INTELLIGENCE **23**, 105–132 (1996)

Intelligence and the Differentiation Hypothesis

IAN J. DEARY
VINCENT EGAN
University of Edinburgh

GAVIN J. GIBSON
ELIZABETH J. AUSTIN
Biomathematics and Statistics, Scotland

CHRISTOPHER R. BRAND
University of Edinburgh

THOMAS KELLAGHAN
St. Patrick's College, Dublin

General intelligence (Spearman's *g*) accounts for over 50% of the reliable variance in a battery of mental tests in a sample of the general population. In a "differentiation hypothesis" originally suggested by Spearman it is hypothesized that the degree to which *g* pervades performance on mental tests is greater at lower ability levels. In addition to providing a critical review, the study presented here tests the differentiation hypothesis: (a) at different ability levels and ages; (b) when groups are selected on the basis of a wide range of criterion abilities; and (c) by developing new statistical techniques for sampling groups of different ability levels. Data used were the Differential Aptitude Test results of over 10,500 Irish schoolchildren aged 14 through 17 years. Of groups selected on the basis of verbal, numerical, or spatial ability, the below-average ability groups had a more pervasive *g* factor, confirming the differentiation hypothesis.

The term *general intelligence*, or *g*, refers to the common variance shared by a battery of mental tests. It is often expressed as the variance accounted for by the first unrotated principal component in the analysis of the intercorrelations of a battery of mental tests. First described by Spearman (1904), *g* has proved to be a durable phenomenon, and compatible with many different theories of intelligence (Carroll, 1993). A major development in theorizing about *g* is the "differentiation hypothesis," according to which the pervasiveness of *g* on abilities

We thank Claudia Pagliari and Mike Anderson for assistance with our coverage of the literature.

Correspondence and requests for reprints should be sent to Ian J. Deary, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ Scotland, U.K. E-mail: <I.Deary@ed.ac.uk>

might vary with IQ level and/or age (Spearman, 1926, 1927; Detterman & Daniel, 1989; Deary and Pagliari, 1991). If correct, the differentiation hypothesis would in part explain why the testing of single IQ and *g* factors has often seemed an unduly limited way of assessing intellectual abilities of people of above-average IQ.

DIFFERENTIATION BY ABILITY LEVEL AND AGE

The notion that the influence of general mental ability might vary with age and ability level occurred to Spearman (1926, 1927), and Burt (1954) claimed to have demonstrated age-related intelligence differentiation by 1919. Deary and Pagliari (1991) described Spearman's (1926, 1927) evidence for the differentiation hypothesis, or the "law of diminishing returns," as he dubbed it. Drawing on others' empirical data, Spearman surmised that, at lower ability levels and at younger ages, *g* accounted for a higher percentage of the variance in a battery of mental tests. Garrett, Bryan, and Perl (1935) found that the first unrotated factor accounted for, respectively, 31%, 32%, and 12% of the variance in a 10-test battery for boys aged 9, 12, and 15 years (and 31%, 24%, and 19% for girls of the same ages). Garrett (1946)—who coined the term *differentiation hypothesis*—stated that "with increasing age there appears to be a gradual breakdown of an amorphous general ability into a group of fairly distinct aptitudes" (p. 375). Filella (1960) found that, among 18-year-old private academic, public academic, and technical school boys, the private academic boys showed the least common variance on subtests from the Differential Aptitude Tests and, by inference, the most differentiation. Filella hypothesized that high socioeconomic status led to greater ability differentiation.

Lienert and Faber (1963) used the term *differentiation* to refer to mental abilities differentiating (and the *g* factor accounting for less variance) as children increase in chronological age, and the term *divergence* to refer to lower correlations between mental abilities found among testees with higher levels of overall IQ. Lienert and Faber reported little support for differentiation, but strong support for divergence. On the German HAWIK test battery children of IQ > 109 showed a smaller first centroid factor than did children of IQ < 91. Lienert and Crott (1964) extended the study of differentiation by testing adolescents (age 10–12 years), young adults (18–20), and older adults (45–60) on 14 ability tests; the percentage variance in the first centroid factor was 45, 41, and 47, respectively. Similarly, using Wechsler normative data, Balinsky (1941) found that differentiation increased from adolescence to early adulthood and then reversed with later adulthood. Moreover, in a longitudinal study, McHugh and Owens (1954) found that the first unrotated principal component accounted for 53.0% of the variance in the Army Alpha Test at age 19, increasing to 63.4% at age 50.

Lienert and Crott (1964) attempted to integrate both divergence and differentiation of mental abilities,

> In the same way as the improvement in performance from childhood to adolescence is accompanied by a differentiation of the underlying ability structure, and as the decline of performance from adolescence to adulthood is accompanied by an integration of structure, a divergence and a convergence of the ability structure takes place from low to high performance level and vice versa respectively. Therefore, it may be assumed that performance level and degree of differentiation are interdependent, and that—regardless of the subject's age—differences in intelligence level presuppose variations in the structure of intelligence in accordance with the divergence hypothesis. (p. 158)

Lienert and Crott (1964) reported that intertest correlations are stronger when performance is impaired by drugs. Lienert (1963) reported also that there is less differentiation of ability among high neuroticism testees (see also Eysenck, 1994; Eysenck & White, 1964).

Anastasi (1970) summarized differentiation studies from the 1950s and 1960s and showed that differentiation was a principal concern in the field of intelligence prior to 1970. The subsequent low profile of this area of research is puzzling. Detterman and Daniel (1989) stated:

> During the 85-year history of this work, it was thought that the positive manifold was uniformly distributed over the full range of ability. That is, it was assumed that the correlation among mental tests would be about the same in a group of low IQ subjects as it would be in a group of high IQ subjects. (pp. 349–350)

Compare this with Filella's (1960) comment on reviewing the research that had taken place on the differentiation of intelligence in the 1940s and 1950s: 'A few decades ago it was thought that the way in which abilities were organized was essentially the same for all individuals and remained constant throughout the individual's life' (p. 119).

RECENT DIFFERENTIATION STUDIES

Carroll (1993) found little evidence for differentiation of intelligence: his reanalysis of the British Ability Scales found no differences in the structure of abilities from age 5 to age 16. Large-scale studies that had found evidence for the differentiation hypothesis (e.g., Atkin et al., 1977) tended to employ school achievement tests or measures of specific knowledge in particular domains. In contrast, Detterman and Daniel (1989) showed that correlations among mental tests were stronger at lower IQ levels, suggesting that *g* exerts a more powerful influence on individual differences across lower levels in the IQ range. In an analysis of the WAIS-R standardization sample, they found that average intersubtest correlations at five levels of IQ were .56 (IQ < 78), .37 (IQ 78–92), .30 (IQ 93–107), .25 (IQ 108–122), and .26 (IQ > 122). An analysis of the WISC-R standardization sample found correlations of .42, .29, .26, .21, and .22 for the five IQ bands, from low to high.

Lynn (1992) examined the Scottish standardization sample of the WISC-R, using similar IQ bands to those employed by Detterman and Daniel (1989) and found mean intersubtest correlations across the five IQ bands of .44, .38, .17, .14, and .20. Therefore, from large data samples, there would appear to be confirmation of the differentiation hypothesis: *g* appears to exert a greater influence on abilities at lower IQ levels. The importance of such an effect for the theory of *g* is outlined by Detterman and Daniel (1989), Lynn (1992), and Anderson (1992).

The possibility exists, therefore, that mental abilities and attainments are differentially accounted for by *g* at different levels of ability and at different ages. The evidence for age differentiation is especially equivocal. Because the structure of intelligence is central to research on the causes and development of ability, and because of the long (though sporadic) history of differentiation ideas, it is important to attempt to confirm or refute the differentiation hypothesis. This is underlined by the fact that the differentiation of ability across IQ levels and ages is important to some theories of intelligence and personality (Anderson, 1992; Brand, 1984; Brand, Egan, & Deary, 1993).

TESTING THE DIFFERENTIATION HYPOTHESIS

Testing the differentiation hypothesis poses a number of methodological difficulties (Burt, 1954), and we now discuss the approach taken in the study presented here (see also Detterman, 1991; Detterman and Daniel, 1989; Deary & Pagliari, 1991; Garrett, 1946).

