

Title: Why is educational achievement so highly heritable?

Eva Krapohl*, Kaili Rimfeld*, Nicholas G. Shakeshaft , Maciej Trzaskowski , Andrew McMillan , Jean-Baptiste Pingault, Kathryn Asbury, Nicole Harlaar, Yulia Kovas, Philip S. Dale, and Robert Plomin

*Joint first authors

Abstract

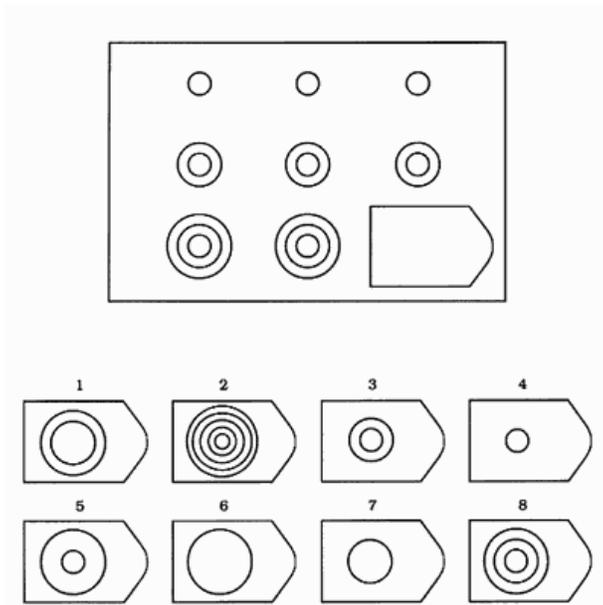
Because educational achievement at the end of compulsory schooling represents a major tipping point in life, understanding its causes and correlates is important for individual children, their families and society. Here we identify the general ingredients of educational achievement using a multivariate design that goes beyond intelligence to consider a wide range of other predictors such as self-efficacy, personality, and behavior problems, in order to assess their independent and joint contributions to educational achievement. We use a genetically sensitive design to address the question of why educational achievement is so highly heritable. We focus on the results of a UK-wide examination, the General Certificate of Secondary Education (GCSE), which is administered at the end of compulsory education at age 16. GCSE scores were obtained for 13,306 16-year-old twins, whom we also assessed contemporaneously on 83 scales that were condensed to nine broad psychological domains*, including intelligence, self-efficacy, personality, wellbeing, and behavior problems. GCSE core subjects (mean of English, mathematics, science) is more heritable (62%) than the nine predictor domains (35% to 58%). Each of the nine domains correlates significantly with GCSE results, and these correlations are largely mediated genetically. The main finding is that, although intelligence accounts for more of the heritability of GCSE than any other single domain, the other domains collectively account for as much GCSE heritability as intelligence. Together the nine domains account for 75% of the heritability of GCSE. We conclude that the high heritability of educational achievement reflects many genetically influenced traits, not just intelligence.

*A description of the 83 scales used to create the nine broad psychological domains is provided below

1. General Cognitive Ability ('g') (2 scales)

- Ravens progressive matrixes (Raven, Court, & Raven, 1996). This is a nonverbal ability task. This test consists of a series of incomplete patterns ("matrixes"); in each case, the twin is asked to identify the missing part of the pattern. The twin does this by clicking on one of 8 possible missing parts.

For example:



- Mill Hill Vocabulary test (Raven, Raven & Court, 2010). The test consists of a series of multiple choice questions. In each question, a single word is presented at the top of the screen. Below it, 6 other words, labelled A, B, C, D, E and F are presented as the response options. Twins respond by clicking on the response they think is closest in meaning to the word at the top of the screen. There is only one correct response for each question.

For example:

Word	Response A	Response B	Response C	Response D	Response E	Response F	Correct Response
Fascinated	ill-treated	poisoned	frightened	modelled	charmed	copied	E

2. Home Environment (10 scales)

- CHAOS (Confusion, Hubbub and Order Scale) at home. (Matheny, Wachs, Ludwig and Phillips (1995). This is a home environment scale, measuring the household disorganization. Data was collected from twins, who had to rate the statements on a 3-point scale from not true to very true. For example: It’s a real „zoo“ in our home; you can’t hear yourself think in our home.
- Parental control, this is a measure of home environment and relationships. Adapted from NICHD early childcare and youth development study. Parental control scale measures how decisions about twins daily lives are reached and how much control parents exercise about them. For example twins had to rate the following questions: How do you make most of the decisions about the following topics? How late can you stay up on school night? Which friends you can spend time with etc. on a scale: My parent(s) decide, my parents decide after discussing with me, we decide together, I decide after discussing it with my parents, I decide all by myself.

