INTELLIGENCE **14**, 375–388 (1990)

Simple Reaction Time as a Suppressor Variable in the Chronometric Study of Intelligence

ARTHUR R. JENSEN
University of California, Berkeley

T. EDWARD REED
University of Toronto

Individual differences in reaction time (RT) to various elementary cognitive tasks (ECTs) reflect variance in both peripheral (sensorimotor) and central components of information processing. Minimizing the variance associated with peripheral processes by controlling simple RT in chronometric studies of more complex ECTs involving choice, discrimination, memory scanning, or other central processes, can increase the correlation between RTs and scores on complex psychometric tests of ability, thereby identifying more clearly the elementary processes involved in psychometric *g* and other abilities. Simple RT can be controlled by subtraction, partial correlation, and multiple correlation. The relative effectiveness of the different methods depends on various features of the chronometric data. The methods are explicated, with examples from a study of simple, choice, and discriminative RTs (the odd man out paradigm) in 213 male students from three colleges.

One of the oldest and least disputed facts in psychology is that reaction time (RT) is a complex variable and is analyzable into a number of components. Donders (1868–1869/1969) was the first to suggest that this basic fact of RT could be used to measure the speed of various processes, such as discrimination and choice. He proposed the *subtraction* method for analyzing RTs into their constituent elements. For example, he devised techniques to measure three types of reaction, labeled *a, b,* and *c.* The *a-reaction* is simple RT, that is, a uniform response to the onset of a single stimulus (usually visual or auditory); it is a measure of stimulus *apprehension.* The *b-reaction* required the subject to *discriminate* the particular stimulus (out of two or more different stimuli) that had occurred and also required the subject to *choose* the appropriate response for the particular stimulus, as each stimulus required a different response. The *c-reaction* required the subject to *discriminate* the particular stimulus (out of two or more stimuli) that had occurred, but the subject was required to respond to only one of the stimuli (always the same one), hence, not requiring any choice of

---

Correspondence and requests for reprints should be sent to Arthur R. Jensen, School of Education, University of California, Berkeley, CA 94720.

response. It was found that the average lengths of RT to the three conditions were $a < c < b$. Donders argued that subtracting the RT for the a condition from RT for the c condition measured the time required for *discrimination,* and subtracting the RT for c from the RT for b measured the time required for *choice* of response.

Donders's subtraction method has since been questioned on the grounds that it assumes perfectly sequential processing and ignores the possibility that some processes may overlap in time or occur simultaneously (*parallel processing*), or that there may be interactions between different processes (Jensen, 1985). Nevertheless, a fundamental distinction can be made between two aspects of *any* type of RT, including simple RT (SRT) as well as choice RT (CRT). As explained by Luce (1986):

> The first thing that simple reaction-time data seems to suggest is that the observed reaction times are, at a minimum, the sum of two quite different times. One of these has to do with decision processes, invoked by the central nervous system, aimed at deciding when a signal has been presented. The other has to do with the time it takes signals to be transduced and transmitted to the brain and the time it takes orders issued by the brain to activate the muscles leading to responses. (p. 94)

That strictly sensory processes, which involve different degrees of *sensory lag,* are a component of RT is evident from the fact that RTs differ for stimuli of different sensory modalities (such as auditory RT being faster than visual RT) partly because of cone transduction delay, and RT varies in any given modality as an inverse function of stimulus intensity. Similarly on the response side, RT varies according to the nature of the required response, and it is known that the time for neural transmission from the motor cortex to the finger is about 20 ms. (Reed, 1988; Rossini, Marciani, Caramia, Hassan, & Cracco, 1986). SRT comprises such elements as attention, sensory lag, afferent neural transmission of the signal to the brain, central encoding of the stimulus, efferent transmission from the brain to the muscles, muscle lag, and response execution time. More complex types of RT, requiring discrimination and/or choice, or other metal processes, involve all of the elements included in SRT and, in addition, the time required for other, presumably central, processes. Hence we can distinguish between *peripheral* and *central* components in any kind of RT. All RTs necessarily contain the times for peripheral components, but various forms of RT, such as Donders's *a, b,* and *c* reactions, contain quite different amounts of time for central components.

