



Speed and the Flynn Effect

Olev Must*, Aasa Must

Institute of Education, University of Tartu, Salmel 1a, 50103 Tartu, Estonia



ARTICLE INFO

Keywords:
Flynn Effect
NIT
Speed
Tork
Estonia

ABSTRACT

We investigated the role of test-taking speed on the Flynn Effect (FE). Our study compared two cohorts of Estonian students (1933/36, $n = 888$; 2006, $n = 912$) using 9 subtests from the Estonian adaptation of the National Intelligence Tests (NIT). The speededness of the items and the subtests was found by determining the proportion of unreached items from among the total number of errors (Stafford, 1971). The test-taking speed of the younger cohort was higher in all 9 of the subtests. This suggests that the younger cohort is able to solve more items than the older one. The lack of measurement invariance at the item and subtest level was quantitatively estimated using a method proposed by Dimitrov (2017). The test-taking speed and the non-invariance of the items was strongly, yet inversely correlated (up to -0.89). The subtests versions that consisted of only invariant items showed no, or a small positive, FE. The subtest versions consisting of only speeded items showed a large positive FE, with cohort differences of up to 50%. If the requirement of measurement invariance is ignored then this effect becomes apparent. The rise in test-taking speed between cohorts can be attributed to an increase in automated responses, which is an outgrowth of modern education (differences in the mandatory age of school attendance, and in the student's readiness to solve abstract items also affected the test-taking speed of the cohorts). We were able to conclude that the younger cohort is faster than the older one.

Recent meta-analysis of IQ test score changes over > 100 years shows an increase of about 3 IQ points per decade. This rise seems to be holding over time, but the strength of the gains has varied by country, intelligence domains, and the time period being studied. Moreover there is also evidence suggesting that IQ test score gains have actually been declining in the latter part of the 20th century (Flynn & Shayer, 2018; Pietschnig, 2016; Pietschnig & Voracek, 2015). Explanations for these processes are still ongoing.

The relationship between the level (power, ability) and speed of intellectual processes has been acknowledged, but is not well understood, nor has it been thoroughly studied. The level/speed relationship poses a special challenge for the empirical researcher, as it frequently becomes necessary to separate the level and speed within the same construct, such as within a test-taker's responses to the items. As classical psychological concepts, the level and speed components make it possible to approach the Flynn Effect (FE) from a new perspective, and to demonstrate that the FE may actually be the manifestation of a change in test-taking speed. This rise in test-taking speed can possibly be attributed to changes in educational practices. Over time modern education has increasingly come to emphasize decontextualized and abstract problem-solving in speeded situations. This may be one of the factors driving the faster responses to IQ test items as well.

1. Introduction

Galton (1883) believed that reaction time (RT) could be used as an elementary cognitive measure of mental processing speed. Jensen (2006) maintained that the main source of g -differences is related to the speed of information processing. The higher the processing speed, the greater the intellectual capacity. The speed with which an individual is able to respond to a test item, has attracted the attention of IQ-theorists for more than a century (Binet & Simon, 1905; for an overview see: Nicolas, Andrieu, Croizet, Santitioso, & Burman, 2013). The discussions pertaining to speed and level were common from the 1920's to the 1950's, and during that time the concepts of level and speed as natural qualities of intelligence were explored by several authors (e.g. Baxter, 1941; Davidson & Carroll, 1945; Freeman, 1931; Kelley, 1927; Lord, 1954; Peak & Boring, 1926; Spearman, 1927; Thorndike, Bregman, Cobb, & Woodyard, 1927). One researcher from that era, Baxter (1941), concluded that speed and level vary independently as factors of intelligence. A decade later Gulliksen (1950) offered a new and more detailed approach by introducing the concept of pure speed and pure power tests to measure mental abilities. According to Gulliksen a pure speed test would consist of items that would be so easy to answer, any test-taker could respond to all of them correctly. A pure power test would be one where all the items could be completed and the resulting

* Corresponding author.

E-mail address: Olev.Must@ut.ee (O. Must).

score would depend upon the number of items that were properly answered. The various concepts and methodological attempts related to time use and test-taking speed have evolved from this approach.

In modern psychology the discussion regarding the role of time in the measurement of cognitive abilities is ongoing and the discourse concerning the relationship between speed and power, speed and level, and speed-error trade-off are continuations of an old debate (De Boeck, Chen, & Davison, 2017; Goldhammer, 2015; Goldhammer & Klein Entink, 2011; Klein Entink, Fox, & van der Linden, 2009; Kyllonen & Zu, 2016; Partchev & De Boeck, 2012; Partchev, De Boeck, & Steyer, 2011; Ren, Wang, Sun, & Schweizer, 2018). There are several authors who take the position that speed and power are separate qualities in the measurement of cognitive abilities. The research of Wilhelm and Schulze (2002) showed that speeded and unspeeded tests, which measure reasoning ability, do not equally assess the same constructs. Their results indicate that the variance of a speeded reasoning test can be more thoroughly explained via the linear function of unspeeded reasoning and mental speed. They concluded, that the use of speeded reasoning tests will more likely lead to overestimations of the relationship between mental speed and reasoning ability. Power and speed cannot be directly observed and must be derived from observable factors.