- Parental monitoring this is a measure of home environment and relationships. Adapted from NICHD early childcare and youth development study. Parental monitoring scale is collected from twins and is measuring how much a parent knows about the lives of the children and what they are exactly doing. For example the twins rated the following questions: How much does a parent or an adult in your home know about: Where you go right after school? Who you spend your time with? Etc.
- Parental discipline, this is adapted from multiple risk factors in the development of externalising behaviour problems (Deater-Deckard, K., Dodge, K. A., Bates, J. E., and Pettit, G. S. (1998). This is a parental discipline scale measure assessing how parent reacts when a child misbehaves. This measure is collected from twins. For example: When I misbehave I am: (1) given a smack; (2) told off or shouted at; (3) Explained what I have done wrong (4) my mom or dad is firm or calm with me. 2 subscales: firm parenting and harsh parenting.
- School engagement: Family support for learning. This is a subscale from School engagement questionnaire adapted from Appleton et al. (2006). This scale measured how much support twins feel they are getting from family. For example the twins rated the following questions: When I have problems at school, my family/carer(s) are willing to help me, When something good happens at school, my family/carer(s) want to know about it etc.
- Inventory of Parent and Peer Attachment. Just the peer attachment was included with 3 subscales, adapted from Armsden, G. C., and Greenberg, M. T. (1987). In all cases twins rated the statements on 5 point scale.

Attachment trust. For example: I can count on my friends when I need to get something off my chest.

Attachment communication scale. For example; My friends encourage me to talk about my difficulties.

Attachment Alienation scale. For example: It seems as if my friends are irritated with me for no reason.

- Paranoid parenting, Items devised by Angelika Ronald and Daniel Freeman. This measure was collected from the parents of the twins and measured how overly worried the parent might be. Parents had to rate statements on 4- point scale. For example: Do you feel that others might try to harm your child? Have you ever worried that your child will be abducted?

3. **School Environment (10 measures)**

- PISA usage of school resources , PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. TEDS measures were selected from the PISA 2000, 2003 and 2006 student questionnaires. Twins were asked how often do they used school resources over the last school year. For example how often did they use library, science laboratories etc.
- Math classroom environment composite measure adapted from PISA and Midgley, Eccles and Feldlaufer (1991). This scale was specifically measuring the mathematics classroom environment. Participants were asked to think about the maths lessons over last year and rate if the statements are true on 4-point scale from never- some lessons- most lessons- every lesson. For example: We get to work with each other in small groups; Some pupils try to be the first ones to answer questions the teacher asks.
- PISA homework behaviour scale, PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. TEDS measures were selected from the PISA 2000, 2003 and 2006 student questionnaires. This scale assessed how children do their homework. Children were asked to indicate which statement applies to them on a 4-point scale. For example: I complete my homework on time; I do my homework while watching television
- PISA homework feedback scale PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. TEDS measures were selected from the PISA 2000, 2003 and 2006 student questionnaires. This scale measured the feedback participants got for the homework handed in. Children were asked to indicate which statement applies to them on a 4-point scale. For example: My teachers make useful comments on my homework

- PISA attitudes to School PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. TEDS measures were selected from the PISA 2000, 2003 and 2006 student questionnaires. This scale measured the attitudes the twins have about school and how it prepares them for life after school. Participants were asked to rate how strongly they agree with statements on 4-point scale. For example: School has done little to prepare me for adult life when I leave school; School has helped give me confidence to make decisions
- School Engagement with 4 subscales: adapted from Appleton et al. (2006). Participants were asked to think back to the last school year and rate how they agree with the statements on 4-point scale from strongly disagree to strongly agree.

Teacher-Student relations. For example: Most teachers in my school are interested in me as a person, not just as a student; I enjoy talking to the teachers at my school.

Control/Relevance of School Work. For example: When I do my schoolwork, I check to see whether I understand what I'm doing; When I do well in school, it's because I work hard.

Peer support for learning. For example: Students at my school are there for me when I need them; Students at my school respect what I have to say.

