

PATHFINDER



The Pathways to Adulthood Newsletter



In Brief

Changes in education test scores in England over time: are estimates from PISA robust?
Article by John Jerrim, Institute of Education, University of London
Page 2

Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept
Article by Anna K. Chmielewski, Michigan State University
Page 4

Author Profile: John Jerrim
Page 7

Author Profile: Anna K. Chmielewski
Page 7

Special issue on school success: Perspectives from Europe and beyond
Page 8

Engagement in STEM learning and teaching with mobile video inquiries and communities
Page 8

New dataset
Page 8

Upcoming event
Page 8

Capacity building and advancing the understanding of productive youth development in an international context are the objectives of the PATHWAYS Post-Doctoral Fellowship Programme funded by the Jacobs Foundation. In our biannual issue of the PATHWAYS newsletter, the PATHFINDER, we report on research conducted by the PATHWAYS fellows and the PATHWAYS team. Issue number 6 focuses on research using evidence from the Programme for International Student Assessment (PISA).

In this issue of the PATHFINDER two studies are featured that used evidence from PISA to examine trends in academic performance over time and the linkage between school tracking and student's academic self concept. PISA is an international study that was launched by the OECD in 1997. It aims to evaluate education systems worldwide by assessing 15-year-olds' competencies in the key subjects: reading, mathematics and science. Every three years, a randomly selected group of fifteen-year-olds take tests in the key subjects: reading, mathematics and science, with focus given to one subject in each year of assessment. PISA is unique because it develops tests which are not directly linked to the school curriculum. The students and their school principals also fill in background questionnaires to provide information on the students' family background and the way their schools are run.

The two papers represented in this issue of the PATHFINDER used assessments of math ability and associated self concepts. John Jerrim reports on his study examining changes in math test scores in England over time which has had a large impact on current public policy debates in the UK. Evidence from the PISA data suggests a dramatic decline in math performance between 2000 and 2009 in England. Examining the robustness of the finding John compared findings from PISA and TIMSS (Trends in International

Mathematics and Science Study) and found that the findings from the two studies contradict each other, with TIMSS showing an improvement in math performance in England over the same time period. He then examined the reasons for this contradiction which include alterations to the target population, survey procedures and problems with non-response. For example, while TIMSS focuses on student's ability to meet an internationally agreed curriculum, PISA examines functional ability, i.e. how well students use skills in 'real life' situation. John's conclusion is that one has to be careful when examining change in test performance over time, especially in England.

Using data collected for different countries at only one time point Anna K. Chmielewski reports on a study she has conducted to examine the role of tracking in shaping student's academic self concept. She differentiates between different types of tracking (between-school; within school; and course-by-course tracking) and hypothesizes different effects on academic self concept based on tracking type. The findings suggest that course-by-course tracking has negative associations with math self concepts among low-track students, while for the other two tracking systems students on the lower track reported higher self concepts. These findings were obtained after controlling for individual and track average mathematics attainment. The findings point towards the role of the reference group against which the students compare themselves, and instructional practices that can differ between tracks. It appears that track effects on academic self-concept are better predicted by the reality of students' daily lives than the meaning attached to tracks by society.

Please visit our website to find out more about our work and the team:
www.pathwaysto adulthood.org

Changes in educational test scores in England over time: are estimates from PISA robust?

John Jerrim
Institute of Education
University of London

One of the major developments in educational research over recent years has been the widespread implementation of the international studies of pupil achievement, PISA (Programme for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading Literacy Study). Each has the explicit aim of producing cross-nationally comparable information on children's abilities at a particular age in at least one of three areas (reading, maths and science) and is widely cited by academics and policymakers. Another goal of these studies is to monitor how countries are performing relative to one another (in terms of educational achievement of school pupils) and to map trends in performance over time. An example is a recent report published by the Organisation for Economic Co-operation and Development (OECD 2010a), which used information from the four waves of PISA to investigate how test scores have changed across countries since 2000. A similar case study was produced for one country, England (Jerrim, 2013), examining change in performance in the international achievement tests. The findings from that study have had a large impact on government officials' thinking and public policy debate.

