1985). The result is clearly apparent to the naked eye. Both the $a$ and the $b$ differ significantly ($p < .001$); the groups differ in $b$ by 0.70σ, which is 37 percent as much as they differ in IQ. These results bear out the theoretical prediction that "rate of gain of information" (to use Hick's terminology for the slope of the Hick function) is related to IQ.

Another comparison, shown in Figure 2, is between students in a vocational college and undergraduates in a selective university; the average IQ difference between the groups is about 1σ (Jensen, 1987, Table 24, SID Nos. 12-15, 1-7, 10, 11). The data are detailed in Table 1 and compared with another measure of individual differences derived from the RT data, intraindividual variability ($s_i$), which is the average standard deviation of an individual's RT over a given number of trials within each level of BITs. In a majority of studies, the $s_i$ has the largest correlation with IQ of any other variable derived from the Hick RT data (Jensen, 1992). We see in Table 1 that the effect size ($d = 0.91$)) for the intercept is almost twice as large as for the slope ($d = .47$), while for intraindividual variability $d =$ 1.03. The point-biserial correlation ($r_{pb}$) is the correlation between the dichotomized groups (quantitized as 0 and 1, respectively) and the particular RT variable. Since the two groups differ, on average, in IQ, this is the only means by which we can estimate the correlation between IQ and each of the RT variables when they are based on data aggregated over individuals within each group. These point-biserial correlations are substantial (and,

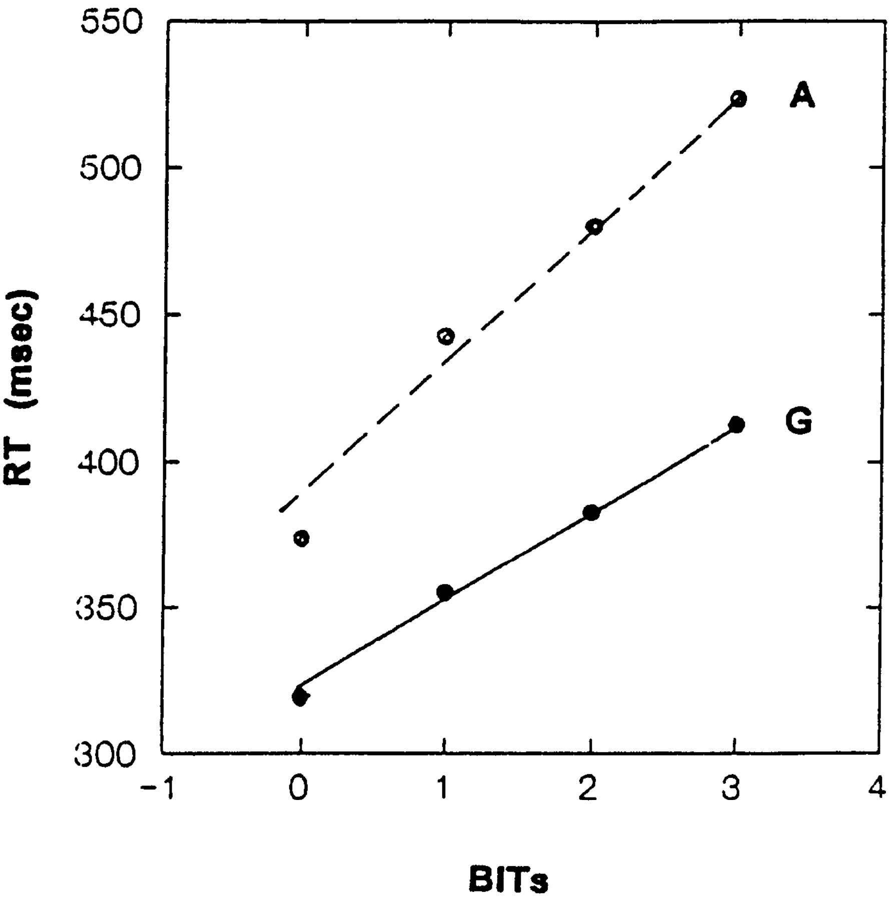


**Figure 1.** The Hick function plotted as the mean for 0 to 3 BITs for groups of average (A) and gifted (G) 7th-grade students. The regression for A is RT = 382.63 + 46.32 BIT; the regression for G is RT = 320.44 + 32.73 BIT. ($N_A = 72$; $N_G = 60$)

as each of the values of *d* is significant beyond the .001 level, so too are the corresponding values of $r_{pb}$), although RT slope again has the smallest correlation with IQ of any RT parameter, as is generally the case for individuals.