Future Aspirations/Goals. For example: I plan to continue my education following school; School is important for achieving my future goals.

4. **Personality (10 measures)**

- Personality Big five-factor model. Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., Widiger, T. A. (2006). Twins were told that these questions are about

personality and we asked to report where they fall on the scale from high to low on a 5-point scale.

Neuroticism. For example: rating anxiousness from fearful, apprehensive to relaxed, unconcerned.

Extraversion. For example: rating assertiveness from dominant, forceful to unassuming, quiet, resigned.

Openness. For example: rating ideas from strange, odd, peculiar, creative to pragmatic, rigid

Agreeableness. For example: rating compliance from docile, cooperative to oppositional, combative, aggressive.

Conscientiousness: For example: rating deliberation from cautious, ruminative, reflective to hasty, careless, rash.

- LOT-R Optimism (Scheier, M. F., Carver, C. S., and Bridges, M. W. (1994). Twins were asked: To what extent do the following statements describe you? Participants were asked to rate the statements on a 5-point scale from very much like me to not like me at all. For example: In uncertain times, I usually expect the best; If something can go wrong for me, it will.
- GRIT, Perseverance measure (Duckworth, A.L, and Quinn, P.D. (2009). Twins were asked: To what extent do the following statements describe you? Participants were asked to rate the statements on a 5-point scale from very much like me to not like me at all. For example: I have difficulty maintaining my focus on projects that take more than a few months to complete; Setbacks don't discourage me.

- Ambition (Duckworth, A.L, and Quinn, P.D. (2009). Twins were asked: To what extent do the following statements describe you? Participants were asked to rate the statements on a 5-point scale from very much like me to not like me at all. For example: I aim to be the best in the world at what I do; I am ambitious.
- Curiosity Explore: (Kashdan, T.B., Rose, P., and Fincham, F.D. (2004). Twins were asked: To what extent do you agree with the following statements? The participants had to rate the statements on a 7-point scale from strongly agree to strongly disagree. For example: I would describe myself as someone who actively seeks as much information as I can in a new situation; I frequently find myself looking for new opportunities to grow as a person (e.g., information, people, resources).
- Curiosity Flow: (Kashdan, T.B., Rose, P., and Fincham, F.D. (2004). Twins were asked: To what extent do you agree with the following statements? The participants had to rate the statements on a 7-point scale from strongly agree to strongly disagree. For example: When I am participating in an activity, I tend to get so involved that I lose track of time; When I am actively interested in something, it takes a great deal to interrupt me.

5. **Wellbeing (17 measures).**

- Meaning of life (wellbeing), (Morgan & Farsides (2009). Participants were asked to read the statements and rate how strongly they agree on 7-point scale from strongly agree to strongly disagree. For example: My life interests and excites me; I find it satisfying to think about what I have accomplished in life
- Basic psychological needs (wellbeing) (Deci, E. L., & Ryan, R. M. (2000). Participants were asked to read the statements, and then select the point on the scale that they feel is most appropriate in describing them on a 7-point scale from not at all true to very true. For example: I feel like I am free to decide for myself how to live my life; People I know tell me I am good at what I do.

- Temporal Experience of Pleasure Scale (TEPS) for Hedonia (Gard, D. E., Germans Gard, M., Kring, A. M., and John, O. P. (2006). The participants were asked to rate the statements according to how much they apply to them, based on their thoughts and feelings over the last month. For example: When something exciting is coming up in my life, I really look forward to it; I don't look forward to things like eating out at restaurants.
- Introvertive Anhedonia (Mason, O., Linney, Y., and Claridge, G. (2005). This scale measures the inability to experience pleasure in what other people may find enjoyable. The twins had to agree or disagree with given statements. For example: Are you too independent to get involved with other people?; Are there very few things that you have ever enjoyed doing?
- Life Satisfaction (Huebner (1994). Multidimensional students' life satisfaction scale. Participants were asked to think about the past few months and rate how much they agree with the statements presented on a 6-point scale ranging from strongly agree to strongly disagree. This measure was divided into five subscales:

Life satisfaction Family. For example: My family is better than most; I enjoy being at home with my family.

Life satisfaction Friends. For example: I have a bad time with my friends; My friends will help me if I need it.

Life satisfaction School. For example: I learn a lot at school; There are many things about school I don't like.

