Trends in preschool lead exposure, mental retardation, and scholastic achievement: Association or causation?

Rick Nevin
National Center for Healthy Housing, USA

Abstract
This study shows that 1936–1990 preschool blood lead trends explain 65% of the 1948–2001 variation in USA mental retardation (MR) prevalence, 45% of the 1953–2003 variation in the average scholastic achievement test (SAT) verbal score, and 65% of the 1953–2003 variation in the average SAT math score. These temporal relationships are characterized by best-fit time lags (highest $r^2$ and $t$-value for blood lead) consistent with lead-induced cognitive damage in the first year of life: a 12-year lag for school-age MR, and a 17-year lag for SAT scores. Recent shifts in age-specific MR prevalence are consistent with recent trends in preschool blood lead. SAT and MR trends by race are consistent with racial differences in how 1960s slum clearance affected childhood exposure to severe lead paint hazards. SAT trends by Hispanic origin are consistent with an especially sharp fall in preschool blood lead in New York City since 1970.

1. Introduction

This study examines whether trends in preschool lead exposure can explain trends in school-age mental retardation (MR) prevalence and scholastic achievement test (SAT) scores. Racial, Hispanic origin, and birth cohort differences in lead exposure are compared with subsequent racial, Hispanic origin, and age-specific shifts in MR prevalence and SAT scores.

The link between MR and lead poisoning was first reported in the 1940s (Byers and Lord, 1943). In the 1960s, lead poisoning was associated only with childhood blood lead above 60 μg/dL (micrograms of lead per deciliter of blood) but subsequent research links lower blood lead to elevated MR risk (Marlowe, 1995; David et al., 1976, 1982). Many studies also show an inverse relationship between blood lead and IQ and academic achievement (US Centers for Disease Control and Prevention, 1991; Lanphear et al., 2005; Miranda et al., 2007). Higher blood lead is associated with larger absolute losses in IQ, but the impact of each μg/dL increase is greater at lower levels. Preschool blood lead of 10μg/dL is associated with a loss of 7.4 IQ points relative to children with blood lead of 1 μg/dL (Canfield et al., 2003). Each μg/dL increase from 10 to 15 μg/dL lowers IQ by 0.323 points and each μg/dL above 15 μg/dL lowers IQ by 0.232 points (Schwartz, 1994). Therefore, blood lead of 40 μg/dL lowers IQ by about 15 points versus blood lead of 1 μg/dL, and blood lead of 60–80 μg/dL reduces IQ by about 20–25 points.

MR is characterized by significant limitations in intellectual functioning and adaptive behavior. The specific IQ associated with “limitations in intellectual functioning” has varied by state, but special education for MR generally requires IQ below 70–75 (National Research Council, 1982). By definition, average IQ is 100, and 25% of the population has IQ below 90, so a loss of 15–25 IQ points due to blood lead of 40–80 μg/dL suggests lead exposure could substantially affect the percent of children satisfying the MR condition of significant intellectual limitations. IQ is a good predictor of academic achievement (Neisser et al., 1996), so the loss of 7.4 IQ points due to blood lead of 10 versus 1 μg/dL also impacts achievement.

Elevated blood lead can be due to lead paint chip ingestion, inhaled air lead, and other pathways, but lead in paint and gasoline had especially pervasive effects due to contaminated dust ingested via hand-to-mouth activity as children crawl (US Environmental Protection Agency, 1986). The lead share of USA pigments fell from near 100% in 1900 to 35% by the mid-1930s (Meyer and Mitchell, 1943), and the USA banned lead paint after 1977, but 80% of pre-1940 and 46% of 1940–1959 homes still had some interior lead paint in 1999 (Jacobs et al., 2002). Average preschool blood lead tracked trends in leaded gas use from the 1930s through the 1980s as air lead fallout contaminated dust while lead paint exposure changed slowly with slow changes in the housing stock (Thomas et al., 1999; Nevin, 2007). Average USA preschool blood lead was 15 μg/dL in the late 1970s, but fell sharply from 1975 to 1987 due to the leaded gas phase-out (Pirkle et al., 1994). Since 1990, preschool blood lead prevalence over 10 μg/dL has tracked trends in lead dust hazards from lead paint (Jacobs and Nevin, 2006; Nevin and Jacobs, 2006; Nevin et al., 2007).

