A SECULAR DECLINE IN SPEARMAN'S G IN FRANCE

RICHARD LYNN AND COLIN COOPER
UNIVERSITY OF ULSTER

ABSTRACT: A hypothesis originally advanced by Spearman stating that intercorrelations between intelligence tests are lower among high ability populations than among low has recently been confirmed. The hypothesis yields the prediction that the increases that have taken place in intelligence during the last half century in a number of countries should be accompanied by a decrease in test intercorrelations. The prediction was tested on the French standardisation samples of the WISC and WISC-R and confirmed. Further analysis showed that the decline only occurred in verbal-performance subtest intercorrelations, and not between verbal or performance subtests.

It was suggested by Spearman (1927) that g may be more powerful, or account for a greater proportion of the variance among a set of tests, among low ability populations than among those of high ability. This hypothesis has recently been tested by Detterman and Daniel (1989) on the standardization samples of the WISC-R and the WAIS-R. They divided the samples into five ability groups and found that the average intercorrelation declined significantly from the low to the high intelligence groups, thereby confirming Spearman's hypothesis. The hypothesis has been further tested by examining whether it holds for the Scottish standardization sample of the WISC-R and it was found that the hypothesis was confirmed (Lynn 1992).

It is well known that the mean IQ of the populations of a number of economically developed nations has increased by around 15 IQ points over the last half century or so (Lynn & Hampson 1986; Flynn 1987). It should therefore follow that the intercorrelations between tests should be falling. The magnitude of the test intercorrelations is a measure of the strength of Spearman's g, and it can therefore be predicted that the strength of g should be in secular decline.

One of the most obvious sources of evidence to examine for a possible decline in the strength of Spearman's g would seem to lie in the American standardization samples of the WISC and the WISC-R. However these data are not well adapted for this purpose because of differences in the populations sampled in

Direct all correspondence to: Richard Lynn, Psychology Department, University of Ulster, Coleraine, Northern Ireland.

Learning and Individual Differences, Volume 5, Number 1, 1993, pages 43–48.

ISSN: 1041-6080

the two standardizations. The WISC was standardized on whites only, whereas the WISC-R included blacks. The inclusion of blacks is likely to increase the variance in the tests and therefore increase the size of the intercorrelations, and this would counteract the predicted fall. In spite of this problem we have calculated the mean intercorrelations for the two tests for the 7, 10 and 13 year olds for which the correlation matrices are given in both manuals. The results show no general trend but because of the difference in the standardization samples they do not provide an adequate test of the hypothesis.

It is however possible to test the hypothesis of a secular decline in *g* from data available from France and this forms the subject of the present article.

METHOD

The original WISC standardization was performed using 100 male and 100 female children in each of 4 age groups. Geographical region was controlled (3 levels) as was community size (6 levels), school type and years of enrollment (3 levels) following Taver (1952). The WISC-R was standardised as part of a wider national study and parental occupation was included as an additional stratification variable. 50 males and 50 females were tested at each of 11 ages. One slightly unusual feature of the WISC-R standardisation sample is that all subjects were tested between 46 and 58 weeks after their birthdays. Thus the "6.5 yr" sample was six and a half years old ± 6 weeks, rather than between 6 and 7 years old as in the WISC standardization sample. Some "un certain nombre" of immigrants were included in the WISC-R sample, but care was taken to ensure that they had a fluent knowledge of French. There is thus good reason to suppose that the two standardization groups were highly similar in composition.

The tests follow closely the American originals with some changes of detail in the verbal subtests to make them more appropriate for French children. The tests were published in 1965 (WISC) and 1981 (WISC-R) and it is assumed that the testing took place in the two years preceding the dates of publication, thereby giving a 16 year period between the two standardizations. The tests can be obtained from the Centre de Psychologie Appliquée in Paris.

The manual of the WISC gives a correlation matrix of the subtests only for 10.5 year olds and a matrix for the same age group is given in the WISC-R. It is therefore possible to compare the intercorrelations for only this age group.