First, because the key indices to be compared are the correlations of mental test scores across different samples within a population, it is important to ensure that the samples do not differ on parameters that might influence the sizes of the correlations. Samples being compared should be similar with respect to their (a) spread of ability on the mental test scores that are being used to select the samples, (b) spread of age, and (c) reliability on all subtests. Detterman and Daniel (1989) used disattenuated correlations to equalize the spread of ability across samples and age-corrected scores to control for age. However, it has been shown that corrections made to correlations to compensate for restriction of range are likely to overestimate "true" values and to have different effects at different IQ levels (Frearson, Barrett, & Eysenck, 1988). This is because the assumption that underlies "correction" is that the relationship expressed by the correlations in a sample of restricted ranges exists equally across the entire range. It is precisely this assumption that the differentiation hypothesis questions, however. Therefore, "corrections" to correlations might distort the within-sample, intersubtest correlations, which are the key data. In the study presented here, the samples were chosen to be very similar for their spreads of age and of mental ability with respect to the test used to select them, obviating the need to correct samples' correlations for restriction of range.

Second, even if the differentiation hypothesis holds, its effects might not be equally apparent over the full IQ range. For instance, the Detterman and Daniel (1989) study found the greatest effect of *g* in the lowest IQ group and less evidence of increasing differentiation within the nonhandicapped IQ range. It will be of interest to discover whether there is evidence of mental differentiation across groups that are widely separated in intelligence, yet within the normal range. This approach has the additional advantage of focusing the search for differentiation within groups that are not so likely to be affected by ceiling or floor effects in test scores. Therefore, for the purposes of this study, it was decided to study two groups whose IQs were around 90 and 110, each with about half the population standard deviation.

Third, Detterman and Daniel (1989) described in some detail the statistical problems associated with selecting samples from the IQ range in order to assess the pervasiveness of *g* within each. If subgroups of participants within a sample are selected for low or high overall scores, this procedure induces lower intersubtest correlations. Because of this, Detterman and Daniel selected subgroups of testees on a single subtest and examined the within-group intersubtest correlations for each subgroup on the entire test battery, including the subtest that was used for selection. In this study a similar but modified approach to sample selection was taken. Samples of lower and higher ability participants were selected from the population on the basis of single subtests, and the variance accounted for by *g* was computed on the basis of the correlations among all the other subtests. This was repeated for each subtest with attention to each test's own *g* loading.

Fourth, the approach taken in this study is demanding in terms of the number of participants required. The aim is to examine four groups of children (high age–high IQ, high age–low IQ, low age–high IQ, and low age–low IQ) and to examine intersubtest correlations to assess the percentage of variance attributable to *g*. To meet the methodological concerns already noted, it was necessary for each of the four groups to contain hundreds of participants and have very similar bivariate normal distributions for age and selection-criterion ability test scores. Because the mean scores of each group would be removed by 10 IQ points from the center of the normal distribution, the majority in any population sample would remain unused in this design. Therefore, it was essential to begin such a study with a sample larger than 10,000 testees, all of whom had been tested on a large, well-recognized battery of mental tests.

Finally, it was considered important to test participants on a test battery that measured diverse mental abilities. Moreover, it was preferable to choose a battery within which each ability was assessed by a substantial test assessing a theoretically important ability. For the study presented here, the Differential Aptitude Tests were chosen. The subtests of this battery largely correspond to important second-order mental ability factors (Carroll, 1993), and, with two exceptions, the individual subtests last from 25 to 30 min, and all have reliabilities (Kuder-Richardson$_{20}$) in excess of .80.

METHOD

Participants

A representative sample of all second-level schools in the Republic of Ireland was selected for the standardization of the British–Irish version of the Differential Aptitude Tests, Form T. A total of 819 schools in the country attended by 272,000 students were stratified by type (secondary, vocational), size (<319 students, ≥319 students), and sex served by the school (boys, girls, or mixed). Within strata, schools were selected randomly. Altogether there were 53 schools in the sample. Students selected for testing were in their second, third, fourth, and fifth years of second-level education (Grades 8–11, normal ages 14–17 years, respectively). All students in the relevant age grades who were present on the day of testing took the test.

There were 3094 Grade 8 students with a mean age of 168.8 months (*SD*, 8.9), 2708 Grade 9 students with a mean age of 181.0 months (*SD*, 14.8), 2715 Grade 10 students with a mean age of 192.4 months (*SD*, 7.5), and 2528 Grade 11 students with a mean age of 204.7 months (*SD*, 10.6). From this total of 11,045 participants those with any missing data were excluded, to leave a total pool of 10,535 participants.

Differential Aptitude (DAT) Tests Battery

The test used was the American Differential Aptitude Tests (DAT), Form T, adapted for use in Britain and Ireland. The British–Irish version is made up of the same content areas as the American version. It differs from the original American version in that alterations were made to individual words and phraseology (Educational Research Centre, 1986). The DAT consists of eight varied tests of mental ability as follows.

Verbal Reasoning. This is a verbal analogies test that lasts 30 min. Items are in the form, "A is to B as C is to D." Participants are given B and C and are required to indicate which of a number of answer options best completes the analogy by providing A and D.

Numerical Ability. This is a test of arithmetic computation that lasts 30 min.

Abstract Reasoning. This test involves the testee identifying the underlying principle of the progression in a series of abstract figures. The task is to select the next member of the series, given the application of the rule. The test lasts 25 min.

Clerical Speed and Accuracy. This subtest requires the participant to search at speed for certain letter and letter–number combinations. It has two subtests each lasting 3 min.

Mechanical Reasoning. In this subtest participants must study a pictorial representation of a mechanical principle in action and answer a question with respect to the operation of the apparatus. The situations represented in the items include levers, pulleys, and gears; however, the test manual states that the principles on which the questions are based do not require specialist knowledge of physics. This subtest lasts 30 min.

Space Relations. In this subtest, which lasts 25 min, participants are required to imagine the shape that would result if an unfolded figure were folded. In addition to imagining the fully constructed form of the "flattened" pattern, an ability to visualize how 3-D objects appear in different rotations is required.

Spelling. This subtest requires testees to indicate whether each of 100 words is spelled correctly or wrongly. The subtest lasts 10 min.

Language Usage. This subtest requires the participant to detect errors of grammar, capitalization, and punctuation in 50 sentences. This subtest lasts 25 min.

Procedure

All participants undertook the DAT within school classrooms under the conditions strictly prescribed in the test manual. The tests were administered by school counsellors or teachers. Answer sheets were returned to a research agency where they were machine scored.

Statistical Procedures for Subgroup Selection

As already stated, the intention was to extract from this very large sample of participants subgroups of different ages and abilities such that each subgroup had near identical bivariate normal distributions of age (at about 170 or 201 months) and criterion ability (equivalent to about IQ 90 or 110). There are no readily available techniques to perform such a selection, and a statistical program was specially designed and written for the study. The technical aspects of subgroup selection are now described.