Goldhammer and Klein Entink (2011) showed that reasoning speed is a unidimensional construct that represents significant individual differences and that reasoning speed and ability are clearly distinguishable constructs. However, their relationship may be dynamic. An understanding of the speed-accuracy trade-off (Dennis & Evans, 1996) is useful for recognizing that the relationships between power and speed may take several forms, both at the within-individual level, as well as at the between-individuals level. The trade-off suggests that a test taker will operate at a certain (effective or exhibited) level of speed and ability depending on the circumstances. For example, in one situation, the test-takers may work accurately and slowly, whereas in another they may work quickly, but make a great number of errors. Both conditions will result in the same individual information processing efficiency (Goldhammer, 2015). Dennis and Evans (1996) have concentrated on the speed-error trade-off in psychometric testing. They were able to conclude that the standard test scoring procedures end up giving scores that are strongly affected by the candidate's compromise between speed and error. In his analysis of the relationship between level and speed, Goldhammer (2015) uses the term *comparable test-taking behavior* to refer to the individual differences in time-management strategies among test-takers. These different test-taking strategies may consequently result in different testing results. These varied strategies may also explain the FE, as different cohorts may use different time-management strategies, which are based on speed and ability. Goldhammer (2015) perceives speededness as being a central property of a test, and one that reflects the degree to which performance is affected by a time limit.

1.1. Findings regarding speed and its role in the Flynn Effect

Classical descriptions of the FE – the secular rise in IQ scores – often make reference to the one-dimensional conception of IQ. There are also, however, FE studies, that separate level and speed, or in other words, take a two-dimensional approach. Nettelbeck and Wilson (2004) studied the relationship between inspection time (IT) and IQ by using the Peabody Picture Vocabulary Test (PPVT). In their comparisons of student results over a 20 year period they utilized separate measures in order to evaluate the speed and level, and found that although the FE could be detected in the PPVT test results, there was no evidence that processing speed had improved over time. The title of their paper summarizes their main conclusion: smarter, not faster. Their finding suggests that FE gains (smarter) were not related to processing speed, nor to any other factors that have been used to correlate IT to general abilities (Williams, 2013). At the same time, because the rise in IT is

interpreted as being a biomarker for cognitive decline (Deary & Ritchie, 2016; Gregory, Nettelbeck, Howard, & Wilson, 2008), its rise over generations may actually be an indicator of a decline in IQ, or a negative FE.

When Woodley, te Nijenhuis, and Murphy (2013) applied a simple visual reaction time to 14 studies and over 100 years of collected data they found that reaction times were actually rising over time (1889–2004). But as reaction time and IQ are inversely correlated, this was interpreted as being evidence of a secular fall in IQ due to the speed component. Madison, Woodley, and Sanger (2016) used a similar process to assess auditory reaction time, which has actually increased in Sweden from 1959 to 1985. The findings of Woodley et al. (2013) have been criticized due to the fact that there is no evidence of a historical increase in RT after differences between the studies are taken into account (eg. Dodonova & Dodonov, 2013; Silverman, 2013).

The dynamics of test-taking patterns may also reflect the speed component. Brand (1996) originally proposed that test-taking behavior has changed over time. According to his findings the role of guessing during the second half of the 20th century has become more apparent and may account for the rise in test scores. Brand believed that guessing in test-taking is a strategy that is prioritized by modern education in order to cope with the time constraints imposed by many contemporary tests and examinations. The idea behind this hypothesis is that people in modern society have learned to work more efficiently. Brand noticed that for FE “... evidence is drawn largely from short, timed, multiple-choice, group-administered tests of IQ on which there is no adjustment for guessing. Scores on such tests may have improved since 1945 not just because of rising g levels but also because of modern educators' encouragements to children to avoid ‘obsessional’ accuracy and ‘pedantic’ attention to detail. Being composed of different sections, each requiring the use of different principles (e.g. series completion, analogies, oddity), most group tests effectively penalize test-takers who strive for accuracy. Such testees spend valuable time trying to be quite sure they are giving correct answers - rather than making use of guesswork” (Brand, 1996, p. 140). Flynn (1990), however, found that Brand's prediction – that there will be minimal IQ gains on the Wechsler verbal subtests because such tests discourage guessing (no time pressure and the item format is not multiple choice) – was not supported by the evidence, because the rise in these subtests scores was actually the highest.

On the other hand, the research of Woodley, te Nijenhuis, Must, and Must (2014) confirmed Brand's hypothesis that the increase in the number of correct answers of the younger cohorts could in fact be attributed to enhanced guessing. Because the harder items tend to encourage increased guessing, the secular gains in IQ stemming from this Brand effect should be positively associated with a subtest's g loadings. When the Estonian National Intelligence Test (NIT) data, which was collected between 1933 and 2006 and includes data on guessing, was analyzed, it was found that the correlation between the gains via the Brand effect and g loadings was actually very strong: 0.95, thereby corroborating Brand's hypothesis.

The possibility that changes in test-taking patterns have contributed to the rise in speed over time has been recognized by several authors. For example, Wicherts et al. (2004) suggested, that contemporary test-taking strategies may indeed have an effect on the scores of intelligence tests. It is possible that modern test-takers more often resort to guessing than participants in earlier times did, and the strategy results in higher scores on multiple-choice tests. Wicherts et al. also argued that the rise of guessing in more recent cohorts may lead to the negation of measurement invariance of IQ between cohorts. According to the 3 parameter models in Item Response Theory (Hambleton, Swaminathan, & Rogers, 1991), guessing is one of the three most important parameters affecting testing results, together with item difficulty and discrimination.

Must and Must (2013) also did research into the role of guessing and test-taking patterns on the FE. They concluded that rapid guessing has

Download English Version:

<https://daneshyari.com/en/article/7292800>

Download Persian Version:

<https://daneshyari.com/article/7292800>

[Daneshyari.com](https://daneshyari.com)