Changes in math test scores in England between 2000 and 2009

It is important to explain the motivation for focusing on England and why this has become such an important (and politically sensitive) issue. Children who took part in the first PISA wave (2000) were born in 1984, and would thus have received most of their compulsory schooling during the years when the Conservative party was in power (who held office between 1979 and 1997). The majority of the most recent (PISA 2009) cohort were, on the other hand, born in 1994, and so spent all their time in school under Labour (who governed between 1997 and 2010). Whether rightly or wrongly, many commentators have thus regarded change in England's PISA ranking since 2000 as an evaluation of the Labour government's educational policy success.

When the PISA 2009 results were released in December 2010, it was therefore England's dramatic decline in performance that grabbed the domestic headlines. Figure 1 highlights exactly why this happened. The solid grey line refers to change in real educational expenditure in England since 2000 with dashed lines referring to mean PISA maths test scores (author's calculations) over the same period. A dotted line is also included to illustrate the change in the proportion of children who obtained 5 or more A*-C grades in their General Certificate of Secondary Education (GCSE) exams including maths and English (GCSEs are important national exams children in England sit at age 15/16, with the 5 A*-C threshold often treated as the minimum target that children should attempt to meet). Figure 1 refers to the percentage change since 2000. As one can see, spending on education rose by around 30% over this period in real terms, and was accompanied by a large increase in the proportion of young people achieving 5 A*-C grades. Yet the PISA data contradict this pattern, suggesting that England's secondary school pupils' average maths performance has been in relative decline.

Public Response

This has since become a widely cited "fact" that has been used for both political benefit and to justify the need for policy change. The Daily Telegraph (a leading English newspaper) ran a commentary stating that:

"This is conclusive proof that Labour's claim to have improved Britain's schools during its period in office is utter nonsense. Spending on education increased by £30 billion under the last government, yet between 2000-09 British school children plummeted in the international league tables."

<http://blogs.telegraph.co.uk/news/tobyyoung/100067092/british-schoolchildren-now-ranked-23rd-in-the-world-down-from-12th-in-2000/>

A sentiment echoed by the Secretary of State for Education (Michael Gove MP) in a recent parliamentary debate:

“I am surprised that the right hon. Gentleman has the brass neck to quote the PISA figures when they show that on his watch the standard of education which was offered to young people in this country declined relative to our international competitors. Literacy, down; numeracy, down; science, down: fail, fail, fail.”

<http://blogs.telegraph.co.uk/news/tobyyoung/100067092/british-schoolchildren-now-ranked-23rd-in-the-world-down-from-12th-in-2000/>

It would thus seem that the change in PISA test scores has had a major impact upon policymakers in England, who are extremely concerned that educational achievement has fallen so much over such a short period of time.

Robustness of the evidence

But is it really true that the achievement of secondary school children has declined rapidly in England relative to other countries? As noted by Brown et al (2007) PISA is just one study, which has its merits, but also its defects. Do other international studies of secondary school children (such as TIMSS) paint a similarly depressing picture of England's lack of progress? And if not, can the difference in results be explained?

Examining the robustness of the finding that secondary school children in England are rapidly losing ground relative to those in other countries Jerrim (2013) compared findings from PISA and TIMSS. The analysis demonstrates that results from PISA and TIMSS do indeed conflict, with the latter suggesting that test scores in England have actually improved over roughly the same period. Yet the fact that these two surveys disagree with