E-mail address: ricknevin@verizon.net

Please cite this article as: Nevin, R., Trends in preschool lead exposure, mental retardation, and scholastic achievement: Association or... Environ. Res. (2009), doi:10.1016/j.envres.2008.12.003
Many city children suffered severe lead poisoning from 1950 to 1970 due to additive exposure to air lead and deteriorated lead paint in city slums. Atmospheric emissions from gas lead affected blood lead even in rural areas, but traffic caused severe city exposure because 55% of emissions settled within 20 km of the roadway (Organization for Economic Cooperation and Development, 1993). Most 1950–1970 slum housing was built around 1900 when the use of heavily leaded interior paint was common. Lead dust hazards are still present in 61% of all homes with deteriorated interior lead paint (Jacobs et al., 2002), and lead paint chip ingestion is often a factor in severe lead poisoning: In 1989–1990, after the leaded gas phase-out, children with X-ray evidence of paint chip ingestion had average blood lead of 63 μg/dL (McElvaine et al., 1992). Slum demolition reduced severe lead paint hazards in the 1960s, but 25% of city children tested still had blood lead over 40 μg/dL when gas lead use peaked around 1970.

Gillsinn (1972) found that substandard housing prevalence explained 95% of local variation in children over 40 μg/dL in 1970, but the percent of children over 40 μg/dL relative to substandard housing prevalence was higher in New York City due to higher air lead in larger cities. Blacks accounted for 15% of central city households in 1960, but occupied 56% of substandard city housing, and the percent of all central city blacks in substandard housing was 25% in 1960 and 16% in 1966 (Kristof, 1968; Koebel, 1996). Per capita gas lead use fell from 1956 to 1962 but hit new highs from 1966 to 1974, when 62% of blacks under 6-year olds lived in central cities, versus 24% of whites under age six (US Census, 1960–90). Average blood lead for black 2-year olds in Chicago and New York City fell about 30% from 1970 to 1978, but the late 1970s USA average for black children ages 6–36 months was still 30% above the white average (Agency for Toxic Substances and Disease Registry, 1988).

2. Data and methods


The calculation of “public school” MR reflects children in special education for MR at the start of each academic year as a percent of total enrollment. Annual data are available for 1976–2006 and straight-line interpolation was used to derive pre-1976 trends from 3- to 6-year interval data (US Department of Education, 1993, 1997–2007).

Age-specific MR prevalence reflects children in an age bracket in special education for MR at the start of each academic year as a percent of total enrollment. Annual data are available for 1976–2006 and straight-line interpolation was used to derive pre-1976 trends from 3- to 6-year interval data (US Department of Education, 1993, 1997–2007).

Age-specific MR trends from 1991 to 2006 are compared with the earlier NHANES data on children ages 1–5 years, and 1997–2004 surveillance data on children under 7 years, with blood lead over 10 μg/dL (US Centers for Disease Control and Prevention, 1997, 2000, 2007a, b).

This analysis reports the SAT data by calendar years at the start of academic years (most students take the SAT in the first half of their senior year). The SAT data are available for 1966–2004 based on the same SAT norm, and comparable 1951–1965 estimates were derived from data based on an older norm (College Board, 1987–2005; US Department of Education, 2004, 1997–2007). Fig. 1 shows trend estimates for the percent of test-takers who speak a foreign language at home and the percent who took SAT prep courses outside of school. Students who speak a foreign language at home have lower scores (especially verbal scores) and prep courses are expected to increase scores. Prep course trends reflect survey data, straight-line interpolation between survey years, and assume no pre-1977 or post-1995 change. Trends for foreign language reflect reported 1986–2004 data and an estimated 1970–1986 rise and no pre-1970 change, based on trends in Hispanic enrollment and language spoken at home. A foreign language was spoken at home by 8% of 18–24-year olds in 1979 and 14% in 1989; Spanish was spoken by 60% of these youths in 1979; and the Hispanic share of 16–24-year-old college enrollment was 3.1% in 1972 (the earliest available data) and 4.5% in 1979 (US Department of Education, 2004, 1997–2007).

Regressions were run with preschool blood lead lags of 1–45 years to identify best-fit lags for MR prevalence and SAT scores, where the “best-fit” lag has the highest significance (R2) for blood lead and percent of variation explained (R2). Regression results are reported for best-fit lags based on MR data through 2001 and the SAT data through 2003. Graphs show the best-fit lag relationship for blood lead versus MR through 2006 and SAT trends through 2004. To illustrate best-fit lags, regression R2 is also graphed across time lags for MR prevalence and SAT scores.