We denote the entire data set by S, where

$$S = \{x_i \mid 1 \leq i \leq 10{,}535\} \tag{1}$$

where each x_i is a vector consisting of nine components. For a data vector x, the first component x_0 denotes the age of a testee in months, and the components

$x_1, \ldots, x_8$ record the scores of that testee in each of eight DAT subtests. For each subtest, j, $1 \leq j \leq 8$, the following sampling procedure was carried out. The mean and covariance matrix of $(x_0\, x_j)$ across the whole data set S was calculated. We denote these quantities by $\mu = (\mu_0, \mu_j)$ and

$$W_j = \begin{pmatrix} W_{00} & W_{0j} \\ W_{0j} & W_{jj} \end{pmatrix} \tag{2}$$

respectively. The purpose of the sampling is to extract four subgroups from S within which the respective distributions of the vector (x_0, x_j) represent the four possible combinations of high and low age and high and low scores on subtest j. Specifically, this was effected by identifying "ideal" distributions for each of the four subgroups (high age–high ability, high age–low ability, low age–high ability, and low age–low ability). For the two groups in which age was "high" the mean age was chosen to be

$$\mu_0^{\text{high}} = \mu_0 + W_{00}^{1/2} \tag{3}$$

(i.e., one standard deviation above the mean age of S). Similarly, for the "low-age" groups the mean age was chosen to be

$$\mu_0^{\text{low}} = \mu_0 - W_{00}^{1/2}. \tag{4}$$

Mean values for subtest scores within the four groups were chosen by first calculating the mean and variance of the score, x_j, *conditional* on x_0 being equal to the mean age for that group [assuming that (x_0, x_j) is distributed in S according to a bivariate normal distribution] and setting the mean of the subtest score to be respectively one standard deviation above and below this conditional mean for "high-ability" and "low-ability" groups. For the case of the "low-age" groups we used standard results on the bivariate normal distribution to calculate the mean and the variance of x_j, conditional on $x_0 = \mu_0 - W_{00}^{1/2}$, to be $\mu_j - \rho W_{jj}^{1/2}$ and $(1 - \rho^2)W_{jj}$ respectively, where

$$\rho = \frac{W_{0j}}{W_{00}^{1/2} W_{jj}^{1/2}}. \tag{5}$$

Hence the score means for the two groups are

$$\mu_j^{low,high} = \mu_j - \rho W_{jj}^{1/2} + (1 - \rho^2)^{1/2} W_{jj}^{1/2} \tag{6}$$

for the "high-ability" group, and

$$\mu_j^{low,low} = \mu_j - \rho W_{jj}^{1/2} - (1 - \rho^2)^{1/2} W_{jj}^{1/2} \tag{7}$$

for the "low-ability" group.

Similarly, mean scores within the "high-age" groups were calculated as

$$\mu_j^{high,high} = \mu_j + \rho W_{jj}^{1/2} + (1 - \rho^2)^{1/2} W_{jj}^{1/2} \tag{8}$$

and

$$\mu_j^{high,low} = \mu_j + \rho W_{jj}^{1/2} - (1 - \rho^2)^{1/2} W_{jj}^{1/2} \tag{9}$$

for the "high-ability" and "low-ability" groups, respectively. The location of mean age and score for each of the four groups (low age–high ability, low age–low ability, high age–high ability, and high age–low ability) is illustrated in Figure 1.

For each of the four groups, the covariance matrix of (x_0, x_j) was chosen to be diagonal (i.e., age and ability designed to be uncorrelated within any group) and invariant among groups. The standard deviation of age within any group was chosen to be one half that of the standard deviation of age in S. The standard deviation of subtest score within any group was selected to be one half of the

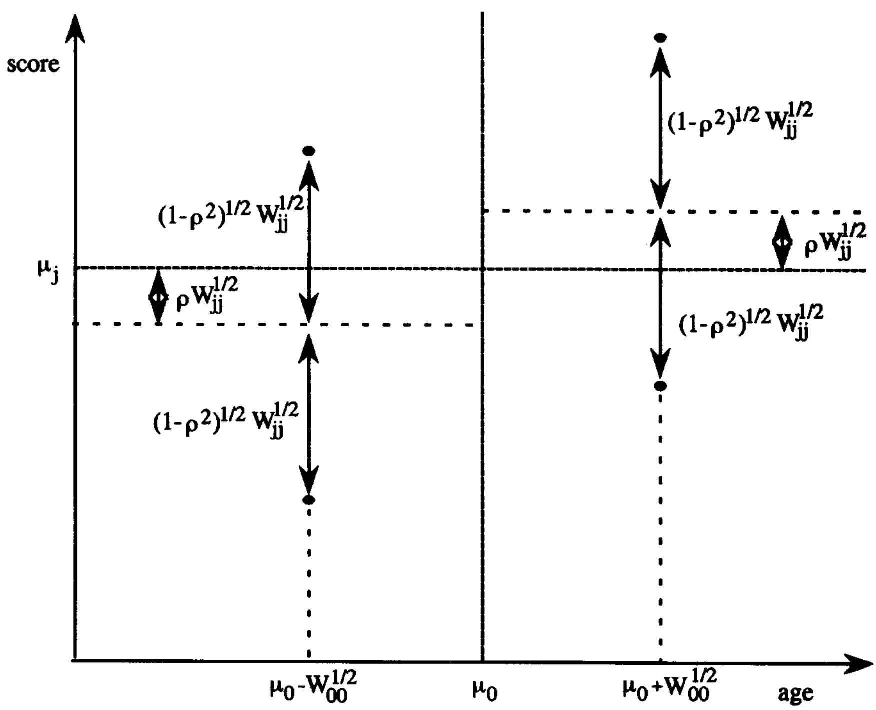


Figure 1. Location of mean age and ability subtest score for each of the four groups as defined in the section on statistical procedures for subgroup selection.

standard deviation of subtest score in S conditional on a fixed value of age (appealing to assumptions of normality). Formally, the covariance, V, of (x_0, x_j) within any group was given by

$$\mathrm{V} = \begin{pmatrix} W_{00}/4 & 0 \\ 0 & W_{jj}(1 - \rho^2)/4 \end{pmatrix}. \tag{10}$$

To generate the four groups (for any given subtest) so that within each group the distributions of (x_0, x_j) had the specified means and covariance matrix V, the following procedure was developed.

(i) Initially all four groups are incomplete (the notion of completeness will be described).
(ii) Select an incomplete group at random, with equal probability for each incomplete group.
(iii) Generate a random vector (y_0, y_j) from a bivariate normal distribution with mean and covariance specific to the selected group, as already described.
(iv) Extract the data vector (i.e., complete data pertaining to an individual participant) from S for which (x_0, x_j) matches (y_0, y_j) most closely. Formally, we select the data vector x, from previously unselected cases, which minimizes the quantity

$$C = \left(\frac{W_{jj}(x_0 - y_0)^2 + W_{00}(x_j - y_j)^2}{W_{00} + W_{jj}} \right). \tag{11}$$

If the minimized value of C is less than some specified tolerance, then x is placed in the selected group, and is thereafter unavailable for reselection. If the tolerance is exceeded, a mismatch for the specified group is recorded. A group is defined as being complete when m mismatches are recorded, where m is an integer specified by the user.
(v) The procedure terminates when all four groups are complete.

We can envisage the qualitative effect of assigning too high a tolerance or too high a value of m in the algorithm described. In such circumstances, group sizes will be larger but at the expense of statistical deviations from the prescribed distributions. Conversely, overly conservative values of tolerance and m will lead to unacceptably small (but perfectly formed!) groups.

Applying the described procedures for each subtest in turn, we constructed, in each case, four disjoint, randomly chosen groups with the following characteristics: mean age and subtest score cover the four permutations of high and low age and ability; the covariance matrix of age and subtest score is the same for all groups; and age and subtest score are uncorrelated within any group. For any subtest, there were inevitably different numbers of participants in each of the four groups, reflecting the nonideal nature of the original data set *S*. As an example of the results of this procedure, Figure 2 shows the result of the group selection exercise when groups were selected on the basis of age and verbal reasoning scores.

