Sophie von Stumm, Robert Plomin

Monozygotic twin differences in school performance are stable and systematic

Article (Accepted version)
(Refereed)

Original citation:
von Stumm, Sophie and Plomin, Robert (2018) *Monozygotic twin differences in school performance are stable and systematic.* Developmental Science, 21 (6). e12694. ISSN 1363-755X

DOI: 10.1111/desc.12694

MONOZYGOTIC TWIN DIFFERENCES IN SCHOOL PERFORMANCE ARE STABLE AND SYSTEMATIC

Sophie von Stumm,

Department of Psychological and Behavioral Science, London School of Economics and Political Science (LSE), London, UK.

&

Robert Plomin,

SGDP Research Centre, King's College London (KCL), Institute of Psychiatry, Psychology and Neuroscience (IoPPN), London, UK.

CORRESPONDING AUTHOR: Sophie von Stumm, Department of Psychological and Behavioural Science, London School of Economics and Political Science, Houghton St, London WC2A 2AE, London, UK. Email: s.von-stumm@lse.ac.uk**.**

ACKNOWLEDGEMENTS: We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant to RP from the UK Medical Research Council (MR/M021475/1 and previously G0901245), with additional support from the US National Institutes of Health (AG046938). The research leading to these results has also received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ grant agreement n° 602768 and ERC grant agreement n° 295366. SvS is supported by a Jacobs Foundation Fellowship. RP is supported by a Medical Research Council Professorship award (G19/2).

RESEARCH HIGHLIGHTS

- Non-shared environmental experiences have long been thought to be idiosyncratic and due to chance
- Here we show for the first time that non-shared environmental effects are systematic and stable over time
- We found that differences in school performance within monozygotic twin pairs are age- and trait-general
- These findings will motivate future research for non-shared environmental causes of children's differences in school performance

ABSTRACT

School performance is one of the most stable and heritable psychological characteristics. Notwithstanding, monozygotic twins (MZ), who have identical genotypes, differ in school performance. These MZ differences result from non-shared environments that do not contribute to the similarity within twin pairs. Because to date few non-shared environmental factors have been reliably associated with MZ differences in school performance, they are thought to be idiosyncratic and due to chance, suggesting that the effect of non-shared environments on MZ differences are age- and trait-specific.

In a sample of 2,768 MZ twin pairs, we found first that MZ differences in school performance were moderately stable from age 12 through 16, with differences at the ages 12 and 14 accounting for 20% of the variance in MZ differences at age 16. Second, MZ differences in school performance correlated positively with MZ differences across 16 learning-related variables, including measures of intelligence, personality and school attitudes, with the twin who scored higher on one also scoring higher on the other measures. Finally, MZ differences in the 16 learning-related variables accounted for 22% of the variance in MZ differences in school performance at age 16.

These findings suggest that, unlike for other psychological domains, non-shared environmental factors affect school performance in systematic ways that have long-term and generalist influence. These findings should motivate the search for non-shared environmental factors responsible for these stable and systematic effects on children's differences in school performance.

Words: 235

Keywords: Monozygotic twin; difference scores; non-shared environment; school performance; learning;

INTRODUCTION

Educational achievement at the end of compulsory schooling regulates the access to subsequent occupational and educational opportunities and thus, affects all important life outcomes. This pervasive, lifelong influence necessitates understanding the causes and properties of children's differences in school performance. Previously, two key findings have emerged. The first shows that children's differences in school performance are relatively stable over time, with school grades' inter-correlations ranging from .40 to .80 across subjects and years (e.g. Kovas, Haworth, Dale, & Plomin, 2007; von Stumm, 2017). Thus, children perform relatively consistently in comparison to their peers throughout school, suggesting that the causes of their differences in performance have long-term effects. The second finding comes from well-powered, genetically sensitive studies that report heritability estimates of up to 65% for children's differences in school performance, while non-shared environments account for about 10% of the variance (e.g. Bartels, Rietveld, van Baal, & Boomsma, 2002; Calvin et al., 2012; Shakeshaft et al., 2013). In other words, the majority of children's differences in school performance can be explained by their genetic differences. This high heritability reflects the effects of many genetically influenced traits (Krapohl et al., 2014; Tucker-Drob et al., 2016), for example cognitive ability, personality, and school attitudes that are all associated with school performance, both contemporaneously and across time (e.g. Bartels et al., 2002; Eccles & Wigfield, 2002; Poropat, 2009). This phenomenon is known as 'generalist genes' (Plomin & Kovas, 2005), which implies that the same genes affect many traits, largely accounting for their phenotypic associations and giving rise to the temporal stability of individual differences.