As the subtests of the WISC and WISC-R are highly reliable (median split half reliabilities of 0.80 and 0.81) the test scores approximate closely to the individuals' "true scores" on each ability. With good sized samples (n = 100 and n = 200) and subtest correlations in the order of 0.4, the standard errors of the correlations between the subtests are acceptably small (in the order of 0.05). The correlations between the subtests are therefore both accurate (because of their small standard errors) and close to the true correlations between the subtests (as the subtests are highly reliable). Correlations between the subtest scores are thus good approximations to correlations between true scores. Hence it is unnecessary to either correct the correlations for unreliability or to take account of possible sampling variation when averaging the transformed correlations.

The French versions of the WISC and WISC-R each comprise 12 subscales. Six measure verbal IQ (information, similarities, arithmetic, vocabulary, comprehension and memory) and six Performance IQ (picture completion, picture arrangement, cubes, object assembly, coding and mazes). There are thus 66 correlations between the subtests which may be usefully broken down into three groups. There are 15 correlations between the Verbal subscales, 15 correlations between the Performance subscales, and 36 correlations between Verbal and Performance subscales.

Taking the WISC and WISC-R in turn, the correlations between the subscales were first normalized by Fisher transformations. The means of these transformed correlations were then computed across:

(a) All 66 correlations between the subscales.
(b) The 15 Verbal subscale intercorrelations.
(c) The 15 Performance subscale intercorrelations.
(d) The 36 correlations between Verbal and Performance subscales.

LISREL analyses were also used to identify any differences in the size of the correlations between the WISC and WISC-R. Unfortunately, however, it was not possible to incorporate corrections for unreliability as close inspection of the French WISC manual indicated that the reliability coefficients quoted there were those from the American standardization sample. Thus Maximum Likelihood analyses were performed to check the equality of the correlation matrices without correcting for unreliability. The method was as outlined in Section V.2 of the Lisrel V manual (Joreskog & Sorbom 1981), and analyses were performed using Version 6 of the package.

RESULTS

The intercorrelations between the subtests were first normalized by Fisher transformations. The means of the normalized subtest intercorrelations were then calculated and the significance of the differences between the means in the WISC and the WISC-R tested by t tests. The results are given in Table 1.

The first row in the table gives the mean for the 66 correlations between all subtests in the WISC and the WISC-R. It will be seen that the value declines from .41 to .36. A one tailed significance test was used because the direction of the change was predicted, and the decline is statistically significant at $p < .05$.

We thought it would be interesting to determine whether the decline in the intercorrelations is present in all sections of the tests when they are divided up into the verbal scale, the performance scale and the verbal-performance cross correlations. These results are given in rows 2, 3 and 4 of Table 1. When the

TABLE 1
Mean Normalized Intercorrelations between Subtests in the French WISC and WISC-R

		WISC		*WISC-R*		
Tests	*N*	*Mean*	*sd*	*Mean*	*sd*	t
All	66	.41	.16	.36	.22	1.66*
Verbal	15	.51	.15	.53	.19	0.55
Performance	15	.36	.16	.32	.22	0.67
Verbal-Performance	36	.38	.14	.29	.28	2.65**

One and two asterisks denote statistical significance using 1 tail tests at the 5 and 1 per cent levels, respectively.

verbal subtests alone are considered, it will be seen that the mean intercorrelations are .51 and .53 and that no significant change has taken place. The same is true for the performance subtests. It is only in the remaining verbal-performance cross correlations that the mean intercorrelation shows a statistically significant decline from .38 in the WISC to .29 in the WISC-R.

The LISREL analyses supported the above conclusions. When used to test the equality of the correlation matrices of the WISC and WISC-R subtests the chi squared "badness of fit" index was 157, with 78 degrees of freedom. The ratio chi squared was thus greater than 2.0, and the LISREL goodness of fit index was 0.86. These figures indicate that there probably are substantial differences between the matrices (Loehlin 1987, p. 67)—a conclusion which substantiates the findings from the t-tests performed on the transformed correlations.

