Learning and Instruction 47 (2017) 43–52

Contents lists available at ScienceDirect

Learning and Instruction

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

Academic tracking is related to gains in students' intelligence over four years: Evidence from a propensity score matching study

Karin Guill [a, *], Oliver Lüdtke [b], Olaf Köller [a]

[a] Leibniz Institute for Science and Mathematics Education, Germany
[b] Leibniz Institute for Science and Mathematics Education, Centre for International Student Assessment, Germany

A R T I C L E I N F O

Article history:
Received 12 November 2015
Received in revised form
24 August 2016
Accepted 5 October 2016

Keywords:
School quality
Tracking
Ability grouping
Intelligence
Cognitive development

A B S T R A C T

Ability grouping or tracking during secondary schooling is widespread. Previous research shows academic track schools are more successful than non-academic track schools in teaching mathematics, reading and foreign languages. Reasons include a more favorable student composition and higher instructional quality. However, there is less evidence that between track differences are even large enough to differentially affect the students' cognitive development. We used data from a large Hamburg panel study to test this hypothesis ($N = 8628$). By employing several propensity score matching algorithms we formed parallelized samples of academic track and either non-academic track students or comprehensive school students. After four years of tracking, academic track students showed considerably higher intelligence scores than their counterparts at the non-academic tracks and slightly higher scores than students at the comprehensive schools. Our results underline the importance of a cognitively stimulating learning environment in school to support students' cognitive development.

1. Introduction

Some schools more effectively teach reading, mathematics and sciences than others. School effectiveness research mainly agrees with this statement (Reynolds et al., 2014). However, increasing students' general cognitive abilities is usually not an explicit goal of schooling (Adey, Csapó, Demetriou, Hautamäki, & Shayer, 2007). Yet, the question arises whether school quality indicators not only result in different subject specific outcomes but also differentially affect students' general cognitive abilities. This question is relevant against the background of broad evidence regarding the meaning of intelligence for numerous factors of life quality such as educational success, employment status, higher income, better health, higher life expectancy, and enduring partnerships (Der, Batty, & Deary, 2009; Gottfredson, 2003; Wrulich et al., 2013). Therefore, and in light of an increasingly complex environment a closely related, albeit not identical construct, that is domain-general problem solving, has received a lot of attention from educational researchers to the point of its inclusion in the PISA 2012 cycle (*Programme for International Student Assessment;* Greiff et al., 2014).

Most recently, to address the question of school quality effects on students' intelligence, Becker, Lüdtke, Trautwein, Köller, and Baumert (2012) took advantage of structural features of the German school system: The explicit between-school tracking during secondary schooling in Germany goes along with significant advantages for the academic tracks in terms of teacher qualification, cognitively demanding instruction and student composition (Klusmann, Kunter, Trautwein, Lüdtke, & Baumert, 2008; Retelsdorf, Butler, Streblow, & Schiefele, 2010; Trautwein, Lüdtke, Marsh, Köller, & Baumert, 2006) and resulted in a clear advance in psychometric intelligence scores for academic track students compared to a matched sample of non-academic track students. Our own study extends the findings of Becker et al. in several directions: First, employing the German adaption of Cattell's Culture Fair Intelligence test (Cattell, 1960; Weiß, 1998), we use a more comprehensive instrument of psychometric intelligence. Second, the sample in our study is eight times larger and considerably more heterogeneous concerning student prior achievement and social background. Third, we not only use students from non-academic tracks but also students from non-tracked comprehensive schools as an additional and more challenging comparison group to the academic track students.

* Corresponding author. Leibniz Institute for Science and Mathematics Education (IPN), Olshausenstr. 62, D-24118 Kiel, Germany.
E-mail addresses: guill@ipn.uni-kiel.de (K. Guill), oluedtke@ipn.uni-kiel.de (O. Lüdtke), koeller@ipn.uni-kiel.de (O. Köller).

http://dx.doi.org/10.1016/j.learninstruc.2016.10.001

1.1. Tracking and student achievement

Many school systems integrate some sort of grouping of students at least during secondary schooling based on the assumption that teaching is easier and more effective in homogenous groups (LeTendre, Hofer, & Shimizu, 2003). Grouping can take place within class, on a course-level (*setting* or *streaming*) or on a school level (*tracking*). The placement of students often depends on their achievement (*ability grouping,* Trautwein et al., 2006). Differences between groups or tracks are expected for two main reasons, *compositional effects* and *institutional effects* (Maaz, Trautwein, Lüdtke, & Baumert, 2008). Compositional effects refer to the more favorable student composition at academic track schools. On average, students show higher achievement and higher cognitive abilities along with a more favorable social background. This allows for interactions between students which are more cognitively activating. Institutional effects refer to the fact that tracks differ in their pedagogical response to the different groups in terms of curricular foci, teacher qualification and instructional quality (Ireson & Hallam, 2001). Concerning the curriculum, in Germany, for example, academic track students are required to learn a second foreign language (Kultusministerkonferenz, 2006). In their language lessons they focus more on literature while in the non-academic track the focus is more on basic linguistic skills (Klieme et al., 2008). Academic teachers have greater content knowledge and greater pedagogical content knowledge. This results in cognitively more activating instruction, for example by encouraging students to discuss and validate different solution paths of a specific task instead of training one correct solution (Baumert et al., 2010; Klusmann et al., 2008; Retelsdorf et al., 2010).