3. Results

Fig. 2 shows public school MR increased from 0.06% in 1914 to 0.38% in 1935, but was little changed from 1935 to 1939, and fell from 1939 to 1947. Fig. 3 shows MR then increased from 0.37% in 1948 to a peak of 2.16% in 1976 and fell to 1.09% in 2006, tracking blood lead with a 12-year lag. Fig. 4 graphs R2 across time lags for which 1936–1999 blood lead is statistically significant, showing the 12-year MR time lag is the “best-fit” (highest R2). Table 1 shows regression results for the best-fit 12-year lag for blood lead versus 1948–2001 MR (SAT results in Fig. 4 and Table 1 are discussed below).

Overall public school MR was relatively stable at 1.2–1.3% from 1991 to 2004, but Fig. 5 shows age 6–11-year-old MR fell from 1% in 1996 to 0.7% in 2006, and age 12–17-year-old MR prevalence fell from 1.27% in 2001 to 1.13% in 2006. The percent of 18–21-year olds in special education for MR was 0.4% in 1996 and 2002–2006. Fig. 6 shows 6–8-year-old MR declined about 35% from 1993 to
2006, as 17-year-old MR rose 7%. Age 6–21-year-old black student MR fell from 2.11% in 1998 to 1.69% in 2006, as age 6–21 white MR fell from 0.78% to 0.63%.


Fig. 4 graphs SAT regression $R^2$ across all time lags for which the independent variables are statistically significant and have the expected sign. $R^2$ peaks at an 18-year lag across SAT math simple regressions, but the prep course and language variables are significant only with a 16- or 17-year blood lead lag, and multiple regression $R^2$ peaks at a 17-year lag. $R^2$ peaks at 23 years for SAT verbal simple regressions, but the prep and language variables are...
significant only with blood lead lags of 12–19 years, and multiple regression $R^2$ peaks at a 17-year lag. Fig. 7 shows the fall and rise of SAT math and verbal scores have tracked the rise and fall of blood lead with a 17-year lag. Table 1 shows blood lead with a 17-year lag is highly significant in regressions with and without SAT prep and foreign language variables. $R^2$ is 64.7% in a simple regression for SAT math versus blood lead with a 17-year lag, and adding the prep and language variables only increases $R^2$ to 67.7%. $R^2$ is 45.4% in a simple regression for SAT verbal versus blood lead with a 17-year lag, and adding the prep and language variables increases $R^2$ to 90%. $R^2$ is 86.1% in a verbal regression with just the language variable and a 17-year lag for blood lead.

Fig. 8 shows per capita gas lead use was little changed as urban renewal projects in execution increased 100-fold from the early 1950s through the 1960s. A single urban renewal project often involved demolition of large numbers of dilapidated slum units. Overall public school MR fell from 2.2% in 1976 to 1.2% in 1992, as black student MR fell from 4.1% to 2.5%. White MR fell from 1.3% in 1976 to 1.07% in 1986 then rose to 1.15% in 1992. Average black SAT scores increased substantially from 1975 to 1987, as white SAT scores fell slightly.

USA preschool blood lead fell after the mid-1970s, but Mexico did not reduce gas lead use until the late 1980s, affecting foreign-born Mexican–Americans. Native-born Mexican–Americans are also disproportionately located in border areas that had cross-border gas lead exposure, whereas the Puerto Rican population is disproportionately located in New York City (US Census, 2001). New York City had extremely high air lead in the 1960s, followed by an especially sharp drop in blood lead after 1970. In 1986, average SAT math and verbal scores for college-bound Puerto

---

**Table 1**

Regressions for preschool blood lead versus mental retardation prevalence and average SAT math and verbal scores.