## Intercept and Slope as Mutual Suppressor Variables

Because there is a mathematically "built-in" negative correlation between the errors of measurement in *a* and *b*, the correlation of either *a* or *b* with IQ (or any independently measured variable), is suppressed by the other parameter (*a* or *b*). A suppressor variable (*z*) is defined as a variable for which, when it is partialled out of the zero-order correlation $r_{xy}$ between two other variables (*x* and *y*), the resulting partial correlation $r_{xy \cdot z}$ is greater than

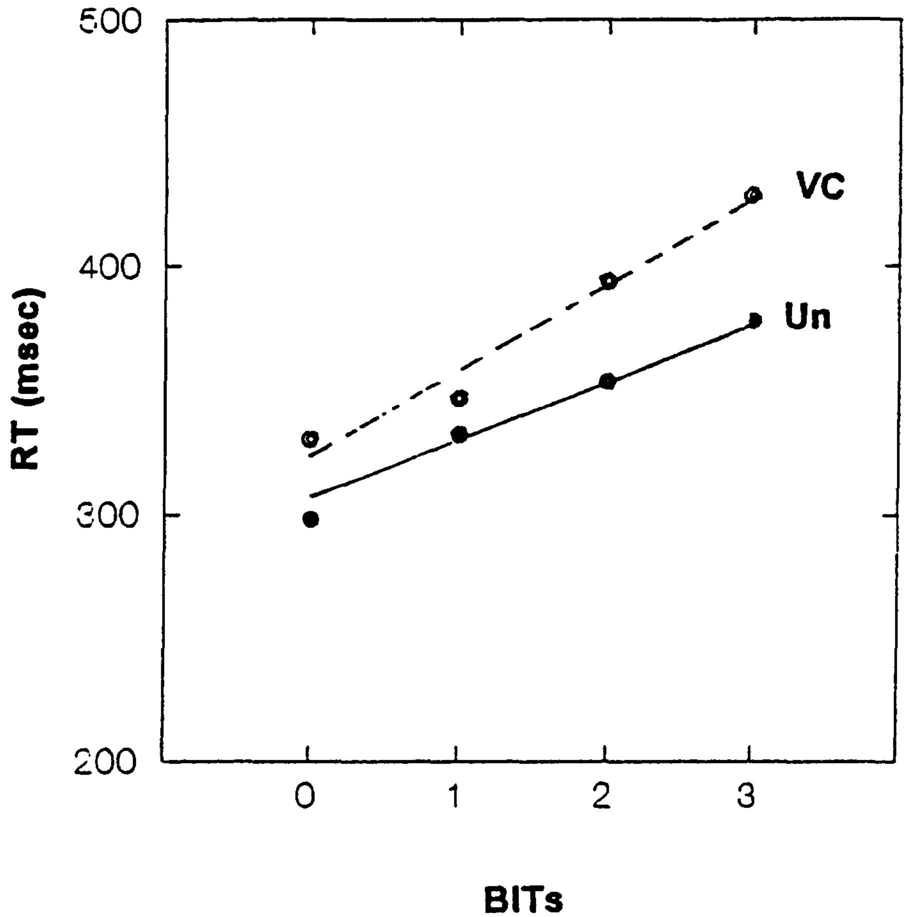


**Figure 2.** The Hick function plotted as the mean RT for 0 to 3 BITs for groups of Vocational College (VC) and University (Un) students. The regression for group VC is RT = 323.63 + 34.38 BIT; the regression for group Un is RT = 301.71 + 26.22 BIT. ($N_{VC}$ = 324; $N_{Un}$ = 530).