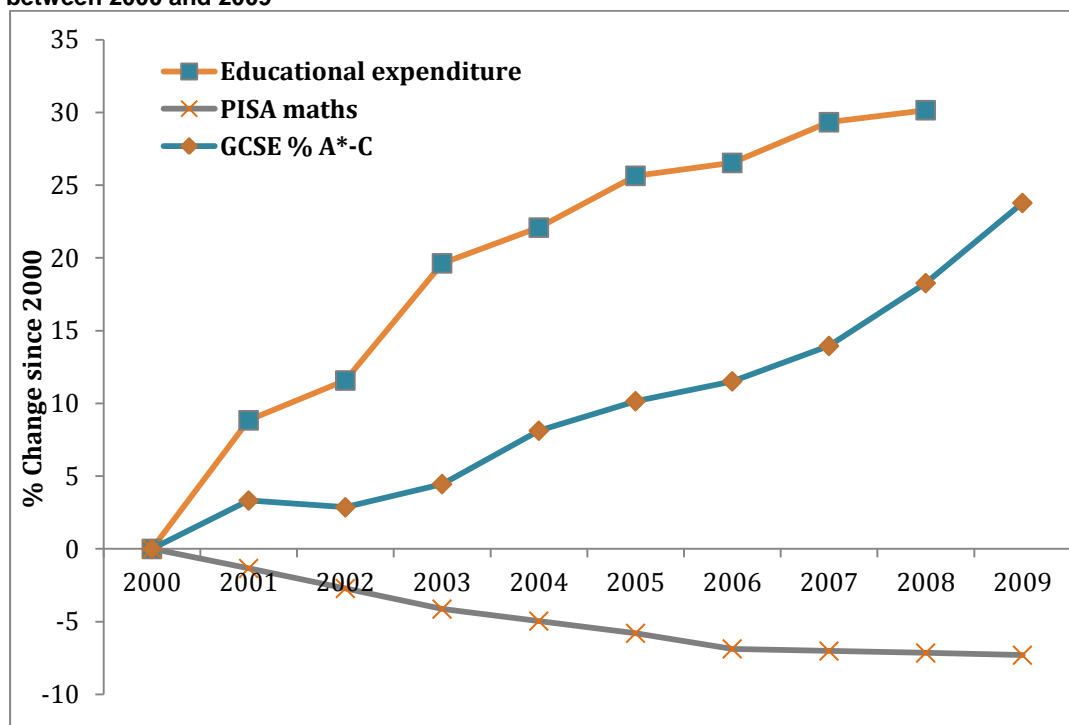
regard to change over time does not seem to be an experience that is shared (at least not to the same extent) by other countries. It is then shown how this may be due to difficulties with the PISA and TIMSS data for England, with a focus on issues such as alterations to the target population, survey procedures and problems with non-response.

Conclusions

Given limitations in the data the following conclusions are drawn:

- Both PISA and TIMSS are problematic for studying change in average test performance in England over time.
- Statements such as those made by the policymakers cited above are based upon flawed interpretations of the underlying data.
- England's movement in the international achievement league tables neither supports nor refutes policymakers' calls for change.

Figure 1: Change in real educational expenditure and mean PISA maths test scores in England between 2000 and 2009



References

- Jerrim, J. (2013). The reliability of trends over time in international education test scores: is the performance of England's secondary school pupils really in relative decline? *Journal of Social Policy* 42(2): 259 – 279.

<http://journals.cambridge.org/action/displayAbstract?fromPage=online&aid=8842935>

You can read more about John's work on his personal website – www.johnjerrim.com

Tracking effects depend on tracking type: An international comparison of students' mathematics self-concept

Anna Katyn Chmielewski
Michigan State University

Hanna Dumont
University of Potsdam

Ulrich Trautwein
University of Tübingen

Most school systems in developed countries around the world practice some form of tracking. *Tracking* - and similar practices known as ability grouping and streaming - groups students into different courses, study programs, or schools according to their achievement. Most research that has looked at the consequences of tracking for academic achievement consistently has found detrimental effects of being placed in a low track. Controversial findings have however emerged for the effects of tracking on students' academic self-concept, that is, students' perceptions of their abilities and competencies. Some research has found that being placed in a low track leads to lower self-concept (Oakes, 1985), while other research has reported positive effects of being placed in a low track (and negative effects of being placed in a high track) on students' academic self-concept (Marsh & Hau, 2003; Marsh, Trautwein, Lüdtke, Baumert, & Köller, 2007).

Why focus on academic self-concept ?

Academic self-concept is not only a desirable outcome in its own right, but is also important because it has significant interlinkages to other outcomes. It has been shown to affect subsequent academic achievement, motivation, effort, and grades (Marsh, Trautwein, Lüdtke, Köller, & Baumert, 2005; Marsh & Yeung, 1997; Valentine, DuBois, & Cooper, 2004), as well as course selection and educational aspirations (Marsh & Yeung, 1998; Mulkey, Catsambis, Steelman, & Crain, 2005). Thus, it is important to know whether educational reforms that do away with tracking, could improve the achievement of low-track students - only to reduce their academic self-concept.

Do different types of tracking create different reference groups?

One possible explanation for the conflicting findings on the relationship between tracking and academic self-concept may be that previous research has taken place in different countries that use a wide range of different tracking policies.

