Intelligence 48 (2015) 51–57

Contents lists available at ScienceDirect

Intelligence

Implications of the Flynn effect for age–cognition relations

Timothy A. Salthouse

Department of Psychology, University of Virginia, Charlottesville, VA 22904, USA

ARTICLE INFO

Article history:
Received 28 August 2014
Received in revised form 3 October 2014
Accepted 16 October 2014
Available online 11 November 2014

Keywords:
Flynn effect
Longitudinal
Cross-sectional
Aging
Cognition

ABSTRACT

Many studies have documented that cognitive performance is often higher among people of the same age who are tested in more recent years, and it is sometimes suggested that this phenomenon will distort the relations between age and cognition in cross-sectional studies. This possibility was examined with data from two large projects involving adults across a wide age range. The results indicated that there were similar time-of-measurement increases in cognitive scores at different ages, which were accompanied by nearly constant cross-sectional age differences, but positively inflated estimates of longitudinal age differences. It is proposed that when the Flynn effect is of comparable magnitude in adults of different ages, longitudinal comparisons of age–cognition relations are more subject to distortion than cross-sectional comparisons.

The Flynn effect (named after James Flynn but originally described many years earlier, cf. Lynn, 2013) refers to the phenomenon that people of the same age who are tested in more recent years tend to have higher scores on cognitive tests than people tested in earlier years. Although many questions remain about the mechanisms for the effect, and its generality across ability domains, cultures, and historical periods, the basic phenomenon has been widely replicated and can be considered to be well established (Trahan, Stuebing, Fletcher, & Hiscock, 2014; Williams, 2013).

A number of researchers have postulated that the Flynn effect will lead to a distortion of cross-sectional relations between age and cognition (e.g., Baxendale, 2010; Hiscock, 2007; Ronnlund & Nilsson, 2009). For example, Flynn stated that "cross-sectional data, as a measure of the effects of aging on IQ, are suspect. … Cross-sectional data compare, for example, 80-year-old subjects with a group of 20-year-old subjects, with both groups being tested at the same time. This makes sense only if current 20-year-olds have the same IQ as 20-year-olds did two generations ago, that is, when today's 80-year-olds were 20" (Flynn, 1987 p. 187).

However, the thesis of this article is that the implications of the Flynn effect for both cross-sectional and longitudinal relations between age and cognition depend on whether the Flynn effect represents a cohort effect or a period (time-of-measurement) effect. Consider the definitions of these terms provided by Schaie (2013)

> "… cohort effects represent the impact of historical effects on a group of individuals who share similar environmental circumstances at equivalent points in their maturation sequence … On the other hand, time-of-measurement effects represent those events that have an impact on all members of the population experiencing a common historical exposure, regardless of cohort membership (p. 25)."

> "Period effects would be ones caused by one or more social innovations that act equally at a point in time on all individuals, regardless of age … A cohort effect would be one acting on children or adults of a particular age, persisting across time (p. 340)."

Based on these definitions, it can be inferred that the distinguishing feature of a period effect is that time-of-measurement influences are similar in people of all ages, and do not vary according to birth year or cohort. Fig. 1 illustrates a situation of this type with cognitive test scores plotted as a function of time

E-mail address: Salthouse@virginia.edu.

http://dx.doi.org/10.1016/j.intell.2014.10.007
0160-2896/

of measurement. The thick dashed line represents period influences, which are portrayed as progressively more positive in successive test years. The vertical rectangles correspond to cross-sectional comparisons, and it can be seen that if the period effects are similar at each age, cross-sectional age differences would be expected to be approximately parallel at each test year. That is, even if the absolute level of performance is higher in successive test years, differences in cognitive tests scores between 25-year-old, 45-year-old, and 65-year-old participants would be expected to be comparable in the 2000, 2005, and 2010 test years. In contrast, if the period effects varied with age, possibly with greater time-related improvements at younger ages, cross-sectional age differences would be expected to be larger in more recent test years.

Inspection of Fig. 1 reveals another implication of the Flynn effect for age–cognition relations, namely, that positive time-of-measurement effects can lead to a distortion of longitudinal comparisons. That is, if the factors contributing to higher scores on more recent test years in different people (portrayed by the thick dashed lines) also operate within the same people (portrayed by the dotted diagonal boxes), then longitudinal comparisons will likely be inflated by the presence of the Flynn effect. In other words, because in longitudinal designs assessments at successive ages necessarily occur in more recent years, some of the age-related longitudinal differences in cognitive performance may be attributable to positive period effects. Moreover, although the distortion of the longitudinal comparisons will be larger in later birth cohorts if the period effects are greater at younger ages, some distortion will be evident whenever period effects are positive.