REACTION TIME AND INTELLIGENCE

Numerous studies reviewed elsewhere (Jensen, 1982, 1987, 1988), have shown negative correlations between various measures of RT and assessments of general intelligence, such as IQ. From the earliest study (Gilbert, 1894) to the present, various forms of CRT have been found to be more highly correlated with IQ than SRT, with few exceptions (e.g., Detterman, 1987). A meta-analysis of 15 independent studies comprising 1129 subjects showed an unweighted mean correlation between SRT and IQ of $-.19$ ($p < .001$) and correlations for CRTs involving 2, 4, and 8 choice alternatives of $-.21$, $-.24$, and $-.26$, respectively (Jensen, 1987, pp. 162–163). In general, as the RT task increases in complexity or the amount of information processing required, the RT increases and so does its correlation with IQ. These relationships seem especially clear when increases in task complexity involve additional cognitive processes, such as discrimination and retrieval of information from short-term or long-term memory, rather than merely an increasing demand on a single process. It is a reasonable assumption that the strictly sensory and motor aspects of RT not involving attention, discrimination, decision, or other central processes, are uncorrelated with psychometric *g* in the normal population.

It is a most plausible hypothesis that the correlation between IQ and CRT (i.e., generally, RT to elementary cognitive tasks of greater complexity than those used to measure SRT) is solely attributable to *central* components of RT and not at all to sensorimotor, or *peripheral,* components. CRT is almost invariably greater than SRT, and CRT also shows greater variance of individual differences than SRT. The greater variance of CRT relative to that of SRT reflects greater variance in central processing components. Since SRT reflects less variance in central processes than any form of CRT, but reflects variance in the same peripheral components as enter into CRT, and if the RT–IQ correlation is due to common central processing components, then, assuming a linear stages model of CRT, *subtracting* individuals' SRTs from their CRTs should, theoretically, not weaken the true CRT–IQ correlation but strengthen it. That is, we should expect IQ to be more strongly correlated with the true-score component of the variable CRT − SRT than with CRT. In other words, SRT would act as a *suppressor variable* in the multiple correlation, *R*, of SRT and CRT with IQ, and in the partial correlation between CRT and IQ with SRT partialed out. A *suppressor variable* is defined as (a) a variable which, when subtracted from another variable, increases that variable's correlation with another variable (i.e., if $r_{xy} < r_{x(y-z)}$, z is a suppressor variable), or (b) as a variable which, when statistically partialed out of a correlation between two other variables, increases the correlation (i.e., if $r_{xy} < r_{xy.z}$, the variable z is a suppressor variable).

Subtraction, partial r, or multiple R? In the case of conventional psychometric tests, the difference between an individual's standardized scores on text X and test Y (i.e., z_{x-y}) represents a difference in the individual's relative standings on the two tests with respect to the distribution of z scores in a particular reference group. But RT measures, unlike psychometric test scores, are on a ratio scale, on which real-time differences are meaningful physical units independent of any reference group. Hence, mental chronometry has all the scientific advantages of a true ratio scale, which is extremely rare in traditional psychometrics.

If, as we have hypothesized, the *peripheral* sources of variance in CRT attenuate the correlation between CRT and IQ (or psychometric g), a correlation which reflects *central* processes, then removing the peripheral variance should enhance the correlation. Since SRT reflects peripheral processes to much the same degree as does CRT but reflects central processes to a much lesser degree, removal of the effects of peripheral processes on the CRT–IQ correlation can be effected to some degree by one of three methods: (1) simple subtraction of SRT from CRT (where RT is the median obtained from a large number of trials), (2) partialing SRT out of the CRT–IQ correlation, and (3) a multiple correlation between SRT and CRT (as the independent variables) and IQ (as the dependent variable). Symbolizing SRT and CRT as s and c, respectively, and IQ as g, the three methods can be represented as (1) $r_{(c-s)g}$, (2) $r_{cg.s}$, and (3) $R_{g:cs}$, respectively. What are their advantages and disadvantages?

Subtraction ($r_{(c-s)g}$) is theoretically compelling, especially in the case of a ratio scale such as RT. The subtraction, CRT − SRT, for each individual is a *within-subject* "correction" which controls for the peripheral time component, and therefore it has higher potential validity as a corrected score than could be obtained with partial correlation, since the correction depends on no assumptions about the linearity (or any other form) of the regression of one variable on another. The possible disadvantage of simple subtraction arises from two factors: (1) measurement error in SRT and CRT and (2) the sizable correlation between SRT and CRT, since both have their peripheral elements in common and probably share some variable fraction of the central processing elements. With imperfect reliability of the two correlated measures, the reliability of the difference between them will be less than the reliability of either one, according to the formula for the reliability of a difference, as follows:

$$r_{(c-s)(c-s)} = \frac{r_{cc} + r_{ss} - 2r_{cs}}{2(1 - r_{cs})} . \qquad \text{(Eq. 1)}$$

