INTELLIGENCE 17, 557–561 (1993)

Skodak and Skeels: The Inflated Mother–Child IQ Gap

JAMES R. FLYNN
University of Otago

Skodak and Skeels (1949) is one of the few studies that gives IQ scores for both adopted children and their biological mothers. All tend to accept the reported result that these children outscored their mothers by 20 points or more. In fact, at least one third of the children's advantage is an artifact of the 1916 Stanford-Binet test norms plus obsolescence. The mother–child IQ gap is estimated at 13 points with a lower limit of 10 points. The children's mean IQ is put at 105.5, but this is a very rough estimate and the value could easily be as high as 110.

As Jensen (1973) reported, Skodak and Skeels (1949) is one of the most frequently cited studies in the literature of developmental and educational psychology, primarily because it allows a comparison between the IQs of adopted children and the IQs of their biological mothers. There has always been controversy about the interpretation of its results but no one has carefully assessed the IQ values reported. Jensen (1973) was typical: He accepted that the children at age 13 had a mean IQ some 20 points above their mothers. Storfer (1990), at 30 points, put their advantage even higher. I will argue that at least one third of the 20-point gap usually assumed is an artifact.

OBSOLESCENCE

Skodak and Skeels (1949) tested 63 adopted children who were still available at the final follow-up and for whom maternal IQs existed. The children and their mothers were both tested on the 1916 Stanford-Binet, the children at an average age of 13.31 years in 1946, the mothers presumably aged 14 years and over circa 1933. The latter date is an average based on the fact that the mothers were tested after the child's birth and before the child's placement for adoption at under 6 months of age. Flynn (1984) showed that IQ gains over time began on the Stanford-Binet at least as early as 1932, so the children profited from being scored against norms 14 years obsolescent, the mothers against norms only 1 year obsolescent.

Correspondence and requests for reprints should be sent to James R. Flynn, Department of Political Studies, University of Otago, P.O. Box 56, Dunedin, New Zealand.

THE 1916 TEST AND THE 1932 NORMS

It is informative to compare the norms of the 1916 Stanford-Binet and the 1937 Stanford-Binet. The latter was published in 1937 but 1932 represents the midpoint of the period during which its standardization sample was actually tested. The date at which its norms were current is relevant and therefore, I will refer throughout to "1932 norms" and to "1932 scores."

Terman and Merrill (1937, p. 50) provided an equivalence table which shows 1916 test scores very close to 1932 scores from ages 5 to 11 and progressively lower for ages 12 to 18. This apparently conflicts with the hypothesis of IQ gains over time in that the earlier test should yield higher scores. However, the answer is that the standardization sample for the 1916 test was clearly elite. When Yerkes (1921, pp. 654, 789) used the 1916 test norms on the 1918 white draftees, these men aged 21 to 31 had an average mental age of only 13.08 years. This led to the logically absurd comment, echoed for many years, that the average adult white male had a mental age of only 13, and that therefore, almost half of them were morons! In fact, when values for young adults in the equivalence table are computed, they show that mental age rose by perhaps 1.88 years between 1918 and 1932, a considerable gain.

The fact that the 1916 Stanford-Binet had elite norms would not affect the IQ gap between the adopted children and their biological mothers if the norms were similarly elite at ages 11 to 16, which includes all but 1 of the children, and ages 14 and over, which covers the mothers. However, if one assumes the 1932 standardization sample was roughly reliable—otherwise the mother–child mean IQs are simply noncomparable—the norms against which the mothers were scored were significantly tougher.

ADJUSTED VALUES

Therefore, we must make two adjustments that reduce the mother–child IQ gap: (a) allow for the fact that the children profited from 14 years of obsolescence as compared to 1 year; and (b) compensate the mothers for being scored against more elite norms by using the Terman-Merrill (1937) equivalence table. When the 63 children are plotted by age against the equivalence table, the average disadvantage inflicted by the 1916 test norms is 3.34 IQ points. The reader can check that this is approximately correct simply by averaging the disadvantages for ages 11 to 16. I have used the values for ages 17 to 18 to compensate the mothers by 6.70 IQ points. This assumes that all those aged 14 and over suffered from that disadvantage and, unfortunately, the table does not go beyond age 18. But the values for ages 14 to 18 are virtually identical, a trend that at least suggests uniformity above 18. This adjustment carries the children from a mean of 106.30 to 109.64, the mothers from 85.75 to 92.45. The original means were calculated from the individual scores provided (Skodak & Skeels, 1949, pp. 122–124).

TABLE 1
IQ Difference Between the Adopted Children and Their Biological Mothers: How the Original Means Were Adjusted

	Original 1916 Stanford-Binet Means	Correction Based on Equivalence Table[a]		Added Correction for Obsolescence[b]	
		(*SD* = 16)	(*SD* = 15)	(*SD* = 15)	(*SD* = 16)
Children	106.30	109.64	109.04	105.54	105.91
Mothers	85.75	92.45	92.92	92.67	92.18
IQ Difference	20.55	17.19	16.12	12.87	13.73

[a]*Equivalence Table.* Purpose: To compensate for the fact that the 1916 test norms penalize older subjects (the mothers) more than younger ones (the children). Detail: children, 106.30 + 3.34 (age deficit) = *109.64*; mothers, 85.75 + 6.70 (age deficit) = *92.45*. To convert a standard deviation of 16 to 15: children, 9.64 × 15 = 144.6, ÷ 16 = 9.04, + 100 = *109.04*; mothers, −7.55 × 15 = −113.25, ÷ 16 = −7.08, + 100 = *92.92*. The results score both groups against Stanford-Binet norms circa 1932.