$r_{xy}$. This can be demonstrated using the following average (individual) correlations based on large samples (from Table 26 in Jensen, 1987, p. 161), where IQ is symbolized by $\mathfrak{I}$: $r_{ab} = -.277$, $r_{a\mathfrak{I}} = -.191$, and $r_{b\mathfrak{I}} = -.165$. The partial correlations, both larger than the zero-order correlations, are: $r_{a\mathfrak{I}.b} = -.250$ and $r_{b\mathfrak{I}.a} = -.231$.

## The Multiple Linear Regression of IQ on RT Variables

Using the same data as were used above in the partial correlations, it is instructive to observe the multiple correlations and standardized regression coefficients of the several RT variables derived from the Hick paradigm when IQ is regressed on various combina-

**Table 1.** Statistical Comparison of Group Mean Differences in Intercept ($a$), Slope ($b$), and Intra-individual Variability ($s_i$) of Reaction Time (msec.) in the Hick Function

| *Group* | *N* | *a* | $SD_a$ | $d^1$ | *t* | $r_{pb}$ |
|---|---|---|---|---|---|---|
| Intercept | | | | | | |
| Vocational College | 324 | 330.6 | 36.45 | | | |
| University | 530 | 301.1 | 29.73 | | | |
| *Mean Difference* | | 29.5 | | 0.91 | 12.26* | -.404 |
| Slope | | | | | | |
| Vocational College | 324 | 32.7 | 16.55 | | | |
| University | 530 | 26.5 | 10.37 | | | |
| *Mean Difference* | | 6.2 | | 0.47 | 6.03* | -.224 |
| Intra-individual Variability | | | | | | |
| Vocational College | 324 | 54.1 | 26.27 | | | |
| University | 530 | 34.6 | 12.59 | | | |
| *Mean Difference* | | 19.5 | | 1.03 | 12.54* | -.533 |

*Note:* [1] Effect size, $d$ = (Mean Difference)/(Average $SD$), where

Average $SD$ $\sqrt{[N_1(s_1^2) + N_2(s_2^2)]/(N_1 + N_2)}$.

* $p < .001$ (2-tailed test)

tions of these RT variables (Table 2). The parameters $a$ and $\bar{X}_{RT}$ are never paired in these regression analyses, because they are so highly correlated ($r$ = .979) as to be mutually redundant and collinear in the regressions. The optimum combination of RT variables accounting for IQ is $\bar{X}_{RT} + s_i$, neither of which is intrinsic to the Hick function per se.

The intraindividual variability, $s_i$, clearly makes the strongest contribution to the multiple $R$ with IQ, even though $s_i$ is an individual's trial-to-trial variability *within* each level of BITs, averaged over levels, a point elaborated elsewhere (Jensen, 1992). However, the average value of $s_i$ increases markedly at each successive level of BITs (although, unlike RT, which is a linear function of BITs, $s_i$ is a linear function of the number of S-R alternatives (Jensen, 1982, p. 104).

## Discussion and Conclusion

The slope, $b$, of RT in the Hick function is a psychometrically poor variable in the study of individual differences in information processes and their relation to IQ or to psychometric $g$. As measured in individuals, it is benighted by its relatively small variance, its low reliability, and its artifactual negative correlation with the intercept, $a$, and the individual's overall RT (averaged over all trials), $\bar{X}_{RT}$. Because $a$ acts as a suppressor variable in the correlation between $b$ and any external variable (e.g., IQ), such correlations, when used in any theory-testing way, should control for $a$ by partial correlation or by the use of corrected $b$ measures derived from its regression on $a$. Even then, $b$ is typically the weakest correlate of psychometric $g$ compared with all of the other Hick-derived variables: $\bar{X}_{RT}$, $a$, and especially $s_i$. Recent studies (e.g., Bates & Stough, 1997), however, suggest that some ingenious improvements in the chronometric procedures themselves can considerably increase the correlation between the RT slope and psychometric $g$.