Statistical Analyses

Analyses followed to a large extent the procedures of Detterman and Daniel (1989). Using a single subtest score to help select four subgroups on the bases of age and ability, as described, scores on the other seven subtests were correlated. Therefore, departing from Detterman and Daniel's procedure, the subtest used for subgroup selection was *not* included in subsequent analyses. The intercorrelations of the seven subtests for each subgroup were subjected to principal components analysis. The percentage of total variance accounted for by the first unrotated principal component was used to indicate the pervasiveness of *g* in each subgroup. The loadings of each subtest on the first unrotated principal component were noted in order to compare the structure of *g* in the different groups. This procedure was repeated for each of the eight subtests: that is, separate participant selection exercises were undertaken for each subtest, and subsequent

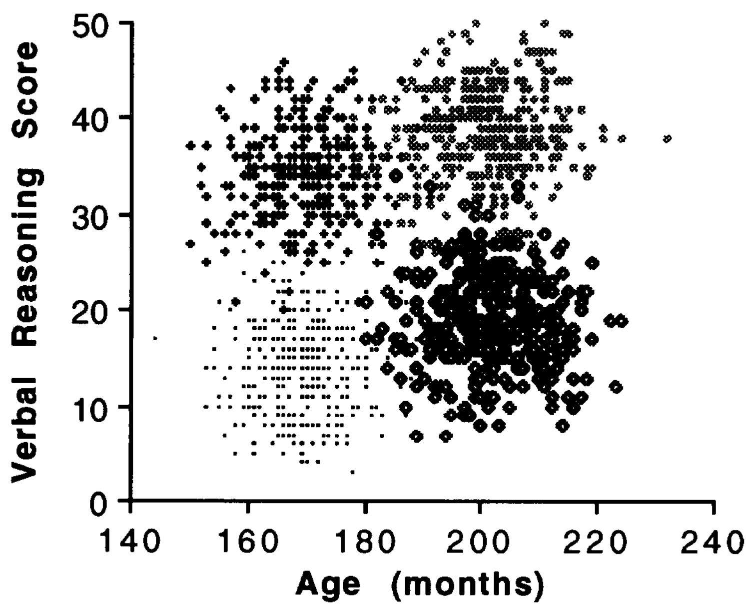


Figure 2. An example of the distribution characteristics of the four groups of participants selected on the basis of age and verbal reasoning scores from the Differential Aptitude Tests. Key to symbols: dots = young–low group; crosses = young–high group; open diamonds = old–low group; and stippled circles = old–high group. Away from the extremes of each distribution each symbol usually represents more than one participant.

analyses were undertaken for each of the eight possible permutations of the remaining subtests. It is important to note, therefore, that the eight analyses reported here (and the eight corresponding appendices) are not based on the same participants; importantly, however, they contain subgroups with highly constrained and similar distribution characteristics.

In addition to reporting the percentage of variance accounted for by the first unrotated principal component, we note also the mean intersubtest correlation $\bar{r}$ for each group of testees in each selection exercise, allowing comparison with the Detterman and Daniel (1989) study. This was done by applying the following formula (Kaiser, 1968),

$$\bar{r} = (\lambda - 1)/(p - 1)$$

where λ is the eigenvalue of the first principal component, and p is the number of variables (which was always seven in our study).

RESULTS

The presentation of results will describe the percentage of variance accounted for by *g* in the different age and ability level groups in each of the analyses undertaken. The results are presented in Appendices 1 to 8, inclusive, which provide the subgroup comparisons based on the group selections for each of the DAT subtests in turn. Each of the appendices has a similar format. The title indicates which of the DAT subtests was used to select the four groups: young–low, young–high, old–low, old–high. (The terms *low* and *high* are merely convenient shorthand used by us to describe the IQ ranges ≤ 90 and ≥ 110, and they are intended to serve as clear and brief indicators of the relative mental ability levels of the groups.) The appendices then show the number of participants that could be placed in each group within the tolerance limits of the statistical selection algorithm. Following this, the percentage of variance accounted for by the first principal component is shown; this is the key result, and these data are collected together for each subtest in Table 1.

Appendices 1 to 8 also show the loadings of each subtest on the first unrotated principal component; this allows the reader to compare the structure of *g* in each analysis. The next two rows of the appendix tables show the means and standard deviations of the ages and selection criterion ability test scores of each group. In almost all cases the selection algorithm constructed large groups of participants with very similar distribution characteristics for age and criterion ability. The last row of each appendix table shows an overall IQ score for the groups that were computed from age-corrected scores for all eight subtests of the DAT. Even though it is not normal to calculate an overall score for performance on the DAT subtests, such a score was computed in the analysis presented here to provide an indication of the overall ability levels of the four groups. These data played no

TABLE 1
Percentage of Total Variance in Differential Aptitude Test Battery Performance Explained by the First Unrotated Principal Component

	Age/Ability Group			
Subtest	**Young/Low**	**Young/High**	**Old/Low**	**Old/High**
Verbal reasoning	44.7 (.355)	37.3 (.269)	42.5 (.329)	35.1 (.243)
Numerical ability	47.7 (.390)	40.8 (.309)	44.7 (.355)	41.0 (.311)
Abstract reasoning	44.1 (.347)	49.1 (.406)	44.8 (.365)	50.7 (.425)
Clerical speed and accuracy	61.4 (.550)	63.5 (.574)	60.5 (.540)	57.1 (.499)
Mechanical reasoning	51.8 (.437)	51.2 (.431)	52.2 (.442)	52.6 (.447)
Space relations	52.5 (.446)	45.7 (.366)	53.2 (.454)	46.8 (.379)
Spelling	49.9 (.416)	49.4 (.409)	53.5 (.457)	54.6 (.471)
Language usage	46.3 (.373)	45.6 (.365)	47.9 (.392)	44.8 (.356)
Mean Variance (%)	49.8	47.8	49.9	47.8

Note. Each row identifies the Differential Apptitude Tests subtest that was used to select the four groups; the principal components analysis was performed on the scores of the remaining seven tests. The average intersubtest correlations are shown in parentheses.

part in the analyses and are not used as outcome data; they are included merely to indicate the overall ability levels of the four groups.

Table 1 summarizes the findings of all eight individual analyses. Age did not appear to affect the size of the first unrotated principal component. Overall, the effect of ability level was the same in both the young and the old samples: for higher ability individuals the first unrotated principal component accounted for 47.8% of the variance in tests scores in both the younger and the older groups; and for lower ability individuals the first unrotated principal component accounted for 49.8% and 49.9% of the variance, respectively, in the younger and older groups.

Table 1 reveals that there are relatively large differences in the percentage of variance accounted for by g depending upon which subtest is used to select the participants. Verbal reasoning has a very high g loading whenever it is included in a principal components analysis (see Appendices 2 to 8). Therefore, when it is used to select participants, the percentage of variance explained in g in each of the four age and ability level groups is relatively low (Table 1 and Appendix 1). In contrast, it may be seen that clerical speed and accuracy has a low g loading whenever it appears in a principal components analysis (see Appendices 1 to 3 and 5 to 8). Therefore, when it is used to select participants, the percentage of variance accounted for by g in each of the four groups is relatively high (Table 1 and Appendix 4). Whereas these results are predictable from the characteristics of the subtests used to make the group selections, and their effects are detectable across all groups, more detailed effects of different selection exercises on individual groups are detectable. Table 1 shows that, when participant group selections are based on verbal reasoning, numerical ability, or space relations, the differences in the percentage of *g* in the high- and low-ability groups are especially marked. If the mean effect of these three tests is taken, and if the old and young groups are combined, the percentage of variance accounted for by *g* is 6.5% greater in the lower ability group than in the higher ability group. On the other hand, group selection based on Abstract Reasoning subtest scores leads to the opposite effect: In this case the higher ability participants have a higher suffusion of *g* (Table 1). The remaining subtests have little effect on the *g* suffusion of other tests.

If the high- and low-age groups are combined, and if the group selection is based upon verbal reasoning scores, the average intersubtest correlation for lower ability participants is .342 and that for the higher ability participants is .256. The corresponding average intersubtest correlations for the selections made on the basis of numerical ability and space relations scores, respectively, were .310 and .372 for the higher ability groups and .372 and .450 for the lower ability groups.

An effect of age is detectable only when group selection is made on the basis of spelling scores. The percentage of variance explained by the first unrotated principal component in the remaining subtests is higher in the older groups, a result that goes against the suggestion that ability differentiates with increasing age (Table 1 and Appendix 7).

An alternative explanation that has been proposed for diminished intersubtest correlations in high-ability groups is that ability tests might have a lower reliability for high- than for low-ability testees (Deary and Pagliari, 1991; Burt, 1954). To test this assumption, the item-level data were obtained for the participants in this study. (A small number of subtest items were missing or unusable. These were mechanical reasoning 53, numerical ability 3, spelling 1 & 3, and space relations 1). Because the results just presented do not show evidence of ability differentiation with increasing age, the effect of age on subtest reliability was not considered.