Life satisfaction Living Environment. For example: There are lots of fun things to do where I live; I wish there were different people in my neighbourhood/area.

Life satisfaction Self. For example: I am fun to be around; There are lots of things I can do well.

- SHS subjective happiness: (Lyubomirsky, S., and Lepper, H. S. 1999). A measure of subjective happiness. Participants were asked to rate in a 5-point scale statements describing themselves. For example: In general I consider myself a very happy person; Compared to most of my peers, I consider myself more happy etc.
- Gratitude. (McCullough, M. E., Emmons, R. A., and Tsang, J. 2002). The grateful disposition measurement. Twins were asked to rate on a 5-point scale from strongly agree, to strongly disagree how well they agree with the statements regarding the gratitude. For example: I have so much in life to be thankful for; If I had to list everything I felt grateful for, it would be a very long list.
- Hopefulness (Snyder et al. 1997). Participants were given sentences that describe how people think about themselves and how they do things in general. Twins were asked to think about how they behave in most situations and what describes them the best; and rate it on 6-point scale from all of the time to none of the time. There are 2 subscales of the hopefulness measure:

Hopefulness Agency. For example: I think I am doing pretty well; I think the things I have done in the past will help me in the future.

Hopefulness Pathways. For example: I can think of many ways to get the things in life that are most important to me; When I have a problem, I can come up with lots of ways to solve it.

- Life events: reduced Coddington Scale (Coddington, R. D. 1972). Twins were given a list of events that might have happened to them recently. The participants were then asked to put a tick in either the 'No' or 'Yes' box if the event has happened in the past year.

Total number of negative events. For example: The loss of a job by my father or mother; Marital separation of my parents

Total number of positive events : For example: Outstanding personal achievement; Becoming a member of a church

Response to negative life events: Twins had to rate how the life events made them feel from very unpleasant to very pleasant on 5-point scale. For example: the death of a close friend or a relative

Response to positive life events: Twins had to rate how the life events made them feel from very unpleasant to very pleasant on 5-point scale. For example: beginning to date

6. Parent-reported behaviour problems (12 measures)

- SDQ Hyperactivity scale (Goodman, 1997; 2001): a dimensional and developmental measure of child mental health. Five items, tapping three key symptom domains of attention-deficit/hyperactivity disorder (ADHD): inattention (2 items), hyperactivity (2 items), and impulsiveness (1 item). 3-point Likert scale (Not true; Quite true; Very true). Example: My child is restless, overactive, and cannot stay still for long.
- SDQ Conduct scale (Goodman, 1997). This is a dimensional and developmental measure of child mental health. Five items measuring externalising behaviour problems on 3-point Likert scale (Not true; Quite true; Very true). For example: Often fights with other children or bullies them; Often has temper tantrums or hot tempers.

- SDQ Prosocial scale (Goodman, 1997). This is a dimensional and developmental measure of child mental health. Five items measuring prosocial behaviour on 3-point Likert scale (Not true; Quite true; Very true). For example: Considerate of other people's feelings; Helpful if someone is hurt, upset or feeling ill; Often volunteers to help others (parents, teachers, other children).
- Autism Quotient Social scale: (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). This is a parent-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 9 items measuring social-communication deficits on 4-point Likert scale (definitely agree; slightly agree slightly disagree; definitely disagree). For example: Finds social situations easy; Finds it hard to make new friends.
- Autism quotient Attention Switching scale (Baron-Cohen et al., 2001). This is a parent-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 8 items measuring social-communication deficits on 4-point Likert scale (definitely agree; slightly agree slightly disagree; definitely disagree). For example: Does not get upset if his/her daily routine is disturbed; Enjoys doing things spontaneously.
- Autism Quotient Imagination (Baron-Cohen et al., 2001). This is a parent-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 6 items measuring social-communication deficits on 4-point Likert scale (definitely agree; slightly agree slightly disagree; definitely disagree). For example: If trying to imagine something, finds it easy to create a picture in his/her mind; When reading a book, can easily imagine what the characters might look like.
- Autism Quotient Attention to detail (Baron-Cohen et al., 2001). This is a parent-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 5 items measuring attention to detail on 4-point Likert scale (definitely agree; slightly agree slightly disagree; definitely disagree). For example: Usually notices car number plates or similar strings of information; Notices patterns in things all the time.