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mental retardation</td>
<td>Intercept</td>
<td>0.00498</td>
<td>0.00007</td>
<td>5.13</td>
<td>&lt;.0001</td>
<td>0.646</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>0.00666</td>
<td>0.00007</td>
<td>9.74</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAT verbal**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>584.8</td>
<td>3.78</td>
<td>154.61</td>
<td>&lt;.0001</td>
<td>0.900</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>−2.1</td>
<td>0.18</td>
<td>−11.81</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAT prep</td>
<td>485.4</td>
<td>114.35</td>
<td>4.25</td>
<td>0.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foreign language</td>
<td>−615.0</td>
<td>84.00</td>
<td>−7.32</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAT verbal**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>574.0</td>
<td>3.27</td>
<td>175.63</td>
<td>&lt;.0001</td>
<td>0.861</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>−1.7</td>
<td>0.18</td>
<td>−9.54</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foreign language</td>
<td>−268.1</td>
<td>22.59</td>
<td>−11.87</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAT verbal**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>550.7</td>
<td>5.12</td>
<td>107.49</td>
<td>&lt;.0001</td>
<td>0.454</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>−2.2</td>
<td>0.35</td>
<td>−6.38</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAT math**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>527.9</td>
<td>3.00</td>
<td>175.70</td>
<td>&lt;.0001</td>
<td>0.677</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>−1.3</td>
<td>0.14</td>
<td>−9.37</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SAT prep</td>
<td>189.6</td>
<td>90.84</td>
<td>2.09</td>
<td>0.0423</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Foreign language</td>
<td>−135.1</td>
<td>66.73</td>
<td>−2.03</td>
<td>0.0485</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**SAT math**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Independent variable</th>
<th>Lag (years)</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-Value</th>
<th>p-Value</th>
<th>$R^2$</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intercept</td>
<td>523.7</td>
<td>1.82</td>
<td>287.24</td>
<td>&lt;.0001</td>
<td>0.647</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Preschool blood lead</td>
<td>−1.2</td>
<td>0.12</td>
<td>−9.48</td>
<td>&lt;.0001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Regressions were run for preschool blood lead versus mental retardation prevalence and average SAT math and verbal scores with time lags of 1–45 years, and the best-fit lag (highest $R^2$ and t-value for blood lead) was 12 years for mental retardation prevalence and 17 years for average SAT math and verbal scores. The percent of students taking SAT prep courses outside of school and the percent who speak a foreign language at home were also significant in the SAT regressions.

---

Please cite this article as: Nevin, R., Trends in preschool lead exposure, mental retardation, and scholastic achievement: Association or... Environ. Res. (2009), doi:10.1016/j.envres.2008.12.003
Ricans were 23 and 21 points lower, respectively, than average scores for Mexican–Americans. In 2002, the average math score for Puerto Ricans was 4 points below the Mexican–American average, and their average verbal score was 8 points higher than the Mexican–American average.

4. Discussion

The increase in MR from 1914 to 1935 is consistent with expanding use of IQ tests in schools (Plucker, 2003), but MR prevalence was little changed from 1935 to 1939 and declined...
from 1939 to 1948. That 1940s MR decline occurred among students born after per capita use of lead in USA paint fell 58% from 1914 to 1930 (Nevin, 2000), and before any substantial rise in gas lead exposure.

Best-fit lags for MR and SAT trends versus 1936–1990 blood lead are consistent with lead-induced cognitive damage in the first year of life. Blood lead coefficients are highly significant at best-fit lags, with the expected positive sign for MR and negative signs for SAT scores. SAT prep and foreign language are also statistically significant with the expected signs, but their inclusion has little impact on the value or significance of the blood lead coefficient. The explanatory power of language and prep is small in the math regression, raising $R^2$ from 64.7% in a simple regression to 67.7% in a multiple regression. Adding SAT prep and foreign language variables does raise verbal regression $R^2$ from 45% to 90%, mostly due to the language variable. (SAT scores fell in 2005, but this appears to be related to a new SAT test that year, as the same academic year saw the rival ACT record its biggest increase in average scores in 20 years.)

Racial trends in MR and SAT scores are consistent with birth cohort slum clearance trends that disproportionately affected black children. Average preschool blood lead peaked around 1970, but severe lead poisoning prevalence must have peaked before slum demolition reduced severe lead paint hazards in dilapidated housing. The 1976 MR peak is consistent with a 12-year lag after a peak in severe lead poisoning around 1964, and the fall in MR from 1976 to 1992 was almost entirely due to a decline among black children. Hauser (1998) shows that the racial difference in National Assessment of Educational Progress (NAEP) scores also narrowed at different times for ages 9, 13, and 17-years old, but black gains at each age were traced to changes across 1962–1973 birth years. The 17-year-old NAEP racial trends are similar to the 1975–1987 racial convergence in SAT scores. These MR, SAT, and NAEP trends are all consistent with a 1960s convergence in the relative severity of black and white preschool lead exposure. In absolute terms, however, black children still have higher MR prevalence and lower average SAT and NAEP scores, consistent with black children continuing to have higher average blood lead.