Hence, to the extent that the reliabilities (r_{cc} and r_{ss}) are low and the correlation (r_{cs}) between SRT and CRT is high, the absolute value of the correlation $r_{(c-s)g}$ will tend to approach or fall below r_{cg}. Then nothing would be gained by subtraction. In other words, the theoretical advantage of subtracting SRT from CRT could be counteracted by the practical disadvantage of using a much less reliable difference score.

Partialing SRT out of the CRT–IQ correlation is a *between-subjects* "correction," based on the linear regression of CRT on SRT in a given sample. What, in effect, is subtracted from a subject's CRT is the least-squares best fitting *mean* CRT of all the other subjects having the same SRT. This regressed value (i.e., CRT minus the mean CRT for a particular value of SRT) is therefore in principle necessarily a less valid correction for an individual than would be direct subtraction, CRT − SRT. But what a regressed score may lose in *potential* validity it may gain in reliability and hence in *realized* validity. In other words, the regression line (being essentially a mean) may be considerably more reliable than are individual measures of SRT. Individual measurement errors cancel each other in the mean. As explicated by Gulliksen (1950, pp. 39–45), the standard error of the difference $y - x$ (assuming for simplicity of formulation that $\sigma_x = \sigma_y$) is

$$SE_{y-x} = \sigma_x(\sqrt{1 - r_{xy}})\sqrt{2}\,, \qquad \text{(Eq. 2)}$$

whereas the standard error of the difference $y - \hat{y}$ (where $\hat{y}$ is the regression estimate of y for a given value of x) is

$$SE_{y-\hat{y}} = \sigma_x\sqrt{1 - r_{xy}}\sqrt{1 + r_{xy}}\,. \qquad \text{(Eq. 3)}$$

It can be seen that $SE_{y-x} > SE_{y-\hat{y}}$. For this reason, $r_{(c-s)g}$ may be smaller (in absolute value) than r_{cg}. So, depending on the reliabilities of SRT and CRT and the size of the correlation between them, $r_{(c-s)g}$ may be either larger or smaller than $r_{cg.s}$. The better method, then, is whichever one—subtraction or partialing—yields the higher correlation.

Like subtraction, partialing has the disadvantage that it removes from CRT not only the peripheral elements that CRT and SRT have in common, but also whatever central elements that are reflected in SRT. Attention or arousal, uncertainty as to when the reaction stimulus will occur, and the decision that it has occurred are all central processes that enter into SRT. In groups of children or samples including adults of below-average IQ, variance in the elementary central processing involved in SRT may constitute a considerable fraction of the total variance in SRT, as indicated by a moderate correlation between SRT and IQ. In such samples, either subtracting or partialing SRT out of CRT would remove not only the unwanted peripheral variance but also enough of the central processing variance as to render $r_{(c-s)g}$ or $r_{cg.s}$ smaller than r_{cg}.

Multiple correlation overcomes this particular problem. In the multiple correlation of CRT and SRT with g, $R_{g:cs}$, only that part of the variance in SRT which it does not have in common with g (that is, the peripheral component) acts as a suppressor variable in the multiple regression equation. Assuming that the central components in SRT are some fraction of all the central components in CRT and that the peripheral components in either SRT or CRT are less correlated with g than are the central components, then SRT in the multiple regression can act as a suppressor variable to the extent that it has peripheral components in common with CRT. Including SRT in the regression equation suppresses the unwanted peripheral variance, thereby making $R_{g:sc} > r_{gc}$. The workings of this may be better understood by expressing the multiple R in terms of partial correlations:

$$R_{g:cs} = (r^2_{gc} + r^2_{gs.c} - r^2_{gc}r^2_{gs.c})^{1/2}. \qquad \text{(Eq. 4)}$$

