Intelligence 38 (2010) 367–384

Contents lists available at ScienceDirect

Intelligence

The Flynn Effect within subgroups in the U.S.: Gender, race, income, education, and urbanization differences in the NLSY-Children data

SiewChing Ang [a], Joseph Lee Rodgers [a,*], Linda Wänström [b]

[a] Department of Psychology, University of Oklahoma, United States
[b] Department of Statistics, University of Stockholm, Sweden

ARTICLE INFO

Article history:
Received 6 July 2009
Received in revised form 21 May 2010
Accepted 22 May 2010
Available online 20 June 2010

Keywords:
Flynn Effect
National Longitudinal Survey of Youth
NLSY-Children data
PIAT-Math
Secular changes in intelligence

ABSTRACT

Although the Flynn Effect has been studied widely across cultural, geographic, and intellectual domains, and many explanatory theories have been proposed, little past research attention has been paid to subgroup differences. Rodgers and Wänström (2007) identified an aggregate-level Flynn Effect (FE) at each age between 5 and 13 in the Children of the National Longitudinal Survey of Youth (NLSYC) PIAT-Math data. FE patterns were not obtained for Reading Recognition, Reading Comprehension, or Digit Span, consistent with past FE research suggesting a closer relationship to fluid intelligence measures of problem solving and analytic reasoning than to crystallized measures of verbal comprehension and memory. These prior findings suggest that the NLSYC data can be used as a natural laboratory to study more subtle FE patterns within various demographic subgroups. We test for subgroup Flynn Effect differences by gender, race/ethnicity, maternal education, household income, and urbanization. No subgroups differences emerged for three demographic categories. However, children with more educated (especially college educated) mothers and/or children born into higher income households had an accelerated Flynn Effect in their PIAT-M scores compared to cohort peers with lower educated mothers or lower income households. We interpret both the positive and the null findings in relation to previous theoretical explanations.

1. Introduction

The Flynn Effect (FE) refers to an increase in intelligence scores across time. The effect was first reported by Lynn (1982) and especially by Flynn (1984), and is commonly known as the Flynn Effect (though it also been referred to as the Lynn–Flynn Effect; see Rushton, 1997). Flynn (2006) observed that increasing IQ scores were documented much earlier, though without much broad attention (e.g., Smith, 1942; Tuddenham, 1948). Despite substantial modern attention to the FE in the intelligence literature (see, e.g., Neisser, 1998 and Flynn, 2007 for summaries), surprisingly little attention has been given to certain critical features of the FE, including how it operates among different population subgroups (Rodgers, 1998; Rodgers & Wänström, 2007; Sundet, Borren & Tambs, 2008). This paper examines how the FE performs within several important demographic subgroups – gender, race/ethnicity, urbanization, mother's education, and household income – using the PIAT-Math subscale from the Children of the National Longitudinal Survey (NLSYC) data.

Many Flynn Effect studies have been reported on various populations worldwide (Colom, AndresPueyo & JuanEspinosa, 1998; Colom, JuanEspinosa & García, 2001; Daley, Whaley, Sigman, Espinosa & Neumann, 2003; Flynn, 1987; Lynn, 1982; Lynn & Hampson, 1986; Rodgers & Wänström, 2007; Sundet, Barlaug & Torjussen, 2004; Teasdale & Owen, 1989, 2000). Although IQ scores have risen systematically, it is questionable whether general intelligence (g) itself has changed (Jensen, 1998; Kane & Oakland, 2000; Must, Must & Raudik, 2003; Flynn, 1987, 2006). Further, the FE appears to operate more strongly within the fluid intelligence domain, and not as much within crystallized intelligence (Flynn, 1987, 2006; Jensen, 1991; Loehlin, 1996; Lynn, 2009; but also see Flynn, 2009b, who identified Flynn Effects on the vocabulary subscale of the WAIS).

* Corresponding author.
E-mail address: jrodgers@ou.edu (J.L. Rodgers).

0160-2896/$ – see front matter
doi:10.1016/j.intell.2010.05.004

Fluid intelligence refers to problem solving or reasoning, whereas crystallized intelligence refers to language acquisition and knowledge attainment (Horn & Catell, 1966).

Many causal explanations of the FE have been proposed, though consensus has not emerged, and some controversy has ensued in regards to certain theories. Flynn Effect scholars have attributed the FE to massive environmental changes, such as the proliferation of movies, television, video games and computers (Greenfield, 1998); urbanization (Flynn 1998; Williams, 1998); increased access to education (Flynn, 1998; Greenfield, 1998); general school factors (Schooler, 1998; Williams, 1998); pre-school education (Teasdale & Berliner, 1991); the development of new math education/curriculum (Blair, Gamsonb, Thorne & Baker, 2005); the development of quality and quantity of schools, teachers and teacher-training programs with increased educational funding (Greenfield, 1998; Williams, 1998); and changes in parental styles (Williams, 1998). Mahlberg (1997) proposed a FE explanation based on a cultural memory interpretation. Studies have also connected the FE with changes in family size (resulting in increased resources per child; Sundet et al., 2008), and increased parental educational levels (Ceci, 1996; Grissmer, Kirby, Berends and Williamson, 1994; Sundet et al). Nutritional mechanisms have provided the basis for another class of popular theories (Lynn, 1989, 1990; Martorell, 1998; Schoenthaler, Amos, Eysenck, Peritz and Yudkin, 1991); Flynn (2009a) provided evidence in opposition to a nutritional interpretation. Steen (2009) favored a "rising tide" hypothesis associated with medical improvements, "because of intervention into medical conditions that depressed intelligence in the past" (p. 129).