These differences are usually not taken into account when studying the effects of tracking. In our approach (Chmielewski, Dumont & Trautwein, 2013) we distinguish between three types of tracking:

1. *Between-school streaming* - students of different abilities go to completely different schools (e.g., several German states)
2. *Within-school streaming* - students are tracked within schools and assigned to different tracks, programs, or streams for all subjects (e.g., Belgium, Portugal, and Luxembourg)
3. *Course-by-course tracking* - students are tracked within schools only with respect to certain subjects. Thus students can be in a high track for one course and in a low track for another (e.g., the United States, the United Kingdom, and Australia).

We hypothesize that, since students in course-by-course tracking are exposed to students in other tracks on a daily basis, the reference group to which they compare themselves is all students of their age in the entire school, and thus they are constantly reminded of the relative status of their track. In contrast, students in within- and between-school streaming most likely compare themselves only to students from their own track and "forget" about students in other tracks after they have been grouped into their tracks at the beginning of secondary school or the school year. Thus, after controlling for students' own achievement and the average achievement of their track, we expect the remaining positive effect of being in the high track on academic self-concept to be largest in course-by-course tracking and smaller in between- and within-school streaming. Similarly, we expect the negative effect of being in the low track on academic self-concept to be largest in course-by-course tracking and smaller in between- and within-school streaming.

Data source

We used data from the 2003 Programme for International Student Assessment (PISA), and focused on the domain of mathematics. 20 OECD countries had available student track information, and of these countries, six practiced more than one type of tracking, and one (Germany) practiced all three types. Thus, we analyzed a total of 28 different tracking systems. We predicted students' mathematics self-concept based on their track, their mathematics achievement, and track average mathematics achievement, while also controlling for teacher-assigned mathematics course grades, gender, socioeconomic status, and ethnicity.

While some scholars caution against using international assessments like PISA to rank the effectiveness of national education systems (e.g., Jerrim 2013), our analyses use PISA to compare students within countries at a single point in time. Regarding mathematics self-concept specifically, previous research has shown that there are no qualitative differences in the development of mathematics self-concept between different cultural settings (Nagy et al., 2010), and the measure of mathematics self-concept used in the PISA questionnaire has been widely used in prior research in many different national settings.

While the average mathematics self-concept differs greatly across countries in PISA, the *variance* of mathematics self-concept is remarkably consistent across countries, giving further validity to studying variation within countries.

Low-track students' mathematics self-concept suffers in course-by-course tracking and benefits in between- and within-school streaming

Our results were consistent with our hypotheses, but even more extreme. As expected, after controlling for individual and track average mathematics achievement, high-track students in course-by-course tracking had much higher mathematics self-concepts than low-track students. But in between- and within-school streaming, instead of this gap being smaller, it was actually reversed: high-track students had much *lower* mathematics self-concepts than low-track students. Figure 2 (p.6) illustrates the differences between high-track and low-track students' mathematics self-concepts in each of the 28 systems. The positive bars on the left are the systems where high-track students have much higher mathematics self-concepts than low-track students, which include all four course-by-course tracking systems. The short bars in the center are systems where high-track and low track students' mathematics self-concepts do not differ much, and the negative bars on the right are systems where high-track students have much lower mathematics self-

concepts than low-track students. The latter are all within- and between-school streaming systems. This is consistent with the idea that students in course-by-course tracking are more aware of the relative status of their track because their reference group is their entire age group in the school, which includes all tracks.