If $r_{gs} = 0$ or some relatively small value and r_{gc} and r_{sc} are considerably larger, in accord with our hypothesis, then $r_{gs.c}$ will have a substantial value (positive, in the case of reaction times), and its role in Equation 4 can be seen to increase the multiple R. The use of multiple R is generally a better method than partial r for ridding the CRT–IQ correlation of the attenuating peripheral variance, because, unlike $r_{gc.s}$, which removes the component that s has in common with g but not with c, $R_{g:cs}$ does not remove any part of s that it has in common with g but not with c. In $R_{g:cs}$, only that component is suppressed that c has in common with s and that neither c nor s has in common with g. In other words, presumably only the peripheral or non-g component of CRT (and SRT) is suppressed in the multiple R. If there is any central (or g) component in SRT that is not in CRT, it will of course increase the value of $R_{g:cs}$. However, the linear state model assumes that all of the components in SRT are also in CRT. The validity of this assumption is an empirical question that can be addressed by means of multivariate statistical analysis. Reaction time data similar to those in the present study show a close fit to this model (Jensen, 1987, pp. 139–141).

The one disadvantage that multiple R shares with partial r is that it is a correction based on *between-subjects* statistics and their assumption of linear regressions of each of the three variables upon each of the others. Hence its potential validity is less than for the subtraction method. But the realized validity of the multiple R method may exceed that of the subtraction method if the reliability of the difference scores (CRT − SRT) is much lower than the regressed scores or if the SRT contains too large a fraction of the same central elements common to CRT and g.

The effects on correlation of removing the predominantly peripheral variance common to both SRT and CRT from the correlation between CRT and IQ has been examined here by means of three RT paradigms—SRT and two forms of CRT that differ considerably in their central processing demands.

METHOD

Subjects

Students from three postsecondary educational institutions in the eastern San Francisco Bay region of California who were male, between 18 and 25 years (inclusive) of age, and of European ancestry, were asked to participate. All 213 subjects were paid volunteers; 123 were university students and 90 were students in two community colleges. All were tested individually under the same conditions in the first author's laboratory.

Tests

Psychometric. In virtually all of the RT studies performed in this laboratory over the past 10 years or so, the single preferred measure of psychometric g has been the Raven Progressive Matrices. Therefore, it seemed desirable for possible comparisons with past studies to use this test in the present study. But there was a question of which form of the Raven to use. Because it was known from previous experience that the Standard Progressive Matrices (SPM) imposes a severe ceiling effect on the score distribution in the university population, and there is risk of a "floor effect" on the Advanced Progressive Matrices (APM) in the community college population, it was decided to use the APM in the university group and the SPM in the community college group. In order to be able to put the scores of both forms on a single common scale, the SPM and APM were equated by administering both forms, along with the nationally standardized Otis–Lennon IQ test, to a large sample of students in another state university which largely encompassed the range of ability found within both the university and community college populations. The equated scores derived from the two forms of the Raven were also equated with the Otis–Lennon IQ, so that the scores of all subjects could be expressed in terms of the nationally standardized Otis–Lennon IQ scale, with a general population $m = 100$, $SD = 16$. The details of this equating have been explicated elsewhere (Jensen, Saccuzzo, Larson, 1988). The Raven test was administered to subjects individually, without time limit, usually taking from 30 to 60 min.

Chronometric. Three chronometric tests were used, which we will call simple RT (or RT1), choice RT (or RT8), and discriminitive RT, or the odd man out paradigm (Oddman, for short), which was introduced in a study by Frearson and Eysenck (1986).

The subject's response consoles for the three tests are shown in Figure 1 (p. 382).

In RT1 (A in Fig. 1), the subject responds to only a single stimulus (the onset of a green under-lighted pushbutton), with no prior uncertainty of the nature or location of the reaction stimulus (RS). The only element of uncertainty for the subject is in not knowing just *when* the green light will go on during the interval (maximum of 4 s) following the preparatory signal (beep). The subject lifts his index finger from the home button (see Fig. 1) and touches the green light to turn it off.

In RT8 (B in Fig. 1), the subject is confronted by a semicircle of eight under-lighted buttons exactly like the RS in RT1, but he is uncertain about which one of the eight buttons will light up at random following the preparatory signal. This uncertainty in RT8 is well known to increase the subject's RT, on average, over that for RT1 (Jensen, 1987).