Various other theories have used processes that were not purely environmental. Lynn (1998) suggested there may exist a genetic component to the FE. An interpretation by Dickens and Flynn (2001) involved an evocative gene–environment interaction model. Mingroni (2007) suggested that heterosis (also known as hybrid vigor or outbreeding) was a plausible explanation (also see Jensen, 1998).

Others have suggested that IQ increases are artifactual. This class of explanations includes indications of factorial invariance violations across time/ages of the measures used to document the FE, identified by Wicherts et al (2004) and Beaujean & Osterlind (2008). (Both of these studies also identified real ability changes as well, however.) Testing issues have also been identified, including increased attention to speeded tests (Brand, 1996; Brand, Freshwater and Dockrell, 1989), and test norming processes (Kanaya, Scullin & Ceci, 2003). Practice effects have been implicated (Williams, 1998), along with other environmental factors (Blake, 1989).

Finally, Jensen (1998) suggested a multiplicity hypothesis, in which many small factors (some likely unspecified) have combined to create the FE. A particular attraction of this interpretation is that it provides explanation for the consistency of the FE in many cultures. The pace and direction of the FE should change substantially over time if one or a very few explanations were accounting for it and if those influences themselves changed over time. The multiplicity hypothesis would explain the persistence of the FE in many geographic settings, even as there have been temporal shifts in a number of its individual putative causes.

Although various causal explanations for the FE have been proposed in the literature, more empirical understanding of the FE is needed before the theories to explain the FE are fully informed by the data patterns. Rodgers (1998) outlined 10 research questions to be answered to support causal evaluation of the FE, one of which focused on identifying differences across demographic categories. A decade later, few of these questions have been yet addressed, and the status of the FE across demographic subgroups is still unclear. Flynn (1998) and Greenfield (1998) studied income differences. Rushton (1999, 2000) considered the FE in relation to race differences. Research by Teasdale and Owen (1989) and by Colom, Lluis-Font and Andres-Pueyo (2005) identifying differential FE in the lower tails of the intelligence distributions had implications for subgroup differences (though indirect). Little other research addressing, for example, basic gender, race, and education differences has been conducted.

The current study is a replication and expansion of the work done by Rodgers and Wänström (2007). Using the NLSYC data from 1986 to 2000, they compared various cognitive assessments at each age from five to 13 years old, using all biological children born to a 1978 sample of mothers whose ages were 36–43 years old on December 31, 2000 (birth cohorts of 1957–1964). The cognitive assessments included were the Wechsler Memory for Digit Span test, the Peabody Picture Vocabulary Test, and three Peabody Individual Achievement (PIAT) Test subscales: PIAT-Math (PIAT-M), PIAT-Reading Recognition and PIAT-Reading Comprehension. Results indicated that the PIAT-M showed the largest FE (the mean slope per year was .30 for the oldest five ages) compared to other subscales. This effect was statistically significant for almost all ages even after controlling for mother's IQ (which was included to adjust for selection bias due to younger mothers having had disproportionately more children). For most other subscales, the FE decreased or disappeared entirely after adjusting for maternal IQ. This result was predicted, because the PIAT-M is much closer to a measure of fluid intelligence, whereas the other tests more strongly reflect crystallized intelligence.

Although the Flynn Effect has been empirically shown in the NLSYC PIAT-M scores, a number of interesting questions remain unanswered. One question, motivated by the critique in Rodgers (1998) over a decade ago, is: Are there differences in these patterns within the subpopulation groups, such as gender, race/ethnicity, education, income, or urbanization? Specifically, this paper focuses on the exploration of the FE within each of these categories over an 18-year period, from 1986 to 2004.

2. Methods

2.1. Sample and designs

The original National Longitudinal Survey of Youth (NLSY79), which contains the mothers of the NLSYC children sample that we use in this research, was based on a household probability sample of 12,686 adolescents aged 14–21 on Dec 31, 1978 (http://www.bls.gov/nls/nlsy79.htm). Since 1986, on a biannual basis, all biological children born to the female NLSY79 respondents have been surveyed, including the administration of cognitive assessments. By 2004, the NLSY79 females were between 39 and 47 years old, and had given birth to 15,359 children. These children's ages ranged between newborn to the mid 30's. Our design involves using the longitudinal structure of the NLSYC to compare children of the same age across time—five-year-olds in 1984 to five-