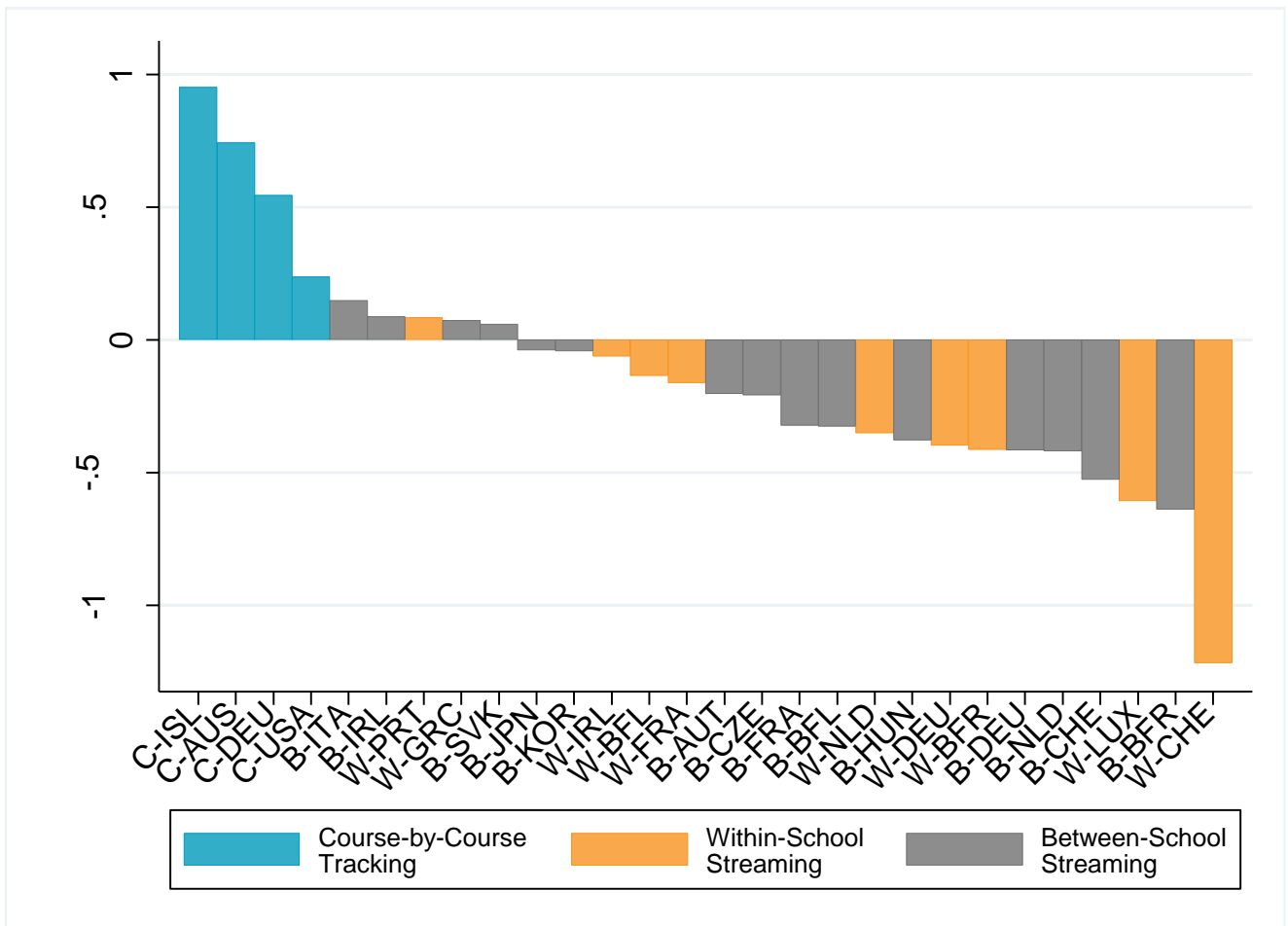
Policy implications

The unexpected negative effects of high track and positive effects of low track on mathematics self-concept in within- and between-school tracking countries suggest that in these systems, instructional practices may differ dramatically between tracks, creating supportive atmospheres in low-track classrooms and competitive atmospheres in high-track classrooms. This raises an important policy implication of the study: If policies to dismantle tracking are to be beneficial for all students, they should be accompanied by instructional practices which bolster the academic self-concept of low-achieving students. Such instructional practices may involve putting less emphasis on competition and ranking of students in the classroom and instead valuing the accomplishments of each individual student. Our study implies that if academic self-concept is not taken into account, lower-achieving students may not be able to fully take advantage of educational opportunities.