[b]*Obsolescence.* Purpose: To compensate for the fact that the mothers were tested in 1933 (when the 1932 norms were 1 year obsolete) and the children in 1946 (when the norms were 14 years obsolete). Detail: children, 109.04 − 3.50 (14 years obsolescence) = *105.54*; mothers, 92.92 − 0.25 (1 year obsolescence) = *92.67*. To convert from a standard deviation of 15 to 16: children, 5.54 × 16 = 88.64, ÷ 15 = 5.91, + 100 = *105.91*; mothers, −7.32 × 16 = −117.28, ÷ 15 = −7.82, + 100 = *92.18*.

Because obsolescence on the Stanford-Binet has been estimated using a white standard deviation of 15, the preceding scores must be translated from the usual Stanford-Binet convention (*SD* = 16). This makes only a small difference, giving 109.04 for the children, 92.92 for the mothers. Allowing for Stanford-Binet obsolescence at a rate of 0.25 IQ points per year (Flynn, 1984, p. 33) and allowing 14 years for the children (1932–1946) and 1 year for the mothers (1932–1933) gives these values: 105.54 as mean IQ for children and 92.67 as mean IQ for mothers. All adjustments are spelled out in Table 1.

The best estimates in Table 1 are those based on current norms, that is, norms corrected for obsolescence. These give 12.87 IQ points as the mother–child gap with a standard deviation set at 15, and 13.73 with a standard deviation set at 16. The former may be a more useful value because a standard deviation of 15 has been popularized by the Wechsler tests. However, the latter must be used to calculate what percentage of the reported mother–child IQ gap is an artifact of the tests. Because a gap of 20.55 points has been reduced to 13.73, only 66.8% of the reported gap is genuine, almost one third being an artifact.

ALTERNATIVE VALUES

Skodak and Skeels (1949, p. 94) gave 100 adopted children both the 1916 Stanford-Binet and the 1937 Stanford-Binet. These were all children still available at the final follow-up and included the 63 children for whom maternal IQs

existed. Both tests were given in one session with the earlier version completed first and the later version immediately after.

The smaller group of 63 subjects scored 115.97 on the 1937 test, which is 6.33 points above the relevant adjusted value in Table 1 (115.97 − 109.64 = 6.33). Probably there was some practice effect, but 6 points is far too high. The usual effect would be about 3 IQ points (Jensen, 1980, pp. 284, 590), but some items common to both tests were given and scored simultaneously. If we allow roughly 2 points for practice effects, this leaves over 4 points unexplained, probably due to random factors and variable test administration. This calls into question the preferred value for the adopted children's mean IQ from Table 1, the value of 105.54 when they are scored against norms corrected for obsolescence with a standard deviation set at 15. The mean could easily be as high as 110 (105.54 + 4.33 = 109.87) and there seems to be no reason for favoring one value over the other.

The Table 1 estimate of 12.87 points as the mother–child IQ gap is not directly affected. As Skodak and Skeels (1949, p. 108) rightly asserted, the mothers probably would have enjoyed much the same score increase from the double administration that their children received. However, the 100 subjects who took both the 1916 and 1937 tests can be used to construct an age-by-age equivalence table to supplement the one provided by Terman and Merrill (1937). Compared to the latter, the former shows that the mothers suffered from an extra 2.07-point (*SD* = 15), age-related deficit thanks to the 1916 test norms. Because the 100 subjects included the 63 for whom maternal IQs were available, the new value could be argued to be more appropriate. On the other hand, the new table gives values only through age 16 as compared to age 18. The most judicious conclusion is that the mother–child IQ gap could be as low as 10.80 points (12.87 − 2.07) with arguments both for and against the lower value.

SUMMARY AND IMPLICATIONS

The IQ advantage of the adopted children over their biological mothers should be put somewhere between 10.80 and 12.87 points. The value for the children's mean IQ should be put somewhere between 105.54 and 109.87. The estimates are carried to two decimal places only because those are the values the calculations yield. As the analysis implies, it would be wrong to think that even the best adjustments can give anything but rough estimates. The children's IQ advantage probably lies between 10 and 13 points. The children's mean IQ probably lies between 105 and 110.

Nonetheless, there is no case for adhering to the old unadjusted values and those who cite Skodak and Skeels (1949) will go astray if they do. For example, Storfer (1990, p. 63) increased the already inflated 20-point mother–child gap to 30 points by using unadjusted 1916 test scores for the former and unadjusted 1937 test scores for the latter (115.97 − 85.75 = 30.22). Storfer (1990, pp. 63,

424) also used 115.97 as his absolute value for the children's mean IQ. He concluded that, although adoptive homes of high quality raised IQ as much as 16 points above the national average in the 1930s, the effect had diminished to perhaps 3 to 8 points by the 1970s. Given that best estimates here put the Skodak and Skeels children at only 5 to 10 points above the national average, the results approximate the range Storfer posited for the 1970s and cannot be cited as evidence for the trend he described.

REFERENCES

Flynn, J.R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.

Jensen, A.R. (1973). Let's understand Skodak and Skeels, finally. *Educational Psychologist, 10,* 30–35.

Jensen, A.R. (1980). *Bias in mental testing.* London: Methuen.

Skodak, M., & Skeels, H.M. (1949). A final follow-up study of one hundred adopted children. *The Journal of Genetic Psychology, 75,* 85–125.

Storfer, M.D. (1990). *Intelligence and giftedness: The contributions of heredity and early environment.* San Francisco: Jossey-Bass.

Terman, L.M., & Merrill, M.A. (1937). *Measuring intelligence.* London: George G. Harrap.

Yerkes, R.M. (1921). Psychological examining in the United States Army. In *Memoirs of the National Academy of Sciences* (Vol. 15). Washington, DC: U.S. Government Printing Office.