Notwithstanding the high heritability of educational achievement, monozygotic twins (MZ), who are genetically identical, differ in school performance, as well as in all other psychological and behavioral outcomes (e.g. Caspi et al., 2004; Kovas et al., 2007; Plomin & DeFries, 1980). For example based on the IQ bell curve with a mean of 100 and a standard deviation of 15, MZ twins differ on average by 6 IQ points, while DZ twins differ on average by 10 IQ points, and randomly selected pairs of individuals differed on average by 17 IQ points (Plomin and DeFries, 1980). The differences in MZ pairs are necessarily caused by non-shared environmental experiences, because MZ twins, who grow up in the same family, share their genotypes and home environments. It is important to note that non-shared environmental experiences refer to effects that environments have on one twin relative to the other twin within a pair, rather than to the actual occurrence of events (Plomin, 2011; Turkheimer & Waldron, 2000). For example, parental divorce is an event that happens for all children of one family (i.e. shared environment) but the children's perception of and response to the event may differ (i.e. non-shared environment).

For the past three decades, behavioral geneticists have attempted to identify specific non-shared environmental experiences that are causally associated with differences in psychological and behavioral outcomes (Plomin & Daniels, 1987; Plomin, 2011). In this context, researchers typically hypothesize that a specific non-shared environment -- very often MZ's differential parenting experience -- is linked to a specific subsequent outcome (for a review see Turkheimer & Waldron, 2000). For example, Pike and colleagues (1996) showed in 93 MZ pairs aged 10 to 18 years that parent negativity was significantly associated with MZ differences in antisocial behavior, accounting for 1% to 4% of the variance. Studying a larger sample of approximately 500 MZ pairs, Caspi and colleagues (2004) found that MZ differences in maternal expressed emotions at age 5 explained 1% to 3% of the variance in MZ differences in antisocial behavior at age 7. By contrast, Burt and colleagues (2006) reported that MZ differences in child-parent conflict at age 11 were not associated with externalizing behavior problems 3 and 6 years later in a sample of 486 MZ pairs. Alternative non-shared environmental effects have also been studied, for example MZ differences in birth weight have been shown to account for up to 2% of the variance in MZ differences in behavior problems and school performance at age 7 (Asbury, Dunn, & Plomin, 2006a). More recently, Ritchie and colleagues (2014) demonstrated that advantages in reading ability at age 7 that were due to non-shared environmental experiences modestly improved later intelligence at the ages 9 and 10 years.

Overall then, numerous studies have evidenced non-shared environmental experiences affect differences in psychological and behavioral outcomes. Because of their small effect sizes, however, the conclusion that "the causal mechanisms underlying the non-shared environmental variability in outcomes remain unknown" (p. 91, Turkheimer & Waldron, 2000) has been upheld to date. Thus, non-shared environmental experiences are widely believed to be idiosyncratic in the sense that they are due to chance, which renders the identification of systematic non-shared environmental factors impossible (Kovas et al., 2007; Plomin, 2011; Plomin & Daniels, 1987).

For the current research, we focused on the pattern of the effects of non-shared environments rather than on identifying specific non-shared environmental experiences. Specifically, little is known about the temporal stability of MZ differences in school performance and about the breadth of MZ differences across phenotypically related traits. To the extent to which MZ differences in school performance are correlated across time, we can conclude that non-shared environmental influences contribute to the stability of children's differences in school performance. Similarly, the extent to which MZ differences in school performance are associated with MZ differences in other learning-related traits, for example intelligence or personality, indexes the generalist effects of non-shared environmental experiences. Using data from a large British twin study, which is representative of the U.K. population (Kovas et al., 2007), we investigated the extent to which the effects of non-shared environmental experiences on school performance are stable and systematic (age- and trait-general), which challenges the earlier 'gloomy prospect' that non-shared environmental influences are unsystematic and idiosyncratic (Plomin & Daniels, 1987; Plomin, 2011).