It is important to note that the uncertain reliability of the original version of the French WISC is unlikely to complicate the interpretation of these results. For under the highly probable assumption that the later WISC-R is, if anything, rather more reliable than the original WISC, the correlations between the WISC subscales would be expected to be *smaller* in the WISC than in the WISC-R. The opposite was found.

DISCUSSION

It has been shown that intelligence tests have lower intercorrelations among high ability groups than among low. From this observation it follows that the increase in intelligence which has taken place during the last few decades in many economically developed nations should be accompanied by a decrease in intercorrelations between tests. Examination of the test intercorrelations in the French WISC and WISC-R shows that the predicted decline has taken place over a 16 year period. More detailed examination shows that the decline has not taken place among the verbal or performance subtests. The decline has only occurred in the verbal-performance subtest correlations. The verbal and performance subtests can be taken as measures of Burt's (1949) two major group factors generally regarded as verbal and visuospatial factors. The results of the present study suggest that the factors responsible for the secular increase in intelligence are also decreasing the correlation between these two group factors.

A possible theoretical explanation for the lower intercorrelations among subtests in higher ability groups was advanced by Maxwell, Fenwick, Fenton and Dollimore (1974). These investigators reported a study involving factor analyses of WPPSI subtests on children who were good and poor readers. Among the good readers, factor loadings on the general factor were lower than among the poor readers. This is apparently another finding of lower test intercorrelations among high ability groups. In order to explain the result, Maxwell et al proposed that individuals with high intelligence use fewer neurones for mental processes. This has the effect that the cognitive abilities involved for different problems become more independent and hence the correlations between tests measuring different abilities are lower among high ability groups. Haier (1990) has suggested that in recent findings people with high intelligence have a lower brain cortex metabolic rate and this is consistent with the hypothesis that they use fewer neurones for the performance of cognitive tasks. This hypothesis provides a possible explanation for the lower intercorrelations between cognitive tests among higher intelligence groups, which the findings reported in the present article show as an empirical phenomenon over time as well as between groups of different abilities at one point in time. Another possible explanation involves social and educational factors. It may be that children who are seen as "bright" are encouraged to develop one or two special talents—e.g., music, computer programming, chess—to a high standard. Less able children may not be encouraged to take up such activities. They would also be less rewarded by initial success should they do so. It seems plausible that the skills that are thus developed and trained may generalize to a small number of primary mental abilities. Computer programming may develop numerical ability and coding, chess spatial skills, and creative writing vocabulary. As different bright children would specialize in different areas, this would lead to an overall lowering of correlations between the primary abilities of these children, as noted by Detterman.

REFERENCES

Burt, C. (1949). "The structure of the mind: A review of the results of factor analysis." *British Journal Educational Psychology, 19,* 100–111.

Detterman, D.K. & W.H. Daniel. (1989). "Correlations of mental tests with each other and with cognitive variables are highest for low IQ groups." *Intelligence, 13,* 349–360.

Flynn, J.R. (1987). "Massive IQ gains in 14 nations: What IQ tests really measure." *Psychological Bulletin, 101,* 171–191.

Haier, R.J. (1990). "The end of intelligence research." *Intelligence, 14,* 371–374.

Joreskog, K.G. & D. Sorbom. (1981). *Analysis of structural relationships by maximum likelihood and least squares methods.* Chicago: International Educational Services.

Loehlin, J.C. (1987). *Latent Variable Models.* Hillsdale, NJ: L.E.A.

Lynn, R. (1992). "Is Spearman's g weaker in high IQ groups? Some evidence from Scotland." *Journal of Genetic Psychology, 153,* 229–230.

Lynn, R. & S. Hampson. (1986). "The rise of national intelligence, evidence from Britain, Japan and the United States." *Personality and Individual Differences, 7,* 23–32.

Maxwell, A.E., P.B.C. Fenwick, G.W. Fenton, & J. Dollimore. (1974). "Reading ability, and brain function: A simple statistical model." *Psychological Medicine 4,* 274–280.

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

Taver, G. (1952). "Statistiques francaises des enfants de 6 a 19 ans." *Revue de Psychologie Appliquée, 2,* 184–195.