In the Oddman paradigm (C in Fig. 1), the subject is confronted by the same semicircle of buttons as in RT8; but in this case, three out of the eight buttons light up simultaneously. Two of the lighted buttons are always closer together in the array than either one is to the third lighted button (i.e., the odd man out), which the subject must touch to turn out all three lights at once. The odd button can be any one of the eight buttons and is randomized across trials. The distances

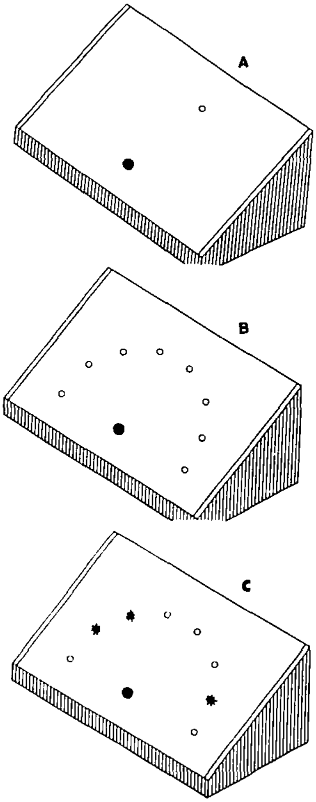

FIG. 1. The response consoles for RT1 (A), RT8 (B), and the odd man out RT (C). The black dot in the lower center of each panel is the "home" button. The open circles, 15 cm from the home button, are green under-lighted pushbuttons. In the RT1 and RT8 conditions (A and B), only one green pushbutton lights up on each trial; on the Oddman task, three pushbuttons light up simultaneously on each trial, with unequal distances between them (shown in C), the remotest one from the other two being the odd man out, which the subject must touch.

between the three lighted buttons varies unsystematically from trial to trial. (With eight buttons in the semicircular array, the total number of possible oddity patterns is 44).

Trials are subject-paced in all tasks. A trial begins when the subject places the index finger of his preferred hand on the home button and keeps it depressed. After 1 s a preparatory signal (a 1-s "beep") occurs followed by a random interval of 1 to 4 s, following which the reaction stimulus (a green under-lighted pushbutton) comes on. The subject responds as quickly and accurately as possible by lifting his finger from the home button to touch the lighted pushbutton, which turns off the light (as well as the other lighted pushbuttons, in the Oddman condition).

RT is the interval, recorded in milliseconds (ms), between the onset of the reaction stimulus and the subject's lifting his finger from the home button. The entire procedure is computerized, and the subject's median RT is automatically computed for *n* test trials; the *n*s for RT1, RT8, and Oddman were 20, 30, and 36 trials, respectively. Each test was preceded by eight practice trials typical of the test trials that followed.

The three tests obviously differ in the complexity of their information-processing demands. Besides measuring such peripheral components of RT as sensory lag, afferent and efferent neural transmission, and muscle lag, RT1 measures stimulus apprehension under near minimal stimulus uncertainty; the only element of uncertainty is the exact moment at which the reaction stimulus will occur during the brief random interval following the preparatory stimulus. RT8 involves all of the elements of RT1 as well as the additional uncertainty as to which one of the eight green pushbuttons would light up, which was a random choice programmed in the computer. The Oddman test involves all of the processing demands of RT8 plus the discrimination of the most remote pushbutton, that is, the "odd man out." The increase in processing demands going from RT1 to RT8 to Oddman, should be reflected in corresponding increases in RT.

RESULTS

Means and Standard Deviations

As can be seen in Table 1, the three chronometric tests do indeed differ markedly in mean RTs, the overall difference RT8 − RT1 being about 51 ms., and Oddman − RT8 being about 139 ms. RT1 and RT8 are in close agreement with previous RT1 and RT8 data obtained in numerous studies of similar groups (Jensen, 1987, Table 3, p. 115). The limited literature on the Oddman test (consisting entirely of Frearson and Eysenck [1986] and Frearson, Barrett, and

TABLE 1

Summary Statistics in University (Un), Community College (CC), and Total Group

	Group						**Mean Diff.**
	Total (*N* = 213)		**Un (*N* = 123)**		**CC (*N* = 90)**		**$\lvert \text{Un} - \text{CC} \rvert$**
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*t*
Age	20.34	2.02	20.61	2.09	19.97	1.85	2.36^a
IQ	118.79	10.60	123.98	8.19	111.69	9.35	9.98^c
RT1	270.60	29.23	275.04	31.84	264.53	24.11	2.74^b
RT8	321.75	33.85	326.31	34.50	315.52	32.09	2.35^a
Oddman RT	460.35	63.41	462.57	68.27	457.32	56.33	<1
RT8 − RT1	51.15	23.87	51.27	24.88	50.93	22.54	<1
Oddman − RT1	189.75	59.12	187.53	64.34	192.79	51.33	<1
Oddman − RT8	138.60	48.40	136.26	52.37	141.81	42.47	<1

$^a p < .05$, 2-tailed; $^b p < .01$, 2-tailed; $^c p < .001$, 2-tailed.