METHODS

Sample

TEDS is a multivariate longitudinal study that recruited more than 11,000 twin pairs born in England and Wales in 1994 through 1996. The recruitment process and the sample are described in detail elsewhere (Haworth et al., 2013). The TEDS sample was at its conception representative of the UK population in comparisons with data obtained by the Office of National Statistics, and remains considerably representative, although it has suffered some attrition over 20 years e (Kovas et al., 2007). Families who continued participating in TEDS at age 16 are on average as often of white ethnicity (93%) as are families at first contact (92%) and as are families in Britain in general (93%). Mothers in families who continued participating are slightly more educated as indexed by their higher A-level completion rate (i.e. higher school leaving certificate; 42%) compared to the rate at first contact (36%) and that in Britain in general (32%). Similarly, maternal employment and paternal employment rates differ slightly in TEDS at age 16 compared to the sample at the first contact and in Britain in general (maternal: 47% at 16, versus 43% at first contact, versus 49% in Britain; paternal: 93% at 16, versus 92% at first contact, versus 89% in Britain). However, these biases are small and unlikely to affect the current study's results and their interpretation.

At age 12, 14 and 16 years, 5,990 monozygotic (MZ) twins and 11,248 dizygotic (DZ) twins participated in the respective assessment waves. We excluded twins (a) who had experienced serious medical problems either before or after birth, (b) whose first language was not English, and (c) who missed data on medical problems and first language from our analyses. The analysis sample therefore consisted of 2,768 MZ pairs and 5,143 DZ pairs.

Measures

School performance was assessed at three ages (i.e. 12, 14, and 16 years), while all other measures were completed at the twins' age of 16 years. Sample sizes vary, because (a) at age 16 data were collected in two waves with more participants completing wave 1 (i.e. GCSE data) and fewer wave 2 (i.e. psychological and behavioural outcomes), and (b) at age 14, only a third of the cohort was assessed due to funding limitations.

School performance. At the ages 12 and 14, teachers reported on the twins' performance in (a) English, including the categories 'speaking', 'reading', and 'writing', (b) Maths, including 'use & applying', 'numbers', and 'shapes, spaces and measures', and (c) Science, including the categories 'scientific enquiry', 'life processes', and 'physical processes'. They rated the twins' performance on 9-point rating scale that corresponds corresponding to National Curriculum Levels (https://www.gov.uk/national-curriculum/overview). At age 16, TEDS twins completed the GCSE, which is a UK nationwide examination that until 2016 marked the end of compulsory education. GCSE courses start typically at the age of 14, and the examinations are taken at the age of 16. Here, we focus on three core GCSE subjects, namely English, Math and Science that all students complete. Grades were coded from 11 (the highest grade, A*) to 4 (the lowest pass grade, G); no information about failed results was available. Twins and their parents reported their GCSE grades as part of a TEDS survey. For 7,367 twins, self- and parent-reported GCSE results were verified using data obtained from the National Pupil database (www.gov.uk/government/uploads/system/uploads/attachment_data/file/251184/SFR40_2013_FINALv2.pdf), yielding correlations of .98 for English, .99 for mathematics, and .95 for science.

Intelligence. Twins completed web-based adaptations of Raven's Standard and Advanced Progressive Matrices and the Mill-Hill Vocabulary Scale using their home computers (Raven, Raven, & Court, 1998; Raven, Court, & Raven, 1998). These measures have been described in detail elsewhere (e.g. Krapohl et al., 2014).

Personality. Twins completed five personality measures. *(a) Five Factor Model Rating Form* (Mullins-Sweatt, Jamerson, Samuel, Olson, & Widiger, 2006). The five domains include Neuroticism, Extraversion, Openness to Experience, Agreeableness and Conscientiousness, with six items for each of the five domains. Each item is rated on a 1 to 5 scale where 1 is extremely low, 2 is low, 3 is neither high nor low, 4 is high, and 5 is extremely high. For example, the neuroticism item anxiousness was assessed with the descriptors "fearful, apprehensive versus relaxed, unconcerned, cool" and the openness factor was assessed with the descriptors "strange, odd, peculiar, creative versus pragmatic, rigid." Neuroticism was scored in a positive direction thus, represented emotional stability. *(b) LOT-R Optimism* (Scheier, Carver, & Bridges, 1994). This 6-item measure asked twins to rate themselves on a 5-point scale ranging from 'very much like me' (5) to 'not like me at all' (1). An example item is "In uncertain times, I usually expect the best". *(c) GRIT* (Duckworth & Quinn, 2009). This 8-item measure assessed perseverance and consistency of interest, for example " I have difficulty maintaining my focus on projects that take more than a few months to complete" and "Setbacks don't discourage me". Twins rated their agreement on the same 5-point scale as above. *(d) Ambition* (Duckworth & Quinn, 2009). This 5-item scale assesses ambitious striving (e.g. "I aim to be the best in the world at what I do") with twins rating their agreement on the same 5-point scale as above. *(e) Curiosity* (Kashdan, Rose, & Fincham, 2004). This 7-item scale assessed seeking out and engaging in new information and experiences on a 7-point scale that ranged from 'strongly agree' (7) to 'strongly disagree' (1). Examples include "I would describe myself as someone who actively seeks as much information as I can in a new situation" and "When I am actively interested in something, it takes a great deal to interrupt me".