Office of Civil Rights (OCR) data for 1978–1979 and earlier MR research found that 75–80% of all MR students were “Educable Mentally Retarded” (EMR), with adaptive limitations (largely based on teacher observations) and IQ above 55 (National Research Council, 1982). Most EMR cases were of unknown cause, and prevalence was high among low-income children who lived in city slums. In 1970, 25% of city children tested had blood lead over 40 μg/dL, and substandard housing explained 95% of local variation in children over 40 μg/dL, so the percent over 40 μg/dL must have been much higher before 1960s slum clearance. Blood lead of 40 μg/dL is associated with a loss of 15 IQ points, so the large percentage of 1960s children over 40 μg/dL must have pushed many of those children below the IQ level associated with EMR. The 1978–1979 OCR data and earlier research found higher EMR prevalence among 10–14-year olds than 5–9-year olds, suggesting EMR diagnosis was directly related to academic demands of school. Age 15–19-year-old prevalence was half the 10–14-year-old prevalence, as EMR students were more likely to drop out (and lose the MR label once they left school). Recent MR data show a dramatic shift in age-specific MR risk, with peak prevalence at 14 years in 1993 and 16 years in 2005, consistent with younger students showing earlier effects of the ongoing decline in preschool blood lead.

There are several limitations to the data used in this analysis. Public school MR prevalence is somewhat affected by temporal changes in the percent of the school-age population enrolled in public schools, but enrollment trends do not appear to explain MR trends: Grade 9–12 enrollment was 70% of the 14–17-year-old population in 1952, 80–85% from 1963 to 2001, and 87% in 2005 (US Department of Education, 1993, 1997–2007). The calculation of age-specific MR also understates total population prevalence for older ages because the numerator excludes MR individuals not enrolled in public school special education, but this mainly affects 18–21-year-old prevalence. The blood lead trends used in this analysis are subject to blood lead survey random error and some estimation error, but potential error is small relative to the large temporal changes in preschool blood lead. This analysis cannot control for temporal changes in public school MR related to changes in MR diagnosis, or education attainment trends that could affect the average parental IQ for students taking the SAT, but trends in education attainment and special education diagnoses (discussed below) do not appear to explain the findings reported here.

4.1. Education attainment trends versus SAT, MR, and other special education trends

The 1970s SAT decline was initially thought to reflect a change in SAT test-taker population, as increased access to college reduced the average ability of students taking the SAT. Long-term trends now show the opposite relationship between achievement, as measured by the SAT, and educational attainment. SAT scores were relatively stable from 1960 to 1968, as the 16–24-year-old status dropout rate (percent not in school and without a diploma or GED) fell from 27% to 16%, and the 16–24-year-old college enrollment rate of high school completers (enrolled within 12 months of diploma or GED) rose from 45% to 55%. SAT scores fell as the status dropout rate declined to 14% in 1973 but stayed at 13.5% or higher until 1983, and as the college enrollment rate stayed below 51% from 1972 to 1980 and did exceed the 1968 rate until 1985. The subsequent rise in SAT scores occurred as the dropout rate fell from 13.7% in 1983 to 9.4% in 2005 and the college enrollment rate rose from 55% in 1984 to 68.6% in 2005.

MR students who left school as dropouts fell 23% from 1993 to 2005 as those graduating with a certificate or diploma rose 97% and 32%, respectively. Some research has suggested falling MR and rising “specific learning disability” prevalence from 1976 to 1990 reflects more restrictive MR criteria shifting mild MR pupils to the specific learning category (US Department of Education, 1991), but this hypothesis implies an increase in average MR severity, which is inconsistent with subsequent gains in MR student attainment.

Shattuck (2006) argues diagnostic substitution could explain the 1994–2003 fall in 6–11-year-old MR and concurrent rise in 6–11-year-old autism, but this hypothesis cannot explain why MR fell from 1976 to 1992. Moreover, diagnostic substitution is not consistent with high autism prevalence among white and upper-income pupils (US Centers for Disease Control and Prevention, 2006) because MR prevalence has always been higher for black and low-income pupils. Also, 17-year-old autism and MR both rose from 1993 to 2005, as more MR students remained in school.