References

- Chmielewski, Anna K., Dumont, Hanna and Trautwein, Ulrich. (2013). Tracking Effects Depend on Tracking Type: An International Comparison of Mathematics Self-Concept. *American Educational Research Journal*, 50(5): 925-957.
- Jerrim, J. (2013). The Reliability of Trends over Time in International Education Test Scores: Is the Performance of England's Secondary School Pupils Really in Relative Decline? *Journal of Social Policy*, 42(2): 259-279.
- Marsh, H. W., & Hau, K. T. (2003). Big-Fish-Little-Pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58, 364. doi: 10.1037/0003-066X.58.5.364
- Marsh, H. W., Trautwein, U., Lüdtke, O., Baumert, J., & Köller, O. (2007). The Big-Fish-Little-Pond Effect: Persistent negative effects of selective high schools on self-concept after graduation. *American Educational Research Journal*, 44: 631-669. doi: 10.3102/0002831207306728
- Marsh, H. W., Trautwein, U., Lüdtke, O., Köller, O., & Baumert, J. (2005). Academic self-concept, interest, grades, and standardized test scores: Reciprocal effects of models of causal ordering. *Child Development*, 76: 397-416. doi: 10.1111/j.1467-8624.2005.00853.x
- Marsh, H. W., & Yeung, A. S. (1997). Causal Effects of Academic Self-Concept on Academic Achievement: Structural Equation Models of Longitudinal Data. *Journal of Educational Psychology*, 89: 41-54. doi: 10.1037//0022-0663.89.1.41
- Marsh, H. W., & Yeung, A. S. (1998). Longitudinal structural equation models of academic self-concept and achievement: Gender differences in the development of math and English constructs. *American Educational Research Journal*, 35(4), 705. doi: http://dx.doi.org/10.3102/00028312035004705
- Mulkey, L. M., Catsambis, S., Steelman, L. C., & Crain, R. L. (2005). The long-term effects of ability grouping in mathematics: A national investigation. *Social Psychology of Education*, 8: 137-177. doi: 10.1007/s11218-005-4014-6
- Nagy, G., Watt, H. M. G., Eccles, J. S., Trautwein, L., Lüdtke, O., & Baumert, J. (2010). The development of students' mathematics self-concept in relation to gender: Different countries, different trajectories? *Journal of Research on Adolescence*, 20: 482-506. doi: 10.1111/j.1532-7795.2010.00644.x
- Oakes, J. (1985). *Keeping track: How schools structure inequality*. New Haven, CT: Yale University Press.
- Valentine, J. C., DuBois, D. L., & Cooper, H. (2004). The relation between self-beliefs and academic achievement: A meta-analytic review. *Educational Psychologist*, 39: 111-133. doi: 10.1207/s15326985ep3902_3

Figure 2: Estimated difference in mathematics self-concept between high and low track students



Note. Models control for individual and mean achievement. C=course-by-course tracking, W=within-school streaming, B=between-school streaming. Country codes are as follows: AUS=Australia, AUT=Austria, BFL=Belgium-Flanders, BFR=Belgium-French, CHE=Switzerland, CZE=Czech Republic, DEU=Germany, FRA=France, GRC=Greece, HUN=Hungary, IRL=Ireland, ISL=Iceland, ITA=Italy, JPN=Japan, KOR=Korea, LUX=Luxembourg, NLD=Netherlands, PRT=Portugal, SVK=Slovak Republic, USA=United States.

Author profile:

John Jerrim

John Jerrim is a Lecturer in Economics and Social Statistics at the Institute of Education, University of London

John's research interests include the economics of education, access to higher education, intergenerational mobility, cross-national comparisons and educational inequalities. He has worked extensively with the OECD Programme for International Student Assessment (PISA) data, with this research having been reported widely in the British media.

He received an ESRC Research Scholarship 2006-2010 and was awarded the prize as the 'most promising PhD student in the quantitative social sciences' at the University of Southampton. In October 2011, he was awarded a prestigious ESRC post-doctoral fellowship to continue his research

into the educational and labour market expectations of adolescents and young adults.

Over the last year John has been successful in both publishing his work in academic journals, and disseminating results to a wide audience. He obtained external funding from the British Academy to work on a new project entitled 'University access amongst socio-economically disadvantaged groups – A comparison across Anglophone countries'. This involves close collaboration with other members of the PATHWAYS Programme, including those from the University of Michigan and Michigan State University.



This year the BBC and the Financial Times have both covered John's research and he was also interviewed on BBC Radio 5 and Sky News.

More information on John Jerrim can be found on his website : www.johnjerrim.com

Author profile:

Anna Katyn Chmielewski

Anna K. Chmielewski is a Post-Doctoral Research Fellow based in the College of Education, Michigan State University working with Professor Barbara Schneider

Anna's research aims to measure socioeconomic disparities in educational achievement and attainment over time and across countries, and to explain those disparities by examining the effects of national educational and social policies, income inequality, and segregation.