- ARBQ Anxiety scale (Eley et al., 2003). This is a parent-reported, 19-item questionnaire on anxiety-related behaviours in children. The items (three-point scale: never, sometimes, often) represent five dimensions: General Distress (or negative mood) (For example: Often seems worked up, on edge or tense ; Separation Anxiety (For example: Is often extremely upset or distressed when parent leaves, wound up or stressed ; Fears (For example: Is afraid of animals or insects (like dogs, spiders, or snakes); Obsessive-Compulsive Behaviours (For example: Tends to check things are done exactly right), and Shyness/Inhibition (For example: Tends to be shy and timid)
- MFQ total scale (Messer et al., 1995). This is a parent-reported Short Mood and Feelings Questionnaire (SMFQ), reliably measuring core depression construct in children and adolescents. For example: [Over the past two weeks, my child...] Felt he/she did everything wrong; Felt lonely; Didn't enjoy anything at all.
- Conners Impulsivity scale (CPRS-R; Conners, Sitarenios, Parker, & Epstein, 1998). This is a parent-reported subscale of the Revised Conners' Parent Rating Scale 9 items (4-point Likert scale) measuring impulsive behaviours. For example: Interrupts conversations.
- Conners Inattention scale: parent-reported subscale of the Revised Conners' Parent Rating Scale (CPRS-R; (Conners et al., 1998). 9 items (4-point Likert scale) measuring inattention problems. For example: Does not follow through on instructions and fails to finish work, schoolwork or chores.
- Conners Emotional Lability (CPRS-R; Conners et al., 1998). This is a parent-reported subscale of the Revised Conners' Parent Rating Scale. 3 items (4-point Likert scale) measuring affective volatility. For example: Changes mood quickly and drastically; Cries often and easily.

7. Behaviour problems child reported (8 measures)

- SDQ Peer problems scale (Goodman, 1997). This is a child-reported, subscale of the Strengths and Difficulties Questionnaire, a dimensional and developmental measure of child mental health. 5 items (3-point Likert scale) measuring peer problems. For example: I am usually on my own. I generally play alone or keep to myself”; I have one good friend or more.
- SDQ Hyperactivity scale (Goodman, 1997). This is a child-reported, subscale of the Strengths and Difficulties Questionnaire a dimensional and developmental measure of child mental health. 5 items (3-point Likert scale) measuring hyperactivity. For example: I am restless, I cannot stay still for long; I am constantly fidgeting or squirming.
- SDQ Conduct scale (Goodman, 1997). This is a child-reported, subscale of the Strengths and Difficulties Questionnaire a dimensional and developmental measure of child mental health. 5 items (3-point Likert scale) measuring anti-social and related conduct problems. For example: I get very angry and often lose my temper; I fight a lot. I can make other people do what I want.
- SDQ Prosocial scale (Goodman, 1997). This is a child-reported, subscale of the Strengths and Difficulties Questionnaire a dimensional and developmental measure of child mental health. Five items measuring prosocial behaviour on 3-point Likert scale (‘Not true’; ‘Quite true’; ‘Very true’). For example: I usually share with others (food, games, pens etc.); I am kind to younger children.
- CASI anxiety total (Silverman, Fleisig, Rabian, & Peterson, 1991). This is a child-reported Childhood Anxiety Sensitivity Index (CASI) 18 items (3-point Likert) measuring anxiety sensitivity (i.e., the belief that anxiety symptoms have negative consequences). For example: I don’t want other people to know when I feel afraid; I get scared when I feel nervous.
- Autism Quotient Social Scale (Baron-Cohen et al., 2001). This is a child-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 9 items measuring social-communication deficits on 4-point Likert scale (‘definitely agree’; ‘slightly agree slightly disagree’; ‘definitely disagree’). For example: I prefer to do things with others rather than on my own; I find social situations easy.

- Autism Quotient Attention to detail scale (Baron-Cohen et al., 2001). This is a child-reported subscale of the Autism Quotient (AQ) a measure of autistic-like traits in the general population. 5 items measuring attention to detail on 4-point Likert scale. For example: I notice car number plates or similar strings of information; I notice patterns in things all the time.
- MFQ total scale (Messer et al., 1995). This is a child-reported Short Mood and Feelings Questionnaire (SMFQ), reliably measuring core depressing construct in children and adolescents. For example: [Over the past two weeks ...] I did everything wrong, I felt lonely, I didn't enjoy anything at all.