A multi-site national study has shown that borderline and mild MR students also recorded IQ averaging 5.6 points lower on the 1994 WISC-III versus their earlier scores on the 1974 WISC-R (Kanaya et al., 2003). This could explain why both 6–11- and 12–17-year-old MR increased in the early 1990s: as more schools adopted the WISC-III in MR assessments, more students recorded IQ scores with MR, when many of those students would have had WISC-R IQ above the level associated with MR. MR student performance on the 1991 WISC-III versus the 1974 WISC-R suggest the 1993–2005 gains in MR student attainment are the result of real gains in cognitive ability in the population classified as MR.
5. Correlation and causation conclusions

This analysis poses a specific variation of the question addressed by Sir Austin Bradford Hill (1965) in “The Environment and Disease: Association or Causation?”

Our observations reveal an association between two variables, perfectly clear-cut and beyond what we would care to attribute to the play of chance. What aspects of that association should we especially consider before deciding that the most likely interpretation of it is causation?

Preschool lead exposure, MR, and SAT trends are reviewed below in the context of the nine “viewpoints” Bradford Hill recommended in assessing evidence of causation.

(1) **Strength:** Bradford Hill illustrates this viewpoint with evidence that cigarette smokers have a lung cancer death rate nine times that of non-smokers, and a coronary thrombosis death rate “no more than twice” that of non-smokers. He notes there is evidence of causation in both cases, but the strength of the lung cancer relationship provides stronger evidence. The association between blood lead and MR trends falls between these two benchmarks: MR risk for children born near the early 1960s peak exposure to air lead and lead paint hazards in slums was five to six times the risk for children born in the late 1930s. The strength of the blood lead, MR, and SAT trend association is also reflected in high regression $R^2$ and the significance of blood lead coefficients.

(2) **Consistency:** Has the association “been repeatedly observed by different persons, in different places, circumstances and times?” (Hill, 1965) MR and SAT trends show a strong association with 1936–1990 USA blood lead trends, encompassing changes in a large, diverse, national population across four time zones and over five decades. Declining MR from 1939 to 1947, tracking earlier declines in lead paint, shows the same association back to 1920s birth cohorts. These trends are consistent with controlled studies showing a causal association between lead poisoning and MR in individual children (Byers and Lord, 1943; Marlowe, 1995; David et al., 1976, 1982). The negative correlation between blood lead and SAT trends is consistent with research demonstrating an inverse relationship between preschool blood lead and IQ, and research showing IQ is an excellent predictor of academic achievement (Schwartz, 1994; US Centers for Disease Control and Prevention, 1991; Lanphear et al., 2005; Miranda et al., 2007; Canfield et al., 2003; Neisser et al., 1996). The long-term nationwide association between blood lead, MR, and SAT trends is consistent with slum clearance trends and subsequent racial trends in MR and in SAT and NAEP scores, and with large Puerto Rican SAT gains after a large fall in New York City preschool blood lead.

(3) **Specificity:** MR and SAT trends show a relationship between preschool blood lead and two manifestations of the specific outcome of impaired cognitive ability, reflecting cognitive impacts across the preschool blood lead distribution. The 1936–1990 average blood lead trend reflects a temporal shift in the entire preschool blood lead distribution, including a rise and fall in severe lead poisoning prevalence especially associated with MR, and in marginally elevated blood lead prevalence associated with marginally lower IQ and academic achievement.

(4) **Temporality:** At a minimum, temporality asks if the suspected cause preceded the effect, but best-fit time lags for MR and SAT scores present an especially compelling case of temporality. While time-series comparisons can result in coincidental correlations, $R^2$ changes in a striking pattern across regressions comparing MR and SAT scores with different time lags for 1936–1990 blood lead trends. MR shows no correlation with blood lead with lags of less than 2 or over 21 years—the blood lead coefficient in such regressions is insignificant—and $R^2$ peaks at a 12-year lag, for students around 12-year olds. SAT math scores show no significant relationship with blood lead with lags of less than 7 or over 26 years, and simple regression $R^2$ peaks at an 18-year lag. With prep and foreign language variables added to the model, $R^2$ peaks at a 17-year lag for both math and verbal scores, for test-takers around 17-year olds. These MR and SAT lags are consistent with lead-induced cognitive damage in the first year of life. Age-specific MR trends reinforce the temporality evidence of causation. Research in the 1980s found MR prevalence peaked in elementary and junior high school, but prevalence peaked at age 14 in 1993 and age 16 in 2006. Declines in 6–11-year-old MR from 1996 and 2006 and age 12–17 MR from 2002 to 2006 track the same documented birth cohort decline in preschool blood lead after the late 1980s.