Anna completed her PhD in Education and MA in Sociology at Stanford University in 2012 working with Sean Reardon and Francisco Ramirez. Her dissertation research compared socioeconomic inequality in achievement internationally in different forms of tracking. Along with Hanna Dumont (University of Potsdam) and PATHWAYS PI Professor Ulrich Trautwein, she recently published a paper in the *American Educational Research Journal* comparing the effects of

different types of tracking on students' mathematics self-concept, which is summarized in this newsletter.

With PATHWAYS alumni John Jerrim and Phil Parker, she recently completed a project examining socioeconomic disparities in access to "high status" post-secondary institutions across Australia, England and the United States.

Currently, Anna is working on a project measuring trends in socioeconomic disparities in academic achievement over the past 47 years in 104 countries.

In spring 2013, Anna was invited to develop and teach a half-day statistical methods module at the Interamerican Development Bank in Washington, DC. The module was



part of a workshop on techniques of collection and analysis of learning assessments for equity studies for an audience of analysts working in Latin American educational ministries. It focused on calculating socioeconomic achievement gaps and trends over time using variables measured categorically and aimed to give the analysts the skills necessary to calculate achievement gaps using national and international databases.

Special issue on school success: Perspectives from Europe and beyond

A special issue of the *European Psychologist* was edited by PATHWAYS members Katariina Salmela-Aro and Ulrich Trautwein (*European Psychologist*, Vol 18(2), 2013, 77-78. doi: 10.1027/1016-9040/a000148).

The special issue focuses on issues related to school success in the European context. It includes five reviews and meta-analyses covering Northern (Finland), Southern (Greece), and central Western Europe (Germany and Austria) as well as from the USA and Australia. The collected papers provide valuable insights and impulses for both the scientific community and for those working in an applied field in the school context. Three papers in this special issue cover intervention programs. The remaining two papers cover timely topics in research on school engagement from a European perspective.

Engagement in STEM learning and teaching with mobile video inquiries and communities

PATHWAYS PIs Katariina Salmela-Aro and Barbara Schneider are a part of a collaborative study between Finland and the United States (supported by the National Science Foundation, Tekes, and the Academy of Finland) that is investigating student engagement in high school sciences classes using the Experience Sampling Method (ESM) and smartphone technology. Recognizing the importance of engagement for student achievement in science, these researchers are testing models to enhance engagement and understand its connection to teacher practices. Pilot data were collected in the spring of 2013 in both countries and analyzed to provide information to teachers on how the students' experiences and emotions tracked during each of their science lessons compared to other classes and out-of-school time. Fall data collection is currently underway in both countries and futures analyses will explore questions such as: What are the preconditions of engagement? How do we know students are engaged? And, How does engagement vary across students and social contexts both in and out of school?

New dataset

A new archived large data set on migration in Germany and Israel is available for PATHWAYS researchers. One of the projects, "Regulation of biographical transitions in second generation immigrants in Germany and Israel", lead by R.K. Silbereisen, Eva Schmitt-Rodermund, Peter Titrmann (now Univ Zurich), has been conducted in Jena, and the overall-direction was also established from there.

For more information, see:
<http://www.gesis.org/en/services/data-analysis/survey-data/migration-and-societal-integration/>

Upcoming event

International Conference on Motivation ICM 2014: Understanding and Facilitating the Passion to Learn

12 - 14 June 2014

Faculty of Behavioural Sciences,
University of Helsinki
Helsinki, Finland.

Keynote speakers include Katariina Salmela-Aro, Ulrich Trautwein and Helen Watt.

For further information see:
www.icm2014.fi



- The major objective of the programme is to promote the next generation of researchers through funding, mentoring and collaboration
- The mission of this Collaborative Post-Doctoral Fellowship Programme is to stimulate innovative, interdisciplinary, and comparative research of productive youth development.
- Our programme brings together experts from the UK, Germany, Finland, Sweden and the US
- We engage with different stakeholders in how best to equip young people or mastering the challenges of growing up in a changing social context.

Principal Investigators and participating institutions

- Lars Bergman – Stockholm University
- Jacquelynne S. Eccles – University of Michigan
- Katariina Salmela-Aro – University of Helsinki
- Barbara Schneider – Michigan State University
- Ingrid Schoon – Institute of Education, University of London
- Rainer K. Silbereisen – University of Jena
- Ulrich Trautwein -University of Tübingen

www.pathwaysto adulthood.org