8. Health related measures (9 measures)

- BMI: Twins' Body Mass Index (BMI), derived from self-reported height and weight.
$$BMI = \frac{\text{weight}(kg)}{(\text{height}(m))^2}$$
- Puberty Status (Petersen, Crockett, Richards, & Boxer, 1988). This is a child-reported Pubertal Development Scale, 9 items measuring pubertal state for both sexes. For example: Would you say that your body hair (underarm & pubic) growth has [Not yet begun/Barely begun/Definitely begun/Completed/Not sure].
- Handedness: child-reported handedness, assessed by two items: "I am...[Left handed/Right handed/Mixed handed]"; "When writing, I use my.. [Left hand/Right hand/Mixed]."
- Insomnia (Bastien, Vallières, & Morin, 2001). This is a child-reported Insomnia Severity Scale measuring perceived sleep difficulties, 7 items (5-point scale). Examples: Difficulty falling asleep; Problems staying asleep.

- Sleep: child-reported Pittsburgh Sleep Quality Index (Buysse, Reynolds III, Monk, Berman, & Kupfer, 1989). 9 items assessing sleep quality. Example: “During the past month, how often have you had trouble sleeping because you: Had bad dreams”.
- Alcohol scale: child-reported questionnaire. Adapted from Avon Longitudinal Study of Parents and children (ALSPAC) www.bristol.ac.uk/alspac/ ‘Life of a 16+ Teenager’ questionnaire. For example: Think back over the last 30 days. How often do you have a drink containing alcohol? ; How many units do you drink on a typical day when you are drinking?
- Tobacco scale: child-reported questionnaire. Adapted from Avon Longitudinal Study of Parents and children (ALSPAC) www.bristol.ac.uk/alspac/ ‘Life of a 16+ Teenager’ questionnaire. For example: Have you ever smoked a cigarette (including roll-ups)? ; How old were you when you first smoked a cigarette?
- Cannabis scale: child-reported questionnaire. Adapted from Avon Longitudinal Study of Parents and children (ALSPAC) www.bristol.ac.uk/alspac/ ‘Life of a 16+ Teenager’ questionnaire. For example : How old were you when you first tried cannabis? ; How many times have you used cannabis, in total?
- Other drugs scale: child-reported questionnaire. Adapted from Avon Longitudinal Study of Parents and children (ALSPAC) www.bristol.ac.uk/alspac/ ‘Life of a 16+ Teenager’ questionnaire. For example: Have you ever tried inhaling or sniffing any of the following within the last twelve months? Aerosols/Gas(butane and lighter refills)/Glue/Solvents(including petrol and paint thinners)/Poppers(also known as amyl nitrates, liquid gold, rush).

9. Self-efficacy and interest in education (5 measures)

- PISA math self-efficacy: child-reported questionnaire, selected from the PISA 2000, 2003 and 2006 student questionnaires PISA (OECD Programme for International Student

Assessment): www.pisa.oecd.org. 8 items. For example: How confident do you feel about having to do the following mathematics tasks? Solving an equation like $3x + 5 = 17$

- PISA math interest: child-reported questionnaire, selected from the PISA 2000, 2003 and 2006 student questionnaires PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. 3 items. For example: I do mathematics because I enjoy it.
- PISA time spent on math: child-reported questionnaire, selected from the PISA 2000, 2003 and 2006 student questionnaires PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. 3 items. For example: Thinking about the past year. How much time did you typically spend per week studying mathematics? Study or homework in mathematics by myself
- Academic self-concept Burden (1998). This is a child-reported 11-item questionnaire shortened from the 20-items 5-point scale from 'Very much like me' to 'Not like me at all'. For example: I like having difficult work to do; I'm clever.
- Total attitude towards key subjects: child-reported questionnaire, selected from the PISA 2000, 2003 and 2006 student questionnaires PISA (OECD Programme for International Student Assessment): www.pisa.oecd.org. 3 items. In general, how important do you think it is for you to do well in the subjects below? [not at all—Very important (on 4-point scale)] Science subjects/Mathematics subjects/English subjects.