(5) **Biological gradient (dose–response relationship):** Controlled studies show a dose–response relationship between preschool blood lead and IQ later in life: a loss of 7.4 IQ points due to blood lead of 10 versus 1 μg/dl, and 15–25 IQ points associated with blood lead of 40–80 μg/dl (Schwartz, 1994; Canfield et al., 2003). Preschool blood lead trends show a corresponding population—dose–response relationship with population MR and SAT scores. Blood lead rose over four-fold from the 1930s through the 1960s and MR prevalence increased by a similar order of magnitude with a 12-year lag. The fall in MR from the 1970s to 1992 was also roughly proportionate to the earlier decline in preschool blood lead, before the introduction of a new IQ test used in school evaluations in the early 1990s.

(6) **Biological plausibility:** Documented neurochemical, subcellular, and cellular effects of preschool blood lead establish a biological basis for impairments in IQ, learning, and behavior. These effects, which are not entirely independent, due to the integrated nature of the nervous system, include the following: (1) inhibition of mitochondria, disrupting energy metabolism; (2) alteration of calcium-mediated neurotransmitter release of postsynaptic receptors; (3) reduction in myelin formation; (4) effects on synaptogenesis, presynaptic terminals and postsynaptic receptors; (5) changes in neurotransmitter release, including but not limited to dopamine; (6) abnormalities in electrophysiological activity for both cortex and brain stem including increased latencies and decreased amplitudes of evoked potentials; and (7) effects on brain development related to immaturity of the blood–brain barrier during gestation (Banks et al., 1997).

(7) **Coherence:** A causal interpretation “should not seriously conflict with the generally known facts of the natural history and biology of the disease” (Hill, 1965). The association between blood lead, MR, and SAT trends reveals an ecological coherence with controlled studies linking individual lead exposure to cognitive damage. The racial convergence in MR prevalence and SAT and NAEP scores across slum clearance birth years is also consistent with the known history of heavily leaded paint use in older housing, and lead dust and paint chip ingestion risks in severely deteriorated city slums. Lead exposure trends also reveal a coherent relationship with international trends in IQ and social behavior that is
correlated with IQ (Nevin, 2000, 2007). The cognitive and behavioral effects of preschool lead exposure can also reconcile the seeming incoherence of controversial research on IQ correlates and recent behavior trends. An analysis of 1979–1990 National Longitudinal Survey of Youth (NLSY)* data found that dropping out of high school and incarceration were both strongly associated with low IQ, independent of race or socioeconomic status (Herrnstein and Murray, 1994). NLSY participants, ages 14–22 years in 1979, constituted a large representative sample of Americans born in the late 1950s and the early 1960s. The high school dropout rate (percent who failed to get a diploma or GED by ages 25–33 years) was 55% for white NLSY participants with IQ below 75, 35% for those with IQ of 75–90, 6% for IQ of 90–110, and less than 1% for IQ over 110. The fall in dropout rates since 1990 means these IQ-related dropout risks cannot be true today. The low NLSY dropout rate for those with IQ above 90 also means the fall in dropout rates must reflect an especially sharp decline for those with IQ below 90. This conclusion is also consistent with the data showing MR student dropouts fell 23% from 1993 to 2005, as MR students graduating with a diploma or certificate increased. Less than 1% of white NLSY males with IQ above 110 were incarcerated prior to 1980, versus 3% of those with IQ of 90–110, and 7% of those with IQ below 90. The fall in USA crime since 1990 means these incarceration risks cannot be true today. The low incarceration rate for NLSY males with IQ above 90 also means the crime decline must reflect an especially sharp decline for those with IQ below 90. The 1990s crime decline was led by a sharp fall in juvenile offending, consistent with the decline in MR led by younger students. The 1960s rise in crime was also led by surging juvenile offending, tracking the earlier rise of preschool blood lead (Nevin, 2007), consistent with the 1948–1976 rise in MR. The 1960s birth years linked to a racial convergence in MR. The 1960s birth years linked to a racial convergence in MR, SAT, and NAEP scores, and juvenile burglary arrest rates. The NLSY IQ-incarceration correlation has been interpreted as evidence that inherited low IQ is a cause of criminal offending (Gottfredson, 1998). The coherence viewpoint casts doubt on this causal interpretation, because inherited IQ cannot explain crime trends, and IQ is largely stable after childhood whereas criminal offending peaks around ages 15–25 years. Adult white matter growth has been linked to a reduced risk of impulsive behavior as people age (Bartzokis et al., 2001) and juvenile delinquency linked to the impact of preschool lead exposure on impulsivity (Needleman et al., 2003). Preschool lead exposure causes gray matter damage linked to IQ loss, and white matter damage related to behavior, suggesting the NLSY IQ–incarceration correlation could reflect separate lead effects. IQ loss may be permanent, but white matter growth to age 50 could ameliorate lead-induced damage, which could explain lower offending rates for older adults, and the impact of preschool blood lead on national crime rates with best-fit time lags that reflect the typical age of offenders (Nevin, 2007). Coherent trends in blood lead, education, and crime are now evident in 2000–2006 incarceration rates by age and race (Sabol et al., 2007; Beck and Harrison, 2001). Fig. 9 shows incarceration rates fell from 2000 to 2006 for all men under age 30, reflecting the decline in preschool blood lead since the mid-1970s, but incarceration rates are rising for men over 40, born when lead poisoning was epidemic. The overall 30–39-year-old male incarceration rate rose slightly from 2000 to 2006, but the 30–39-year-old rate for black males fell 12%, reflecting the same slum clearance birth years associated with a racial convergence in MR, SAT, and NAEP scores, and juvenile burglary arrest rates.

