Personality and Individual Differences 54 (2013) 871–876

Contents lists available at SciVerse ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

In the Netherlands the anti-Flynn effect is a Jensen effect

Michael A. Woodley [a,b,*], Gerhard Meisenberg [c]

[a] Umeå University, Department of Psychology, Umeå, Sweden
[b] Vrije Universiteit Brussel, Center Leo Apostel for Interdisciplinary Studies, Brussels, Belgium
[c] Ross University School of Medicine, Department of Biochemistry, Dominica

ARTICLE INFO

Article history:
Received 4 November 2012
Received in revised form 7 December 2012
Accepted 14 December 2012
Available online 8 February 2013

Keywords:
Anti-Flynn effect
Dysgenics
Flynn effect
Jensen effect
Co-occurrence model

ABSTRACT

In this study, 63 observations of secular IQ changes (both Flynn and anti-Flynn effects) are collected from three demographically diverse studies of the Dutch population for the period 1975–2005 (representing the 1950–1990 birth cohorts), along with data on *g* loadings and subtest reliabilities. The method of correlated vectors is used to explore the association between Flynn and anti-Flynn effect magnitudes, both independently and together, and the *g* loadings of subtests. Despite a positive vector correlation the Flynn effects are not associated with the Jensen effect ($r = .307$, *ns*, $N = 36$), however the anti-Flynn effects are ($r = .406$, $P = .05$, $N = 27$). Combined, the vector correlation becomes negative but non-significant ($r = -.111$, *ns*, $N = 63$). Declines due to the anti-Flynn effect are estimated at −4.515 points per decade, whereas gains due to the Flynn effect are estimated at 2.175 points per decade. The *N*-weighted net of these is a loss of −1.350 points per decade, suggesting an overall tendency towards decreasing IQ in the Netherlands with respect to these cohorts. The Jensen effect on the anti-Flynn effect suggests that it may be related to bio-demographic changes within the Netherlands which have reduced 'genetic-*g*', despite the presence of large, parallel gains on subtests that may be relatively more sensitive to cultural-environmental improvements.

1. Introduction

Flynn effects are associated with an increase in IQ of approximately three points a decade. They have been detected in cohorts dating back to the early decades of the 20th century, but in continental Europe and Japan were most pronounced in the post-war period (Flynn, 2009a). Flynn effects do not occur equally throughout the full range of abilities, as they are most pronounced on measures of fluid intelligence (such as the Raven's matrices – where gains of between five and seven points per decade have been detected), and are typically least pronounced on measures of crystallized intelligence (where they typically occur at a rate of between one and three points per decade depending on the measure (Flynn, 2009a)). Despite the affinity of the effect for apparently highly *g*-loaded measures of fluid intelligence, such as the Raven's matrices, there is debate about whether the secular gains actually occur on *g* or not. Jensen's (1998) method of correlated vectors permits the Flynn effect's affinity with *g* to be determined by simply correlating the magnitude of the gains with the *g* loadings of the subtests on which they occur. When a vector correlation is found to be significantly positive (i.e. where the association between IQ and a source of individual or group differences is more pronounced on subtests exhibiting higher *g* loadings), the source in question can be described as a Jensen effect (Rushton, 1998). Jensen effects are remarkable in that they principally concern sources of individual and group differences that are either wholly or in part biological in origin. These include race and sex differences (Jensen, 1998; Nyborg, 2005; Rushton & Jensen, 2010), inbreeding depression, reaction times, evoked potentials, brain pH, subtest heritabilities (Jensen, 1998; Rushton, 1999; Rushton & Jensen, 2010), brain size (Rushton & Ankney, 2009), fluctuating asymmetry (Prokosch, Yeo, & Miller, 2005) and dysgenic fertility differentials (Woodley & Meisenberg, in press) amongst others. The preponderance of studies indicate that the Flynn effect is not a Jensen effect, in that its effect magnitude is typically either uncorrelated or negatively correlated with the *g* loadedness of the subtests on which it occurs (Must, Must, & Raudik, 2003; Rushton, 1999; Rushton & Jensen, 2010; te Nijenhuis, in press; te Nijenhuis & van der Flier, 2007). Given that *g*-loadings vary positively monotonically with subtest heritabilities (Rushton & Jensen, 2010), this finding is consistent with the idea that the rapidity of secular gains indicates a primarily cultural–environmental cause for the effect (Neisser, 1997). The Flynn effect therefore resembles to some extent other effects that are purely cultural–environmental, such as IQ gains accrued from retesting the magnitudes of which are inversely related to *g* subtest loading (te Nijenhuis, van Vianen, & van der Flier, 2007).