References

- Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427–445. doi:10.1016/j.jsp.2006.04.002
- Armsden, G. C., & Greenberg, M. T. (1987). The inventory of parent and peer attachment: Individual differences and their relationship to psychological well-being in adolescence. *Journal of Youth and Adolescence, 16*(5), 427–454. doi:10.1007/BF02202939
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17.
- Bastien, C. H., Vallières, A., & Morin, C. M. (2001). Validation of the Insomnia Severity Index as an outcome measure for insomnia research. *Sleep Medicine, 2*(4), 297–307. doi:10.1016/S1389-9457(00)00065-4
- Burden, R. (1998). Assessing Children's Perceptions of Themselves as Learners and Problem-Solvers The Construction of the Myself-as-Learner Scale (MALS). *School Psychology International, 19*(4), 291–305. doi:10.1177/0143034398194002
- Buysse, D. J., Reynolds III, C. F., Monk, T. H., Berman, S. R., & Kupfer, D. J. (1989). The Pittsburgh sleep quality index: A new instrument for psychiatric practice and research. *Psychiatry Research, 28*(2), 193–213. doi:10.1016/0165-1781(89)90047-4
- Coddington, R. D. (1972). The significance of life events as etiologic factors in the diseases of children—II a study of a normal population. *Journal of Psychosomatic Research, 16*(3), 205–213.

- Conners, C. K., Sitarenios, G., Parker, J. D. A., & Epstein, J. N. (1998). The Revised Conners' Parent Rating Scale (CPRS-R): Factor Structure, Reliability, and Criterion Validity. *Journal of Abnormal Child Psychology*, *26*(4), 257–268. doi:10.1023/A:1022602400621
- Corsi, P. M. (1973). *Human memory and the medial temporal region of the brain*. ProQuest Information & Learning.
- Deater–Deckard, K., Dodge, K. A., Bates, J. E., & Pettit, G. S. (1998). Multiple risk factors in the development of externalizing behavior problems: Group and individual differences. *Development and Psychopathology*, *10*(03), 469–493.
- Deci, E. L., & Ryan, R. M. (2000). The “What” and “Why” of Goal Pursuits: Human Needs and the Self-Determination of Behavior. *Psychological Inquiry*, *11*(4), 227–268.
doi:10.1207/S15327965PLI1104_01
- Duckworth, A. L., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit–S). *Journal of Personality Assessment*, *91*(2), 166–174. doi:10.1080/00223890802634290
- Eley, T. C., Bolton, D., O'Connor, T. G., Perrin, S., Smith, P., & Plomin, R. (2003). A twin study of anxiety-related behaviours in pre-school children. *Journal of Child Psychology and Psychiatry*, *44*(7), 945–960. doi:10.1111/1469-7610.00179
- Farrell Pagulayan, K., Busch, R. M., Medina, K. L., Bartok, J. A., & Krikorian, R. (2006). Developmental Normative Data for the Corsi Block-Tapping Task. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 1043–1052. doi:10.1080/13803390500350977
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A Research Note. *Journal of Child Psychology and Psychiatry*, *38*(5), 581–586. doi:10.1111/j.1469-7610.1997.tb01545.x
- Halberda, J., Mazocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*(7213), 665–668. doi:10.1038/nature07246

- Haworth, C. M. A., Davis, O. S. P., & Plomin, R. (2013). Twins Early Development Study (TEDS): A Genetically Sensitive Investigation of Cognitive and Behavioral Development From Childhood to Young Adulthood. *Twin Research and Human Genetics, 16*(Special Issue 01), 117–125.
doi:10.1017/thg.2012.91
- Huebner, E. S. (1994). Preliminary development and validation of a multidimensional life satisfaction scale for children. *Psychological Assessment, 6*(2), 149.
- Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and Exploration: Facilitating Positive Subjective Experiences and Personal Growth Opportunities. *Journal of Personality Assessment, 82*(3), 291–305. doi:10.1207/s15327752jpa8203_05
- Kovas, Y., Haworth, C. M. A., Dale, P. S., & Plomin, R. (2007). The genetic and environmental origins of learning abilities and disabilities in the early school years. *Monographs of the Society for Research in Child Development, 72*(3), vii, 1–144. doi:10.1111/j.1540-5834.2007.00439.x
- Lyubomirsky, S., & Lepper, H. S. (1999). A Measure of Subjective Happiness: Preliminary Reliability and Construct Validation. *Social Indicators Research, 46*(2), 137–155. doi:10.1023/A:1006824100041
- Mason, O., Linney, Y., & Claridge, G. (2005). Short scales for measuring schizotypy. *Schizophrenia Research, 78*(2–3), 293–296. doi:10.1016/j.schres.2005.06.020
- Matheny, A. P., Wachs, T. D., Ludwig, J. L., & Phillips, K. (1995). Bringing order out of chaos: Psychometric characteristics of the confusion, hubbub, and order scale. *Journal of Applied Developmental Psychology, 16*(3), 429–444. doi:10.1016/0193-3973(95)90028-4
- McCullough, M. E., Emmons, R. A., & Tsang, J.-A. (2002). The grateful disposition: A conceptual and empirical topography. *Journal of Personality and Social Psychology, 82*(1), 112–127.
doi:10.1037/0022-3514.82.1.112