(8) Experimental evidence: If, “because of an observed association some preventive action is taken ... and it does” in fact prevent ... Here the strongest support for the causation hypothesis may be revealed” (Hill, 1965). Regulations to eliminate the use of lead in gasoline and to reduce lead paint hazards in older housing anticipated societal benefits in excess of regulatory costs because of observed associations between preschool lead poisoning and MR, between even small elevations in blood lead and IQ, and between IQ and educational attainment and lifetime earnings. Ecological trends examined here suggest the magnitude of preschool

![Fig. 9. 2000–2006 Percent change in age-specific male incarceration rates. Declining incarceration rates from 2000 to 2006 for males under 30 reflect declining preschool blood lead since the mid-1970s. Incarceration rates are still rising for those over age 40, born when childhood lead poisoning was epidemic. The overall age 30–39 male incarceration rate rose slightly from 2000 to 2006, but the age 30–39 rate for black males fell 12%, reflecting slum clearance birth years also associated with a racial convergence in MR prevalence and juvenile burglary arrest rates (Sabol et al., 2007; Beck and Harrison, 2001).](image-url)
lead exposure's impact on MR and education achievement and attainment trends may have been underestimated in the past, but the direction of that impact was recognized as the basis for deliberate regulatory interventions that have achieved predicted results.

(9) Analogy: “With the effects of thalidomide and rubella before us we would surely be ready to accept slighter but similar evidence with another drug or another viral disease in pregnancy” (Hill, 1965). Similarly, the effects of blood lead trends before us should lend urgency to eliminating the remaining risk of preschool lead exposure and blood lead levels once considered safe. Some nations still use leaded gas, and children and pregnant women in many nations are still exposed to industrial lead emissions, occupational lead hazards, lead-glazed ceramics, and lead-contaminated home remedies and cosmetics. Lead paint hazards are the greatest remaining USA risk, and a simple window replacement strategy would yield lead hazard reduction benefits plus energy savings from efficient windows that far exceed window replacement costs (Jacobs and Nevin, 2006; Nevin and Jacobs, 2006; Nevin et al., 2007).

(10) Bradford Hill concluded by addressing The Case for Action, with words that clearly apply to the case for moving aggressively to eliminate all preschool lead exposure:

All scientific work is incomplete—whether it is observational or experimental. All scientific work is liable to be upset or modified by advancing knowledge. That does not confer upon us a freedom to ignore the knowledge we already have, or to postpone the action that it appears to demand at a given time (Hill, 1965).

References