* Corresponding author at: Umeå University, Department of Psychology, Umeå, Sweden. Tel.: +1 44 1798 872 021.
E-mail address: M.A.WoodleyPhD@gmail.com (M.A. Woodley).

0191-8869/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.paid.2012.12.022

Recently, a wholesale reversal of the Flynn effect (or anti-Flynn effect) has been detected in some countries. These include the UK amongst 14–15 year olds on the Raven's matrices (Flynn, 2009b), and amongst students assessed on Piagetian type intelligence tests (Shayer & Ginsburg, 2009; Shayer, Ginsburg, & Coe, 2007), and also in Denmark and Norway amongst military conscripts (Sundet, Barlaug, & Torjussen, 2004; Teasdale & Owen, 2005, 2008). Many more studies have found anti-Flynn effects on specific subtests, frequently occurring in tandem with Flynn effects on other subtests. Example countries include the Netherlands (te Nijenhuis, in press; te Nijenhuis & van der Flier, 2007; Wicherts et al., 2004), the US (Kanaya & Ceci, 2010), Sudan (Khaleefa, Sulman, & Lynn, 2008), Estonia (Must et al., 2003; Must, te Nijenhuis, Must, & van Vianen, 2009) and Spain (Colom, Juan-Espinosa, & García, 2001). Theories as to the cause of anti-Flynn effects can be split into four main categories: cultural–environmental, statistical, biological and hybrid, which combine different explanations:

(1) *Cultural–Environmental* explanations posit that the effect is basically the mirror image of the Flynn effect in that it is caused by diminishing (as opposed to improving) cultural and environmental quality (Russell, 2007).
(2) *Statistical* explanations revolve around the idea that the effect results from factors such as items becoming obsolete, and thus harder in a way that doesn't necessarily reflect *g* (Teasdale & Owen, 2000), or alternatively from the existence of error stemming from improper sampling (Wicherts et al., 2004).
(3) *Biological* explanations are based on the idea that dysgenic bio-demographic factors have reversed the Flynn effect (Lynn & Harvey, 2008; Nyborg, 2012; Teasdale & Owen, 2008). Dysgenics describes the expectation that IQ (and other socially valuable traits) diminish over time given the presence of reproductive patterns unfavorable to the maintenance of constant or increasing levels of the trait within a population (Galton, 1869). There are two routes through which dysgenics can operate. The first is associated with the presence of negative correlations between IQ or its proxies and fertility *within* populations. It has been observed that throughout the whole of the 20th century in the West, the fertility trend with respect to IQ and its proxies (such as education and socio-economic status) has been dysgenic (Lynn, 2011; Nyborg, 2012; Skirbekk, 2008). The second involves migration from higher-fertility, lower-IQ countries to lower-fertility, higher-IQ ones. This process of 'replacement migration' (Coleman, 2002) has been occurring to an increasing degree in the West since the 1950's.
(4) A *Hybrid* effect has also been proposed, based on the idea that small dysgenic losses in IQ might have diminished cultural–environmental quality in such a way that amplifies the losses further (Meisenberg, 2003; Woodley, 2012a).

A technique for disentangling these effects is the method of correlated vectors. Primarily cultural–environmental driven anti-Flynn effects are likely to simply be Flynn effects 'in reverse' hence the less heritable and also less *g* loaded abilities should be declining as rapidly or more rapidly than those with higher *g* loadings. In this case anti-Flynn effects should exhibit negligible or negative vector correlations with *g*.