To compare subtest reliability for high- and low-ability testees, the item data for each subtest were analyzed by generating subgroups of participants from various parts of the total score distribution. The method employed was essentially a one-dimensional analogue of the subgroup selection technique described previously, and it results in a group of participants with total scores normally distributed with chosen values of the mean and standard deviation. The programming method differed slightly from that described, employing standard methods for selecting and transforming distributions using a random number generator (Press, Flannery, Teukolsky and Vetterling, 1986). The measure of reliability used was the split-half, that is, the correlation between total scores on even- and odd-numbered items. Because of this, and because there were some unusable items in some subscales, the absolute levels of reliability given here will be underestimates; however, it is the relative reliability levels that are important with respect to testing the differentiation hypothesis.

Reliability coefficients show strong range effects; to present this variation and any possible difference between high- and low-ability groups, reliability was plotted against standard deviation using different symbols for groups with means above and below the global mean subtest score. For completeness and comparison, an average ability group was added to the plots. These plots, displayed in Appendices 9 to 16, show that subtest reliability is not significantly different between high- and low-ability groups for clerical speed and accuracy and verbal reasoning. For abstract reasoning, language usage, mechanical reasoning, space relations, and spelling there is evidence for slightly higher reliability for high-ability groups. Only for numerical ability is there evidence of slightly reduced reliability for high-ability groups. From this, we conclude that varying subtest reliability cannot explain the observed variation in subtest intercorrelations, because only one subtest has a reliability variation in the appropriate direction. Indeed, higher reliabilities for high-ability groups on five of the eight subtests will work in the opposite direction and make the detection of ability differentiation harder by inducing a small increase in subtest intercorrelations for high-ability groups. For comparison with the reliabilities shown in Appendices 9 to 16, the reliability values for the whole sample are verbal reasoning = 0.87, numerical ability = 0.86, abstract reasoning = 0.87, clerical speed and accuracy = 0.98, mechanical reasoning = 0.78, space relations = 0.85, spelling = 0.93, language usage = 0.82.

DISCUSSION

Evidence for Differentiation

A substantial percentage of the variance in human mental ability can be accounted for by a general intelligence factor, often called *g*. However, the percentage of mental ability variance explained by *g* has sometimes been found to vary across different ability and age levels. Although this observation was first made over half a century ago (see Deary & Pagliari, 1991), it is only recently that empirical studies of good quality have addressed this important issue (Detterman & Daniel, 1989) and a theory has been proposed to explain the phenomenon (Anderson, 1992).

The study presented here offers partial support for the notion that the first unrotated principal component accounts for less of the variance in high-ability groups than in low-ability groups. In addition, it was shown that the differentiation effect was not caused by lower subtest score reliability among higher scoring testees; if anything, more able participants tended to show higher reliabilities in most subtests, which may have slightly depressed the extent of the differentiation effect reported here. The overall effect was small: The difference in variance between groups that differ by the equivalent of 20 IQ points is just over 2%. Moreover, across the narrow age range investigated (14.2 years to 16.7 years), this study finds little evidence to suggest that human mental ability is more differentiated in older groups.

However, the more highly *g*-loaded tests did show an effect in the direction predicted by the differentiation hypothesis. This tendency is especially convincing if the results of the selection based on abstract reasoning scores are excluded. Such exclusion can be justified because there were few bright, high-scoring participants on this test (see Appendix 3). From the pool of over 10,000 available participants, 1164 met the statistical criteria for the lower ability groups, whereas only 303 met the criteria for the higher ability groups. Such unusual disparity in available participants between the higher and lower ability groups suggests that participants tend to bunch at relatively low scores on abstract reasoning. This skewing of scores suggests that, on this subtest, the relatively rare high scores are more meaningful with respect to reliable individual differences and thus still strongly *g*-loaded. This is supported by the fact that, in the only three instances where participants were relatively scarce (the young–low group on numerical ability selection, the old–high group on abstract reasoning selection, and the old–high group on spelling selection) the first unrotated principal component accounts for a relatively large percentage of the mental test variance in the relevant group.

The reason for the overall small size of the effect in the study presented here, by comparison with the studies of Detterman and Daniel (1989) and Lynn (1992), might be the range of ability used here. Their strongest effects tended to be in the lowest IQ groups: although they found small systematic results indicating that *g* was less pervasive as ability level increased, they found disproportionately large effects of *g* in the lowest IQ groups whose participants were below IQ 80, and many of whom would be classified as mentally handicapped. However, previous research and this report agree that there is a detectable differentiation effect across ability levels within the normal range of IQ.

A second possible explanation for the small overall effect of differentiation found here is that the magnitude of the effect varies depending on which ability score is used to make the participant group selections. Take the example of the selection that was based on verbal reasoning, which is comparable to the Detterman and Daniel (1989) analysis in which selections were based on vocabulary and information scores. If the results of the high and low ages are combined, the average intersubtest correlation for lower ability participants is .330, and that for the higher ability participants is .246. These results are similar to Detterman and Daniel's (1989) results for adult WAIS-R data where the intersubtest correlations for groups with similar IQ to ranges to those tested here were: IQ 78–92 = .37, IQ 93–107 = .30, and IQ 108–122 = .25. Therefore, our results, using a different ability test battery and different participant group selection criteria, closely support earlier findings with respect to ability differentiation when selections of subgroups are made on the basis of verbal ability. Similar results were found when the group selections were made on the basis of numerical ability or space relations scores. Selections based on other subtest scores showed very little differentiation effect, and the results of the selection based on abstract reasoning scores revealed an effect in the opposite direction, that is, the higher ability participants intersubtest correlations were higher and the effect of *g* was stronger in that group.

It appears clear that there is a marked increase in the effect of *g* in lower ability groups when these groups are defined in terms of some abilities, that is, vocabulary (WAIS-R or WISC-R), verbal reasoning (DAT), and information (WAIS-R or WISC-R). In this study, such an effect was found for participant group selections based on verbal reasoning (DAT), numerical ability (DAT), and space relations (DAT). Other DAT subtests failed to demonstrate the differentiation effect. Whether the differences between subtests are lawful or the result of random error must await further research. Rather than the differentiation effect being only a phenomenon that distinguishes very low ability groups from others, the study presented here establishes that it is detectable across the normal range of *g* so long as sensitive and normally distributed measures of *g* are used.

Mechanisms of Differentiation

In mechanistic terms, how might differentiation of ability operate? Anderson's (1992) computer-oriented theory is couched in cognitive terms. Spearman's theory to account for the differentiation of intelligence is based on the performance of engines and couched in biological terms (Deary & Pagliari, 1991). Despite these different frames of reference, the two are very similar. It is difficult to distinguish Anderson's Basic Processing Mechanism (BPM) from Spearman's *g*, and his Specific Processors (SP) bear a resemblance to the major group factors in human intelligence. For Spearman, *g* was energy that fuelled different engines, that is, different mental abilities. The positive manifold arose because all abilities needed some of this general energy. Anderson explained the positive manifold by setting the BPM's processing speed as a limitation that applies to all specific mental abilities. Spearman's explanation of the differentiation of intelligence at higher ability levels was based on the fact that doubling the fuel in an engine's boiler does not cause a doubling of the engine's speed. Anderson explained differentiation by assuming that, the slower the speed (or efficiency) of the BPM, the more it constrains and dominates the output of the specific processors.

Detterman's (1993) explanation of intelligence differentiation resembles Spearman's (1926, 1927) and Anderson's (1992) accounts:

> The differences in correlation across IQ level are consistent with the systems theory of intelligence . . . In this theory, lower IQs result from deficits in important cognitive processes. Because these processes are part of a system, they affect the functioning of other parts of the system. If these important processes are deficient, the

> parts dependent upon them will also be impaired. Essentially, a deficit in an important process will put an upper ceiling or limit on the efficiency of the operation of other parts of the system. That means that all parts of the system will be more similar if there is a deficiency in an important process. This forced similarity in abilities is what causes a higher correlation among IQ subtests for low IQ subjects. (pp. 27–28)

Therefore, common to all of these accounts is the idea that higher correlations among different ability tests might arise where the efficiency of some general cognitive process is relatively low. At higher levels of general processing efficiency, scores on a given task are more strongly influenced by the efficiency of the specific processes required to perform the task.