**Table 2.** Standardized Regression Coefficients β and Multiple Correlation ($R$) for the Linear Regression of IQ on Various Combinations of the Hick Paradigm[a]

| *RT Variables*[b] | β | $R^c$ |
|---|---|---|
| $a$ | -.256 | .296 |
| $b$ | -.236 | |
| $a$ | .001 | .41 |
| $s_i$ | -.417 | |
| $b$ | .040 | .418 |
| $s_i$ | -.436 | |
| $b$ | -.023 | .390 |
| $\bar{X}_{RT}$ | -.381 | |
| $s_i$ | -.284 | .441 |
| $\bar{X}_{RT}$ | -.196 | |
| $s_i$ | -.305 | |
| $\bar{X}_{RT}$ | -.202 | .444 |
| $b$ | .054 | |

*Notes:* [a] The zero-order correlations from which the β and $R$ were calculated are based on the unweighted averaged data from a number of studies (total $N > 700$); the correlations have been corrected for attenuation and for range restriction of IQ (Jensen, 1987, Table 26).
[b] $a$—intercept, $b$—slope, $\bar{X}_{RT}$—mean $RT$, $s_i$—intra-individual variability (i.e., the standard deviation of an individual's $RT$s over $n$ number of trials within each level of BITs averaged over all of the levels of BITs).
[c] All of the values of β and R greater than |0.10| are significant at $p < .001$, 2-tail test.

The Hick parameter $b$ serves a theoretically useful purpose. When its correlation with IQ, or psychometric $g$, is measured in a way that lessens its psychometric handicaps, by aggregation of individual measurements, disattenuation, and regressing out its suppressor, $a$, it very clearly accords with the theoretical prediction that for each unit of increase in task complexity, as measured in BITs, the uniform increments in RT are inversely related to individuals' speed of information processing and to $g$. Related to this is the fact that the magnitude of the correlation itself between IQ and RT increases as a function of the number of BITs of information in the set of S-R alternatives (Jensen, 1982, p. 109; 1987, p. 164).

Also, studies of the Hick paradigm based on group comparisons of children in different age brackets shows that the RT slope decreases systematically with increasing age and "mental growth" (Jensen, 1987, p. 150, Table 20). Extended amounts of practice on the Hick task also leads to greater efficiency or speed of processing, with a consequent decrease in $b$ as the amount of practice increases. Teichner and Krebs (1974, Figure 8), in their study of the effects of differing amounts of practice (after $10^0$, $10^1$, $10^2$, $10^3$, $10^4$, and $10^5$ trials) on the mean slope of RT across six levels of complexity (1 to 6 BITs), showed a marked, systematic decrease in both the intercept and slope with increased amount of practice. In this sense, differing amounts of practice might be said to simulate the observed RT results of age differences among children and of IQ differences between participants when all are given the same number of trials on the Hick task. Whether the causes of these two classes of similar-appearing effects—individual differences and practice effects—have processes in common is open for speculation and empirical investigation.

## NOTE

1. The reliability of the sum (or mean) of two measures $(X + Y)$ is: $r_{(X+Y)(X+Y)} = (r_{XX} + r_{YY} - 2r_{XY})/2(1 - r_{XY})$. The reliability of the difference $(X - Y)$ is: $r_{(X-Y)(X-Y)} = (r_{XX} + r_{YY} + 2r_{XY})/2(1 + r_{XY})$.

## REFERENCES


Bates, T., & Stough, C. (1997). Processing speed, attention, and intelligence: effects of spatial attention on decision time in high and low IQ subjects. *Personality and Individual Differences, 23*, 861-868.

Cohn, S.J., Carlson, J.S. & Jensen, A.R. (1985). Speed of information processing in academically gifted youths. *Personality and Individual Differences, 6, 621-629.*

Jensen, A.R. (1982). Reaction time and psychometric *g*. In H.J. Eysenck (Ed.), *A model for intelligence* (pp. 93-132). New York: Springer.

Jensen, A.R. (1987). Individual differences in the Hick paradigm. In P.A. Vernon (Ed.), *Speed of information-processing and intelligence* (pp. 101-175). Norwood, NJ: Ablex.

Jensen, A.R. (1992). The importance of intraindividual variability in reaction time. *Personality and Individual Differences, 13*, 869-882.

Lohman, D.F. (1994). Component scores as residual variation (or why the intercept correlates best). *Intelligence, 19*, 1-11.

Marascuilo, L., & Levin,, J.R. (1983). *Multivariate statistics in the social sciences: A researcher's guide.* Monterey, CA: Brooks/Cole.

Teichner, W.H., & Krebs, M.J. (1974). Laws of visual choice reaction time. *Psychological Review, 81*, 75-98.