Research on the effects of tracking has shown, that academic track students indeed reach a higher level of achievement than students on other, more vocationally-oriented tracks, even when controlling for intake differences between tracks. This effect is most pronounced for mathematics achievement (Becker, Lüdtke, Trautwein, & Baumert, 2006; Guill & Gröhlich, 2013; Opdenakker & Van Damme, 2006), but can also be found for French (Neumann et al., 2007) as a foreign language. Findings for reading achievement are less consistent and if track differences exist, effect sizes are lower (Retelsdorf, Becker, Köller, & Möller, 2012).

1.2. Tracking and intelligence

Increasing students' general cognitive abilities is neither just another subject in school nor an explicit aim of systematic instruction (for a criticism, see Adey et al., 2007; similar for domain-general problem solving Greiff et al., 2014).

When speaking of students' cognitive abilities or their intelligence we think of their „ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought" (Neisser et al., 1996, p. 77). In some models, it is differentiated in a crystallized component, that is acquired abilities, and a fluid component, the capacity to analyze and solve novel problems independent of cultural experiences and acquired abilities (Cattell, 1963; Horn, 1994). There is also evidence that fluid intelligence coincides with the *g* factor, the common factor resulting from factor analyses of broad ranges of intellectual tasks (Jensen, 2002). According to Cattell's Investment theory this is because at the beginning of an individual's development her or his fluid intelligence is invested in all kinds of complex learning tasks resulting in high correlations between acquired, crystallized abilities (Valentin Kvist & Gustafsson, 2008). From a developmental perspective following a Piagetian tradition, fluid intelligence is also modelled as developing through four reconceptualization cycles. School-age children are either in the cycle of rule-based reasoning (6–11 years) or principle-based reasoning (11–18 years). Each cycle consists of two phases, the latter implying the full mastery of the thinking possibilities of the new cycle. Growth through these cycles is characterized by change in the nature of representations and their inferential interlinking (Christoforides, Spanoudis, & Demetriou, 2016).

There is no doubt about substantial influence of the genetic disposition on an individual's intelligence (Plomin, 2003). However, we know from various fields that a cognitively stimulating environment also has positive effects on individual cognitive abilities. This could e.g. be shown for challenging work environments (Schooler, Mulatu, & Oates, 1999), memory training programs (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008), music practice (Schellenberg, 2006) and direct or content-based training programs (Adey et al., 2007). Last but not least there is strong evidence regarding the impact of quantity of schooling on students' intelligence. As Ceci (1991) documented especially when using natural experiments, every year of schooling brings with it substantial IQ score gains of 2–6 points. However, it remains unclear whether school quality differences are substantial enough to affect the students' general cognitive abilities differentially. In tracked school systems the compositional and institutional effects described above, consistently work across all academic subjects. Concerning compositional effects following Vygotski's concept of *mediated learning experiences* the interaction with peers being slightly ahead in terms of cognitive functioning should stimulate learning processes (Adey et al., 2007) and these peers are more likely to be found at the academic tracks. Concerning institutional effects across all subjects there is more stimulation of advanced reflection at the academic tracks e.g. when learning to identify the common structure of a drama in different plays or when learning the requirements of valid mathematical proofs. It is known from content specific training programs that they transfer to the students' fluid intelligence and can either improve the students' efficiency of reasoning on a given developmental cycle (Papageorgiou, Christou, Spanoudis, & Demetriou, 2016) or accelerate the transition to the following cycle (Christoforides et al., 2016). In sum, because of the more activating environment in academic tracks one might expect a positive influence of academic tracks on their students' intelligence.

Until now, the effect of tracking on students' intelligence development has been investigated several times. Findings from Swedish (e.g. Balke-Aurell, 1982; Härnqvist, 1968), Israeli (Shavit & Featherman, 1988) and US American studies (Rosenbaum, 1975) during the last decades show consistently higher intelligence scores for students on academically oriented tracks compared to students on vocationally oriented tracks. Cliffordson and Gustafsson (2008) could demonstrate advantages for different academic profiles (social sciences vs. technical) on the respective components of an intelligence test. All of these studies found systematic differences in the social and cognitive composition of the students at the onset of tracking. They usually controlled for at least some of these intake differences using standard least-square regression analyses. However, they all have been criticized either for controlling only a few variables and potentially failing to control all the selection bias or for relying on regression analyses without fulfilling its preconditions, e.g. by extrapolating results for subjects without comparable individuals in the control group (Becker et al., 2012; Brody, 1992).

In their study, Becker et al. (2012) made considerable efforts to overcome these disadvantages. In Germany, after primary school students continue on different formal educational tracks, these being either vocational (further: non-academic track) or academic. In the Becker et al. study tracking started after six years of primary