Eysenck [1988]) does not provide data that would permit direct comparison with that of the present study.

An apparent anomaly in these data is the fact that on the chronometric measures the university and community college groups differ either in the unexpected direction or nonsignificantly, as indicated by the two-tailed t tests in Table 1, despite the fact that the groups differ about 12 points in IQ, as would be expected from the differing academic admission standards of the two types of institutions. We can offer no explanation for this apparent anomaly. Although unlikely, it is possible that the equating of the Standard and Advanced forms of the Raven in another college sample, though perfectly correct technically, is for some reason not strictly applicable to the self-selected groups used in the present study. However, the IQs derived from the equating procedure were highly correlated with the Raven raw scores *within* each group (+.99 in the university group; +.97 in the community college group). The small but significant age difference between the groups could not have significantly affected the group differences on the chronometric tests, as these showed no significant or systematic relation to age in either group. In any case, all the analyses in this study were performed separately on the two groups as well as on the combined groups.

Correlation Between RTs and IQ

The zero-order Pearson correlations among all the variables in the study are shown in Table 2. They are the basis of the partial and multiple correlations shown in Table 3.

Table 3 shows the correlations of RT8 and of Oddman with IQ and the correlations with IQ when (a) RT1 is subtracted from RT8 (or from Oddman RT), (b) RT1 is partialed out of the correlation between RT8 (or Oddman) and IQ, and

TABLE 2
Zero-Order Correlations Between Chronometric Variables and IQ in the University (Un), Community College (CC), and Total Group

Variable	Group	RT8	Oddman	IQ
	Un	.721^c	.353^c	+.153
RT1	CC	.713^c	.412^c	−.110
	Tot	.723^c	.372^c	+.142^a
	Un		.660^c	−.073
RT8	CC		.664^c	−.183
	Tot		.658^c	−.007
	Un			−.263^b
Oddman	CC			−.265^a
	Tot			−.190^b

$^a p < .05$, 2-tailed; $^b p < .01$, 2-tailed; $^c p < .001$, 2-tailed.

TABLE 3
Correlations of Chronometric Variables with IQ in the University (Un), Community College (CC), and Total Group

Variable	Group		
	Total	Un	CC
1. RT8	−.007	−.073	−.183
2. RT8 − RT1	−.184[b]	−.297[c]	−.143
3. RT8 partial RT1	−.161[a]	−.268[b]	−.150
4. Mult. *R* (RT1 & RT8)	.213[b]	.305[c]	.186
5. Oddman	−.190[b]	−.263[b]	−.265[a]
6. Oddman − RT1	−.274[c]	−.355[c]	−.239[a]
7. Oddman partial RT1	−.264[c]	−.346[c]	−.242[a]
8. Mult. *R* (RT1 & Oddman)	.297[c]	−.371[c]	.265[a]
9. Oddman − RT8	−.244[c]	−.295[c]	−.213[a]
10. Oddman partial RT8	−.246[c]	−.287[b]	−.195
11. Mult. *R* (RT8 & Oddman)	.246[c]	.295[c]	.265[a]

[a] $p < .05$, 2-tailed; [b] $p < .01$, 2-tailed; [c] $p < .001$, 2-tailed.

(c) when both RT1 and RT8 (or Oddman) are entered into a multiple regression correlation (R) with IQ.

In row 2 of Table 3, we see that subtracting RT1 from RT8 has increased the correlation of RT and IQ for the total group and university group. Hotelling's (1940) t test (with 2-tailed values) for the difference between correlations based on the same sample shows the increase in correlation to be significant in the total group ($t = 2.72$, $p < .01$), and in the university group ($t = 2.49$, $p < .02$), but not in the community college group ($t < 1$). Similarly, we see in row 6 of Table 3, subtracting RT1 from Oddman RT significantly increases the correlation with IQ in the total group ($t = 2.70$, $p < .01$) and in the university group ($t = 2.28$, $p < .03$), but not in the community college group ($t < 1$).

Rows 3 and 7 in Table 3 show the effect of partialing RT1 out of the correlation between IQ and RT8 and between IQ and Oddman RT. In the total group and the university group the partial r is not as large as the r obtained by the subtraction method. Both subtraction and partial r decrease the RT × IQ correlation in the community college group, probably because RT1 has more central components in this group and subtraction or partialing, besides removing variance associated with peripheral processes, may remove too much of the variance associated with the central processes that are associated with the covariance of IQ with RT1, RT8, and Oddman.