Schaie (e.g., 2013, pp. 192–193) recognized that longitudinal comparisons might be influenced by positive time-of-measurement effects and proposed that adjusted longitudinal change could be estimated by subtracting the estimated time-of-measurement effect from the observed change. Initial analyses of this type were reported in Salthouse (1991), but they were limited by the data available at that time.

There were three major goals of the current study. The first was to investigate whether the magnitude of the Flynn effect was similar at different ages in adulthood. The second goal was to examine cross-sectional comparisons in different test years

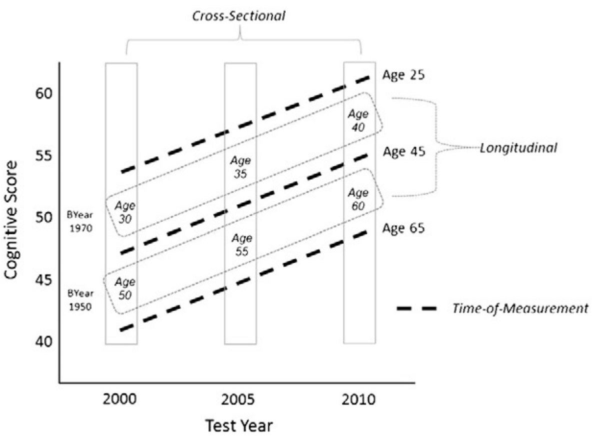

Fig. 1. Data matrix illustrating relations among time-of-measurement (Flynn Effect) influences and cross-sectional and longitudinal age comparisons.

to determine whether the relative age differences were remaining constant, or increasing over time. The third goal was to estimate longitudinal change after adjusting for positive time-of-measurement effects. The rationale was that because they are theoretically independent of particular social or environmental conditions, these estimates of adjusted change may more accurately reflect age trends in cognitive functioning. The analyses were based on published summary data from the Seattle Longitudinal Study reported in Schaie (2013), and summary data from the Betula project reported in Ronnlund and Nilsson (2008, 2009) and Ronnlund, Nyberg, Backman, and Nilsson (2005).

1. Seattle Longitudinal Study (SLS)

Participants in the SLS were recruited from a Health Maintenance Organization, with similar recruitment procedures each year (Schaie, 2013, pp. 37–38). Many of the participants returned for repeat testing at 7-year intervals, and thus, the data were organized into 7-year age-groups. The cross-sectional sample with the primary cognitive battery consisted of 4,850 adults, of whom 2,777 returned for a 7-year longitudinal assessment (43% attrition). The cross-sectional sample with the latent constructs consisted of 2,038 adults (Schaie, 2013, pp. 38, 43), of whom 1,257 returned for a 7-year longitudinal assessment (38% attrition).

The same five tests were administered to new samples of adults between 22 and 77 years or older from 1956 to 1998, with between 500 and 997 new participants recruited each test year. The primary tests (i.e., series completion reasoning, spatial rotation, number arithmetic, multiple-choice vocabulary, and word fluency) were described by Schaie (2013, pp. 52–55), and all had time limits between 4 and 6 min each. Beginning in 1984, new tests were added to the assessment battery, and analyses were reported at the level of latent constructs based on factor scores across two or more tests. These tests also had time limits ranging from 1.5 to 6 min.

The cognitive scores were reported in *T*-score units (mean of 50, standard deviation of 10) based on the initial assessment of the complete sample of 4,850 across all test years for the five primary tests, and on the sample of 2,038 for the latent constructs. Data for the primary variables were obtained from Table 4.2 of Schaie (2013), and data from the latent constructs from Table 4.4. The observed 7-year longitudinal changes across all test years were obtained from Table 5.1 for the primary variables, and from Table 5.10 for the latent constructs.

2. Results

Reasoning scores of participants in four age-groups across the seven test years are portrayed in Fig. 2. Only four ages are illustrated for clarity, but the pattern was similar at all ages. Note that there were nearly parallel increases in reasoning performance as a function of test year at each age.

Regression analyses were conducted to predict the scores of each cognitive measure with age, test year, and their interaction (based on the cross-product of centered age and test year variables) as predictors. (Quadratic test-year effects were also examined, but they were not significant, and thus were not included in the final equations.)

The results of these analyses are presented in the top panel of Table 1, where it can be seen that most of the age coefficients were negative, indicating lower scores at older ages.