Primarily statistically driven anti-Flynn effects resulting from factors such as item obsolescence or inadequate sampling should not be associated with the Jensen effect either, as the first effect should be most pronounced on items and subtests whose *g* loadings are depleting with time (i.e. if a word is sufficiently obsolete, knowledge of it won't discriminate well between different levels of *g*), whereas the second effect should be indiscriminate with respect to *g* loadedness.

Wholly or partly biologically driven anti-Flynn effects should on the other hand be associated with the Jensen effect, as both dysgenic fertility differentials within populations and ethnic differences in IQ are strongly associated with the effect (Rushton & Jensen, 2010; Woodley & Meisenberg, in press). If Flynn and anti-Flynn effects can be shown to be distinct with respect to the Jensen effect in a way that is indicative of some degree of biological causation for the latter, this would provide a strong piece of supporting evidence for the 'co-occurrence model' of dysgenesis and the Flynn effect developed in Woodley (2012a, 2012b). The 'co-occurrence model' posits that rather than large 'phenotypic' gains in IQ via the Flynn effect simply masking smaller 'genotypic' losses due to dysgenics (Lynn, 2011), Flynn effects and dysgenic effects actually co-occur, albeit concentrated on subtest specific sources of variance in the case of the former, and on 'genetic *g*' (Rushton and Jensen's (2010) term for the monotonic Jensen effect on subtest heritabilities) in the case of the latter. In addition to the aforementioned finding that dysgenic fertility differentials (which are not observed IQ declines) are strongly associated with the Jensen effect, unlike Flynn effects, other indirect evidence for this 'co-occurrence' model comes from the study of the temporal correlates of reconstructed genotypic and Flynn effect trends in IQ. Specifically, eugenic or dysgenic trends in IQ appear to predict the per-capita rate of major innovations over 55 decades, while the Flynn effect appears to be related instead to per-capita GDP growth (Woodley, 2012a). This indicates that by assuming co-occurrence, both effects can be regarded as plausible determinants of real and distinct trends in the social evolution of Western societies.

2. Methods

For this analysis, four Dutch studies were selected in which both Flynn and anti-Flynn effects are reported on different subtests and in different samples, spanning 1975–2005. The demographic diversity of these samples indicates that they can be considered broadly representative of cohorts born in the period 1950–1990. The four studies selected include te Nijenhuis (in press), which reports IQ gains/losses for a sample of 1091 applicant bus drivers in the period 1983/85, compared with 110 sampled in the period 1975/76, and 212 sampled in the period 1988/92, compared with the 1975/76 cohort (study 1). It also reports these data for a sample of 270 16-year-old students, sampled in 1985, compared with a sample of 130 sampled in 1975, and 498 in 2005, compared with the 1975 cohort (study 2). All samples were assessed using the Dutch translation of the General Aptitude Test Battery (DGATB).

The third study is sourced from Wicherts et al. (2004), who analyzed student performance on the Dutch translation of the Differential Aptitude Test (DAT) between the years 1984 and 1994/95. The 1994/95 cohort was split into three ability groups – MAVO (lowest, $N = 272$), HAVO (intermediate, $N = 397$) and VWO (highest, $N = 188$). These were then compared against the whole 1984 cohort ($N = 1100$).

The fourth study is also sourced from Wicherts et al. (2004), and involves two cohorts of Dutch five year olds tested in 1981/82 ($N = 207$) and 1992/93 ($N = 415$). Wicherts et al. (2004) further split the cohorts based on age standardizations conducted at 59–62 months ($N = 208$) and 63–71 months ($N = 207$). These cohorts were tested using the Revise Amsterdamse Kinder Intelligentie Test (RAKIT).

Data on the *g* loadings of the DGATB are provided by te Nijenhuis (in press – see also references contained therein). Student *g* loadings for each of the subtests of the Dutch DAT are sourced from te Nijenhuis, Evers, and Mur (2000). The *g* loadings of the RAKIT