Rows 4 and 8 in Table 3 show the multiple correlation when RT1 and RT8 (or RT1 and Oddman RT) are the independent variables and IQ is the dependent variable. (Multiple R is always positive, since it is calculated as the square root of R^2, which, as a proportion of variance, is necessarily always positive.) The multiple *R*s are higher than all the other correlations, because whatever variance components the two predictors (RT1 and RT8, or RT1 and Oddman RT) have in common with IQ contribute to the R^2 and the variance components that the RT variables have in common with each other but not with IQ are suppressed, or partialed out, in effect. Although the squared multiple correlation, $R^2_{g:sc}$, can never be smaller than the squared partial correlation, $r^2_{gc.s}$, it is theoretically possible for $R^2_{g:sc}$ to be smaller than the squared zero-order correlation $r^2_{g(c-s)}$ obtained from the subtraction method (see Discussion). The R^2s indicate that only a small proportion of the total IQ variance (averaging about .08 for RT1 + RT8 and about .10 for RT1 + Oddman) is associated with these particular RT variables in the present samples, which have a quite restricted range of ability (see below). The multiple *R* for all three RT variables (RT1, RT8, and Oddman RT) for predicting IQ is .298 ($p < .01$) for the combined groups, .378 ($p < .01$) for the university group, and .265 ($p < .10$) for the community college group. Corrected for restriction of range on IQ (using McNemar's [1949, p. 126] Formula 50) based on a population $\sigma = 16$, the corresponding corrected correlations are .426, .624, and .426, respectively. (The corresponding unbiased [shrunken] estimates of the uncorrected *R*s are .245, .341, and .203, respectively.) Including RT8 in the multiple regression increases *R* by a negligible amount over the *R* yielded by only Oddman and RT1. Obviously, RT8 has virtually no variance in common with IQ that is not contained in Oddman.

DISCUSSION

Controlling for individual differences in simple RT, either by subtraction or by partialing, increased the correlations of choice RT (RT8) and discrimination RT (Oddman) with IQ measured with the Raven Progressive Matrices in the university group but not in the community college group. As noted in the Introduction, controlling for simple RT (SRT) will have different effects on the correlation between more complex forms of RT and IQ depending on several variables. When the reliabilities of the RT measures are low, neither subtraction nor partialing of SRT stands much chance of increasing the correlation of the more complex RT with IQ. Also, the *higher* the correlation between SRT and complex RT, the less is the chance of increasing the correlation by the subtraction method. On the other hand, the *lower* the correlation between SRT and complex RT, the less effective is the partialing method relative to the subtraction method. Multiple correlation, however, will always be a more effective means of controlling SRT than will partial correlation, because multiple *R* does not remove any of the variance that SRT may have in common with IQ, as does the partial *r*. But if there is very high reliability of the RT variables, a relatively low correlation between SRT and complex RT, and an even lower correlation between SRT and IQ, the subtraction method can yield a higher correlation than the multiple *R*. This is because the subtraction method is a within-subjects control of SRT, which can be more precise than the between-subjects regression on which the multiple R depends. Under typical conditions, however, the multiple R provides the best control, with SRT (or some part of its variance) acting as a suppressor variable in the multiple regression equation.

The only other data on the Oddman paradigm are two studies done in Eysenck's laboratory. The subtraction method was not used in these, but the published information permits us to calculate the partial and multiple correlations. From the first study (Frearson & Eysenck, 1986), based on a heterogeneous ability sample of 37 adults, we have the following correlations between RTs and Raven scores (figures in parentheses are the correlations corrected for attenuation):

RT1 × Oddman: $r = +.52$ (+.59)
RT1 × Raven: $r = -.28$ (−.32)
Oddman × Raven: $r = -.62$ (−.71)
Oddman × Raven, partial out RT1: $r = -.58$ (−.68)
Multiple correlation: $R = .62$ (.72)

The fact that multiple R is not larger than the zero-order r between Oddman and Raven means, of course, that RT1 contributes nothing to the prediction of IQ that is not predicted by Oddman. Hence its action as a suppressor variable is virtually nil in this case, a result which can best be understood in terms of Equation 4. (*Note:* The squared partial correlation between Raven and RT1 [with Oddman partialed out] is +.004, which corresponds to $r^2_{gs.c}$ in Eq. 4.)