A different account of the differentiation hypothesis uses an economic metaphor and construes intelligence in terms of resources for investment (see Brand, 1984). In this view intelligence is spent and invested like money: at low levels of income, increments are predictably directed toward housing, food, clothing, and the care of children. At higher levels of income spending is more differentiated: disposable income may be directed at a near-infinite range of targets. At higher levels of intelligence, ability is probably more directed by interest, motivation, and choice. Personality traits, too, have been thought to show more extremes and more dimensions at higher levels of *g* (Brand, Egan, & Deary, 1994).

Insofar as differentiation occurs according to IQ rather than chronological age our results provide more support for the "cognitive architecture" than for the "economic investment" model. However, the age range in this study is narrow, and it is one across which inspection time, the variable that is chiefly responsible for individual differences in Anderson's (1992) BPM efficiency, does not change (Nettelbeck & Willson, 1985). The accounts of differentiation mechanisms offered above are largely metaphorical; now that the differentiation effect has been established it will be necessary to devise studies to test specific hypotheses about the causes of differentiation.

Another hypothesis about differentiation which will require further research is that the genetic contribution to intelligence is higher at low-ability levels (Detterman, Thompson, & Plomin, 1990), though not all studies are agreed upon this finding (Bailey & Revelle, 1991; Cherny, Cardon, Fulker, & DeFries, 1992). Other research suggests a larger influence of shared family environment at lower IQ levels, an effect that is probably stronger and more stable than any differences in heritability (Thompson, Detterman, & Plomin, 1993).

In conclusion, this report has critically reviewed and partly confirmed the differentiation hypothesis about human intelligence. When ability levels are measured by verbal tests, the mental abilities of higher ability subjects are more differentiated, and the effect of general intelligence or Spearman's *g* is less pervasive. The extent of this effect—comparing the percentage variance accounted for by the first unrotated principal component in groups of average IQ 90 and 110—

can be as great as 7%. However, when selection is made on the basis of some other types of ability, the effect is not seen, and in one case it is reversed. As so often in mental measurement, much depends on whether *g* itself provides the basis on which ability distinctions are made.

REFERENCES

Anastasi, A. (1970). On the formation of psychological traits. *American Psychologist, 25,* 899–910.

Anderson, M. (1992). *Intelligence and development: A cognitive theory.* Oxford, England: Blackwell.

Atkin, R., Bray, R., Davison, M., Herzberger, S., Humphreys, L., & Selzer, U. (1977). Ability factor differentiation grades 5 through 11. *Applied Psychological Measurement, 1,* 65–76.

Bailey, M.J., & Revelle, W. (1991). Increased heritability for lower IQ levels? *Behavior Genetics, 21,* 397–404.

Balinsky, R. (1941). An analysis of the mental factors of various age groups from nine to sixty. *Genetic Psychology Monographs, 23,* 191–234.

Brand, C.R. (1984). Intelligence and inspection time: An ontogenic relationship? In C.J. Turner & H.B. Miles (Eds.), *The biology of human intelligence.* Humberside, England: Nafferton.

Brand, C.R., Egan, V., & Deary, I.J. (1993). Personality and general intelligence. In G.L. van Heck, P. Bonaiuto, I.J. Deary & W. Nowack (Eds.), *Personality Psychology in Europe. Volume 4.* Tilburg, The Netherlands: Tilburg University Press.

Brand, C.R., Egan, V., & Deary, I.J. (1994). Intelligence, personality and society: Constructivist versus essentialist possibilities. In D.K. Detterman (Ed), *Current topics in human intelligence 4.* Norwood, NJ: Ablex.

Burt, C. (1954). The differentiation of intellectual ability. *British Journal of Educational Psychology, 24,* 76–90.

Carroll, J.B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* Cambridge: Cambridge University Press.

Cherny, S.S., Cardon, L.R., Fulker, D.W., & DeFries, J.C. (1992). Differential heritability across levels of cognitive ability. *Behavior Genetics, 22,* 153–162.

Deary, I.J., & Pagliari, C. (1991). The strength of *g* at different levels of ability: Have Detterman and Daniel rediscovered Spearman's "law of diminishing returns." *Intelligence, 15,* 247–250.

Detterman, D.K. (1991). Reply to Deary and Pagliari: Is *g* intelligence or stupidity? *Intelligence, 15,* 251–255.

Detterman, D.K. (1993). Giftedness and intelligence: One and the same? In *Ciba Symposium 178: The origins and development of high ability.* Chichester: Wiley-Interscience.

Detterman, D.K., & Daniel, M.H. (1989). Correlations of mental tests with each other and with cognitive variables are highest for low-IQ groups. *Intelligence, 13,* 349–359.

Detterman, D.K., Thompson, L.A., & Plomin, R. (1990). Differences in heritability across groups differing in ability. *Behavior Genetics, 20,* 369–384.

Educational Research Centre. (1986). *Differential Aptitude Tests. Form T, Manual.* Dublin: Educational Research Centre.

Eysenck, H.J. (1994). Personality and intelligence. In R.J. Sternberg & P. Ruzgis (Eds.), *Personality and Intelligence.* Cambridge: Cambridge University Press.

Eysenck, H.J., & White, P.O. (1964). Personality and the measurement of intelligence. *British Journal of Educational Psychology, 34,* 197–202.

Filella, J.F. (1960). Educational and sex differences in the organisation of abilities in technical and academic students in Colombia, South America. *Genetic Psychology Monographs, 61,* 115–163.

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

Garrett, H.E. (1946). A developmental theory of intelligence. *American Psychologist, 1,* 372–378.

Garrett, H.E., Bryan, A.I., & Perl, R. (1935). The age factor in mental organisation. *Archives of Psychology,* No. 176.

Kaiser, H.F. (1968). A measure of the average intercorrelation. *Educational and Psychological Measurement, 28,* 245–247.

Lienert, G.A. (1963). Die Factoren struktur der Intelligenz als Funktion des Neurotizismus. *Zeitschrift fuer experimentalle und angerwandte Psychologie, 10,* 140–159.

Lienert, G.A., & Crott, H.W. (1964). Studies on the factor structure of intelligence in children, adolescents and adults. *Vita Humana,m 7,* 147–163.

Lienert, G.A., & Faber, C. (1963). Über die Faktorenstruktur des HAWIK auf verschiedenen Alters- und Intelligenzniveaus [On the factor structure of the HAWIK at various levels of age and intelligence]. *Diagnostica, 9,* 3–11.

Lynn, R. (1992). Does Spearman's *g* decline at high IQ levels? Some evidence from Scotland. *Journal of Genetic Psychology, 153,* 229–230.

McHugh, R.B., & Owens, W.A. (1954). Age changes in mental organisation—a longitudinal study. *Journal of Gerontology, 9,* 296–302.

Nettelbeck, T., & Willson, C. (1985). A cross-sequential analysis of developmental differences in speed of visual information processing. *Journal of Experimental Child Psychology, 40,* 1–22.

Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T. (1986). *Numerical recipes.* Cambridge, U.K.: Cambridge University Press.

Spearman, C. (1904). "General intelligence" objectively determined and measured. *American Journal of Psychology, 15,* 201–293.

Spearman, C. (1936). Some issues in the theory of "G" (including the law of diminishing returns). (Address to the British Association Section J—Psychology, Southampton, 1925. Bound in Collected Papers, Psychological Laboratory, University College, London)

Spearman, C. (1927). *The abilities of man.* London: Macmillan.

Thompson, L.A., Detterman, D.K., & Plomin, R. (1993). Differences in heritability across groups differing in ability, revisited. *Behavior Genetics, 23,* 331–336.