- Messer, S. C., Angold, A., Costello, E. J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents: Factor composition and structure across development. *International Journal of Methods in Psychiatric Research*, *5*, 251–262.
- Midgley, C., Eccles, J. S., & Feldlaufer, H. (1991). Classroom environment and the transition to junior high school. In B. J. Fraser & H. J. Walberg (Eds.), *Educational environments: Evaluation, antecedents and consequences* (pp. 113–139). Elmsford, NY, US: Pergamon Press.
- Morgan, J., & Farsides, T. (2009a). Measuring Meaning in Life. *Journal of Happiness Studies*, *10*(2), 197–214. doi:10.1007/s10902-007-9075-0
- Morgan, J., & Farsides, T. (2009b). Psychometric Evaluation of the Meaningful Life Measure. *Journal of Happiness Studies*, *10*(3), 351–366. doi:10.1007/s10902-008-9093-6
- Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., & Widiger, T. A. (2006). Psychometric Properties of an Abbreviated Instrument of the Five-Factor Model. *Assessment*, *13*(2), 119–137. doi:10.1177/1073191106286748
- Murphy, M. M., & Mazzocco, M. M. M. (2008). Mathematics Learning Disabilities in Girls With Fragile X or Turner Syndrome During Late Elementary School. *Journal of Learning Disabilities*, *41*(1), 29–46. doi:10.1177/0022219407311038
- Opfer, J. E., & Siegler, R. S. (2007). Representational change and children's numerical estimation. *Cognitive Psychology*, *55*(3), 169–195. doi:10.1016/j.cogpsych.2006.09.002
- Petersen, A. C., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. doi:10.1007/BF01537962

- Price, T. S., Freeman, B., Craig, I., Petrill, S. A., Ebersole, L., & Plomin, R. (2000). Infant zygosity can be assigned by parental report questionnaire data. *Twin Research and Human Genetics*, 3(03), 129–133. doi:10.1375/twin.3.3.129
- Raven, J, Raven, J., & Court, J. (1998). *Mill Hill vocabulary scale*. Oxford: OPP.
- Raven, JC, Court, J., & Raven, J. (1996). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. Oxford: Oxford University Press.
- Raven, JC, Court, J., & Raven, J. (1998). *Manual for Raven's Progressive Matrices*. London: HK Lewis.
- Scheier, M. F., Carver, C. S., & Bridges, M. W. (1994). Distinguishing optimism from neuroticism (and trait anxiety, self-mastery, and self-esteem): a reevaluation of the Life Orientation Test. *Journal of Personality and Social Psychology*, 67(6), 1063–1078.
- Silverman, W. K., Fleisig, W., Rabian, B., & Peterson, R. A. (1991). Childhood Anxiety Sensitivity Index. *Journal of Clinical Child Psychology*, 20(2), 162–168. doi:10.1207/s15374424jccp2002_7
- Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, M., ... Stahl, K. J. (1997). The Development and Validation of the Children's Hope Scale. *Journal of Pediatric Psychology*, 22(3), 399–421. doi:10.1093/jpepsy/22.3.399
- Wiig, E. H., Secord, W., & Sabers, D. (1989). *Figurative Language subtest. From the Test of Language Competence - Expanded Edition*. San Antonio, TX: The Psychological Corporation.
- Woodcock, R. W., McGrew, K., & Mather, N. (2001). *Woodcock-Johnson tests of achievement*. Itasca, IL: Riverside Publishing.