The second study from Eysenck's laboratory (Frearson, et al. 1988) is based on a somewhat less heterogeneous sample of 89 adults. The simplest RT used was RT2 (i.e., a 2-alternative light/buttons condition on essentially the same apparatus used in the present study). So we must control RT2 in the correlation between Oddman and Raven IQ, as follows:

RT2 × Oddman: $r = +.67$
RT2 × Raven: $r = -.16$
Oddman × Raven: $r = -.48$
Oddman × Raven, partial out RT2: $r = -.51$
Multiple correlation: $R = .53$

In this case, both the partial r and the multiple R are larger than the zero-order r between Oddman RT and Raven score, meaning that RT2 acted as a suppressor variable in this sample. Since RT2 probably involves all the same peripheral factors as RT1, and it is unlikely that RT2 would have any variance in common with the Raven that is not also shared by Oddman, the partial r if RT1 had been removed would probably be slightly larger than when RT2 is partialed out, while the multiple R would be virtually the same whether it included RT1 or RT2 along with the Oddman RT.

We recommend that a measure of simple RT be obtained in all studies of the relationship between complex chronometric variables and psychometric *g* or other abilities, and that the relationship between the chronometric and psychometric variables be examined not only by means of their zero-order correlation, but also when variance in SRT is directly or statistically controlled to rid the correlation of possible attenuation by peripheral, noncognitive factors. Since the optimal method of control cannot feasibly be known beforehand, we suggest that all three methods—subtraction, partial *r*, and multiple *R*—be used in every case.

REFERENCES

Detterman, D.K. (1987). What does reaction time tell us about intelligence? In P.A. Vernon (Ed.), *Speed of information-processing and intelligence.* Norwood, NJ: Ablex.

Donders, F.C. (1969). Over de snelheid van psychische processen [On the speed of mental processes]. In W.G. Koster (Ed. and Trans.), *Attention and performance* (Vol. 2), *Acta Psychologia, 30,* 412–431. (Original work published 1868/69)

Frearson, W.M., Barrett, P., & Eysenck, H.J. (1988). Intelligence, reaction time and the effects of smoking. *Personality and Individual Differences, 9,* 497–517.

Frearson, W.M., & Eysenck, H.J. (1986). Intelligence, reaction time (RT) and a new "odd-man-out" RT paradigm. *Personality and Individual Differences, 7,* 807–817.

Gilbert, J.A. (1894). Researches on the mental and physical development of school children. *Studies from the Yale Psychology Laboratory, 2,* 40–100.

Gulliksen, H. (1950). *Theory of mental tests.* New York: Wiley.

Hotelling, H. (1940). The selection of variates for use in prediction, with some comments on the general problem of nuisance parameters. *Annals of Mathematical Statistics, 11,* 271–283.

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

Jensen, A.R. (1985). Methodological and statistical techniques for the chronometric study of mental abilities. In C.R. Reynolds & V.L. Willson (Eds.), *Methodological and statistical advances in the study of individual differences.* New York: Plenum.

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

Jensen, A.R. (1988). Speed of information processing and population differences. In S.H. Irvine & J.W. Berry (Eds.), *Human abilities in cultural context.* New York: Cambridge University Press.

Jensen, A.R., Saccuzzo, D.P., & Larson, G.E. (1988). Equating the Standard and Advanced forms of the Raven Progressive Matrices. *Educational and Psychological Measurement, 48,* 1091–1095.

Luce, R.D. (1986). *Response times: Their role in inferring elementary mental organization.* New York: Oxford University Press.

McNemar, Q. (1949). *Psychological statistics.* New York: Wiley.

Reed, T.E. (1988). A neurophysiological basis for the heritability of vertebrate intelligence. In H.J. Jerison & I. Jerison (Eds.), *Intelligience and evolutionary biology.* Heidelberg: Springer-Verlag.

Rossini, P.M., Marciani, M.G., Caramia, M., Hassan, N.F., & Cracco, R.Q. (1986). Transcutaneous stimulation of motor cerebral cortex and spine: Noninvasive evaluation of central efferent transmission in normal subjects and patients with multiple sclerosis. In R.Q. Cracco & I. Bodis-Wollner (Eds.), *Evoked potentials: Frontiers of clinical neuroscience* (Vol. 3). New York: Alan R. Liss.