APPENDIX 1

Results of Group Selection Based on Verbal Reasoning

	Age/Ability Group			
	Young/Low	Young/High	Old/Low	Old/High
Participants (n)	382	353	370	366
Variance in 1st Unrotated Principal Component (%)	44.7	37.3	42.5	35.1
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Numerical ability	.765	.645	.720	.698
Abstract reasoning	.776	.643	.765	.718
Clerical speed and accuracy	.329	.454	.531	.295
Mechanical reasoning	.730	.545	.654	.543
Space relations	.620	.672	.620	.700
Spelling	.619	.601	.588	.504
Language usage	.729	.686	.656	.572
Mean age (SD) in months	169.8 (7.3)	169.9 (7.8)	201.8 (8.3)	200.8 (8.3)
Mean (SD) of raw score of test used for group selection	14.1 (4.7)	34.2 (4.7)	19.1 (5.0)	38.6 (4.6)
Mean (SD) IQ	90.4 (8.5)	110.3 (7.6)	89.8 (8.1)	107.3 (7.6)

APPENDIX 2

Results of Group Selection Based on Numerical Ability

	Age/Ability Group			
	Young/Low	Young/High	Old/Low	Old/High
Participants (n)	70	739	1106	950
Variance in 1st Unrotated Principal Component (%)	47.7	40.8	44.7	41.0
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.853	.784	.824	.814
Abstract reasoning	.743	.629	.778	.697
Clerical speed and accuracy	.419	.305	.411	.192
Mechanical reasoning	.718	.568	.581	.634
Space relations	.662	.686	.640	.728
Spelling	.705	.611	.617	.543
Language usage	.658	.766	.742	.676
Mean age (SD) in months	170.0 (6.7)	169.5 (7.7)	201.5 (7.9)	202.0 (7.8)
Mean (SD) of raw score of test used for group selection	7.8 (3.6)	23.4 (4.4)	12.0 (4.2)	28.7 (4.3)
Mean (SD) IQ	91.2 (8.8)	109.6 (8.2)	90.7 (8.4)	105.2 (8.5)

APPENDIX 3

Results of Group Selection Based on Abstract Reasoning

	Age/Ability Group			
	Young/Low	Young/High	Old/Low	Old/High
Participants (*n*)	649	207[a]	515	96[a]
Variance in 1st Unrotated Principal Component (%)	44.1	49.1	44.8	50.7
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.829	.851	.825	.791
Numerical ability	.759	.754	.772	.816
Clerical speed and accuracy	.193	.552	.237	.246
Mechanical reasoning	.629	.504	.556	.681
Space relations	.518	.642	.617	.680
Spelling	.721	.765	.701	.783
Language usage	.774	.765	.791	.810
Mean age (*SD*) in months	170.0 (7.1)	169.4 (7.8)	202.2 (7.8)	201.0 (8.5)
Mean (*SD*) of raw score of test used for group selection	19.3 (4.8)	39.0 (4.6)	22.7 (4.7)	41.3 (4.4)
Mean (*SD*) IQ	92.1 (7.3)	110.8 (9.2)	89.2 (7.8)	108.3 (9.1)

[a]The effect of the relatively small number of "high" testees is examined in the discussion.

APPENDIX 4

Results of Group Selection Based on Clerical Speed and Accuracy

	Age/Ability Group			
	Young/Low	Young/High	Old/Low	Old/High
Participants (*n*)	389	298	399	385
Variance in 1st Unrotated Principal Component (%)	61.4	63.5	60.5	57.1
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.882	.881	.882	.863
Numerical ability	.811	.828	.826	.824
Abstract reasoning	.782	.826	.742	.801
Mechanical reasoning	.707	.735	.764	.654
Space relations	.767	.773	.713	.738
Spelling	.731	.740	.724	.646
Language usage	.795	.786	.782	.736
Mean age (*SD*) in months	169.6 (7.5)	170.3 (7.8)	201.2 (12.2)	200.3 (8.2)
Mean (*SD*) of raw score of test used for group selection	29.6 (6.5)	57.0 (6.3)	35.3 (6.4)	61.4 (6.4)
Mean (*SD*) IQ	93.9 (10.4)	108.2 (10.0)	92.9 (10.3)	105.2 (9.5)

APPENDIX 5

Results of Group Selection Based on Mechanical Reasoning

	Age/Ability Group			
	Young/Low	**Young/High**	**Old/Low**	**Old/High**
Participants (*n*)	343	319	348	337
Variance in 1st Unrotated Principal Component (%)	51.8	51.2	52.2	52.6
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.859	.846	.835	.819
Numerical ability	.794	.782	.779	.807
Abstract reasoning	.726	.664	.713	.669
Clerical speed and accuracy	.377	.493	.459	.426
Space relations	.555	.605	.582	.658
Spelling	.777	.769	.772	.772
Language usage	.821	.786	.834	.839
Mean age (*SD*) in months	169.5 (7.3)	170.0 (7.8)	201.7 (8.1)	200.7 (8.4)
Mean (*SD*) of raw score of test used for group selection	26.7 (5.1)	48.5 (5.0)	31.2 (5.3)	51.9 (5.1)
Mean (*SD*) IQ	91.9 (8.2)	108.0 (8.8)	91.0 (9.2)	104.2 (9.2)

APPENDIX 6

Results of Group Selection Based on Space Relations

	Age/Ability Group			
	Young/Low	**Young/High**	**Old/Low**	**Old/High**
Participants (*n*)	253	238	252	237
Variance in 1st Unrotated Principal Component (%)	52.5	45.7	53.2	46.8
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.858	.847	.839	.820
Numerical ability	.796	.782	.767	.820
Abstract reasoning	.745	.537	.743	.570
Clerical speed and accuracy	.362	.370	.481	.404
Mechanical reasoning	.586	.510	.550	.460
Spelling	.764	.762	.811	.775
Language usage	.834	.777	.831	.794
Mean age (*SD*) in months	169.8 (7.4)	169.9 (7.2)	201.6 (8.0)	200.9 (8.5)
Mean (*SD*) of raw score of test used for group selection	14.8 (4.8)	35.5 (4.9)	18.8 (5.1)	39.0 (5.0)
Mean (*SD*) IQ	91.5 (8.8)	109.7 (8.2)	90.6 (9.0)	106.1 (8.5)

APPENDIX 7

Results of Group Selection Based on Spelling

	Age/Ability Group			
	Young/Low	**Young/High**	**Old/Low**	**Old/High**
Participants (*n*)	368	311	321	101
Variance in 1st Unrotated Principal Component (%)	49.9	49.4	53.5	54.6
Loadings on 1st Unrotated Principal Component for Each DAT Subtest				
Verbal reasoning	.819	.834	.866	.838
Numerical ability	.750	.733	.765	.851
Abstract reasoning	.809	.805	.791	.711
Clerical speed and accuracy	.287	.392	.231	.309
Mechanical reasoning	.773	.649	.784	.748
Space relations	.712	.713	.786	.794
Language usage	.647	.701	.706	.779
Mean age (*SD*) in months	169.7 (7.2)	169.2 (7.7)	201.9 (7.4)	201.2 (8.8)
Mean (*SD*) of raw score of test used for group selection	40.7 (8.6)	75.6 (8.9)	50.5 (8.3)	83.5 (8.0)
Mean (*SD*) IQ	93.5 (8.5)	109.7 (8.4)	91.1 (8.8)	106.0 (9.5)

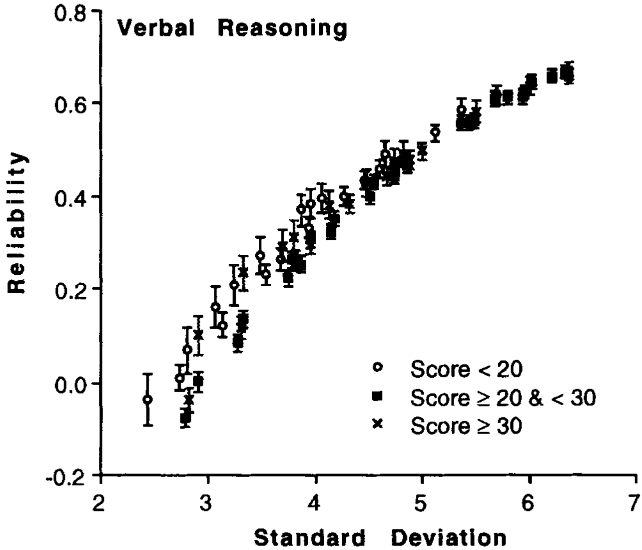
APPENDIX 8

Results of Group Selection Based on Language Usage

	Age/Ability Group			
	Young/Low	**Young/High**	**Old/Low**	**Old/High**
Number of subjects	317	288	312	308
Percent variance in 1st unrotated principal component	46.3	45.6	47.9	44.8
Loadings on 1st unrotated principal component for each DAT subtest				
Verbal Reasoning	.777	.798	.832	.794
Numerical Ability	.777	.723	.814	.785
Abstract Reasoning	.796	.802	.833	.706
Clerical Speed and Accuracy	.171	.302	.289	.274
Mechanical Reasoning	.788	.671	.719	.661
Space Relations	.697	.793	.719	.759
Spelling	.511	.466	.434	.550
Age (*SD*) in months	169.4 (7.4)	170.2 (7.9)	201.7 (8.0)	200.7 (8.3)
Mean (*SD*) of raw score of test used for group selection	21.6 (4.3)	40.0 (4.2)	26.4 (4.4)	44.0 (4.3)
Mean (*SD*) IQ	93.4 (7.6)	111.7 (8.0)	90.7 (8.8)	109.4 (8.2)

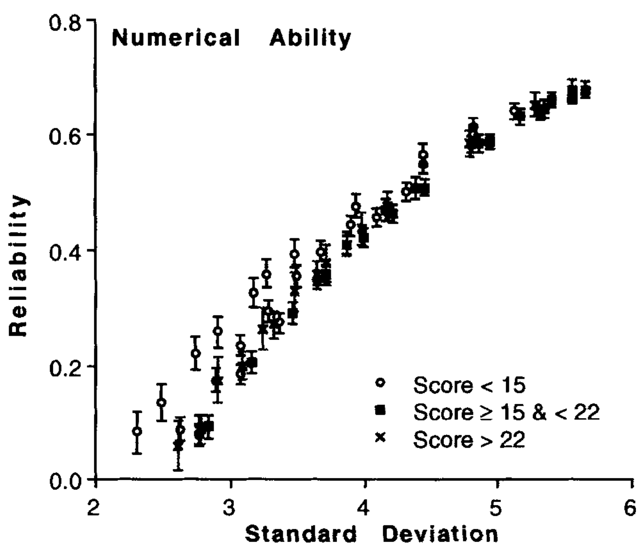
APPENDIX 9

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Verbal Reasoning



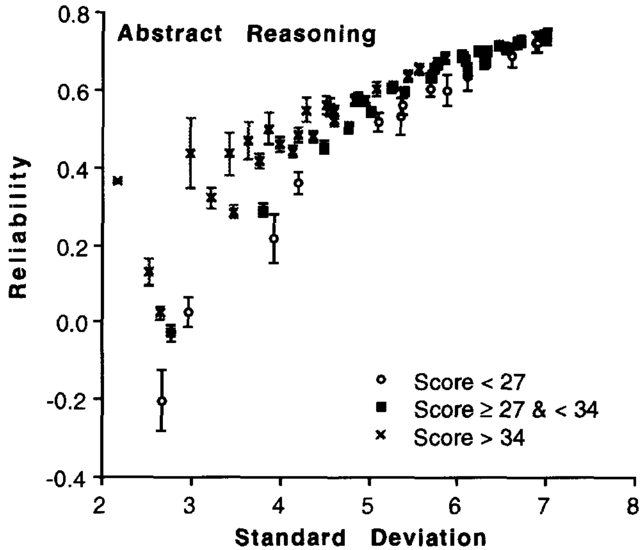
APPENDIX 10

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Numerical Ability



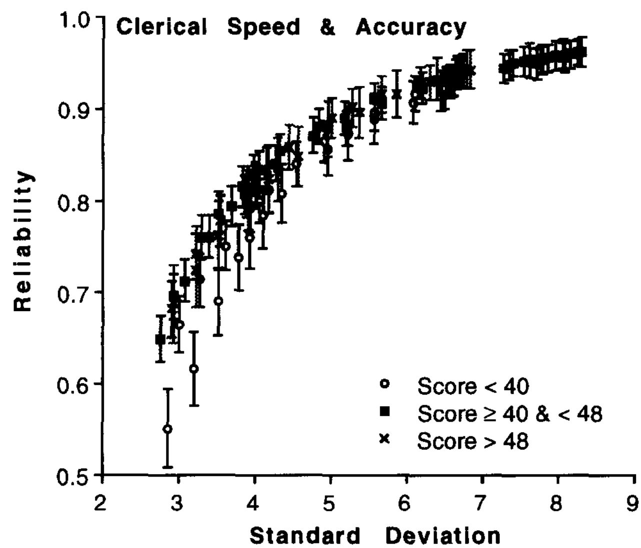
APPENDIX 11

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Abstract Reasoning



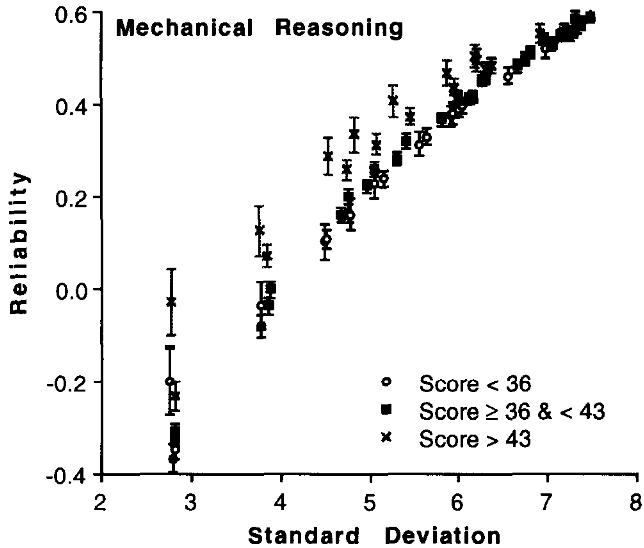
APPENDIX 12

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Clerical Speed and Accuracy



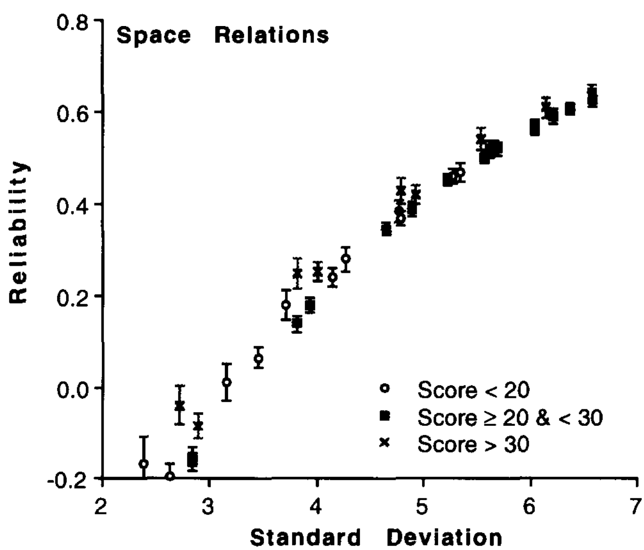
APPENDIX 13

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Mechanical Reasoning



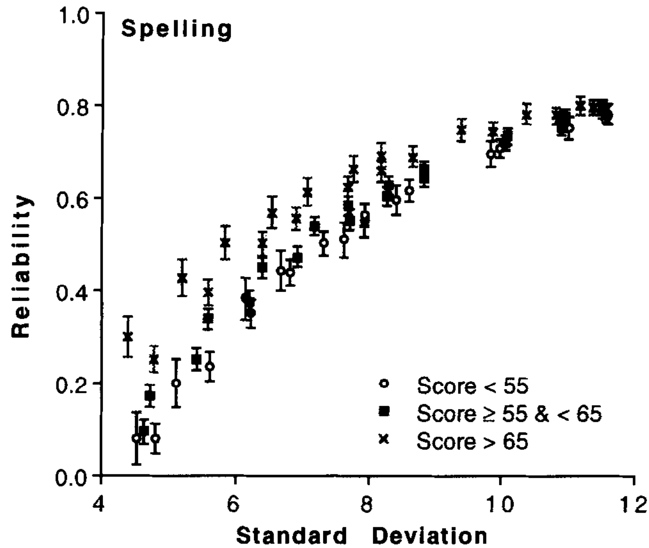
APPENDIX 14

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Space Relations



APPENDIX 15

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Spelling



APPENDIX 16

Reliabilities (± Standard Errors) for Low-, Medium- and High- Ability Groups as a Function of Standard Deviation for Language Usage