School environment. Various measures of the school environment were adapted from the PISA 2000, 2003 and 2006 student questionnaires (OECD Programme for International Student Assessment: www.pisa.oecd.org), and completed by the twins. *(a) PISA usage of school resources* (4 items). The twins were asked how often do they used school resources, for example the library or science laboratories, over the last school year on a 5-point scale ranging from 'never' (0) to 'several times per week' (4). *(b) Classroom environment*. This measure consisted of 19 items and assessed the maths classroom environment, for example "We get to work with each other in small groups" and "Some pupils try to be the first ones to answer questions the teacher asks". Twins were asked to think about the maths lessons over last year and rate if the statements are true on 4-point scale ranging from 'never' (0) to 'every lesson' (3). *(c) Homework behavior and feedback*. This 5-item scale assessed how the twins do their homework and understand the feedback that they receive, for example "I complete my homework on time" and "My teachers make useful comments on my homework". The twins were asked to indicate which statement applies to them on a 4-point scale ranging from 'never' (0) to 'always' (3). *(d) Attitudes to school*. This 4-item scale measured the attitudes the twins have about school and how it prepares them for life after school, which they rated on 4-point scale ranging from 'strongly disagree' (1) to 'strongly agree' (4). An example item is "School has done little to prepare me for adult life when I leave school".

School Engagement: This 19-item scale was adapted from Appleton, Christenson, Kim, and Reschly (2006) to assess various aspects of school engagement at twins' age 16, including teacher-student relations (e.g. "I enjoy talking to the teachers at my school"), control over school work (e.g. "When I do well in school, it's because I work hard"), peer support for learning (e.g. "Students at my school are there for me when I need them"), future aspirations and goals (e.g. "School is important for achieving my future goals"), and family support regarding school (e.g. " When I have problems at school, my family/carer(s) are willing to help me"). Twins 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 (1) to strongly agree (4).

Academic self-concept (Burden, 1998). Eleven items were adapted from the original 20-item scale to assess twins' conceptions of themselves as learners and problem-solvers scale on a 5-point scale that ranged from ‘Very much like me’ (5) to ‘Not like me at all’ (1). An example reads, "I like having difficult work to do".

Analysis

We summed up school subject grades to build composites of school performance at the ages 12, 14 and 16. We z-transformed the scores of the two intelligence tests and averaged them to build one composite. We computed relative difference scores for school performance composites and all learning-related variables, whereby we randomly assigned MZ twins within a pair and subtracted the scores of twin 2 from those of twin 1. Relative difference scores are recommended for the analysis of MZ twin differences (Turkheimer & Waldron, 2000). The reliability of relative difference scores increases with higher reliabilities of their constituent variables but decreases if the variables are strongly correlated (Trafimow, 2015). MZ correlations are high for school performance (i.e. $r = .88$ in the current study) but their reliability is even higher (i.e. internal consistency of .98 in the current study), suggesting that the reliability of the MZ difference score is about .80 (Trafimow, 2015). For the other psychological and behavioral measures in TEDS, MZ correlations are lower but the measures' reliabilities are high so that reliability of the MZ difference scores is adequate. Overall, we estimate that the average reliability of our MZ differences score is .60. While this value may appear modest for directly measured scales, it suggests that our MZ differences scores are sufficiently reliable to detect meaningful associations among them, and it is a conservative bias in terms of detecting correlations between MZ differences across age or across variables.

To test the temporal stability of MZ differences in school performance, we first computed correlations across ages. Next, we fitted a linear regression model, using the R package lavaan (Rosseel, 2017), which uses Full Information Maximum Likelihood (FIML) estimation to handle data missing at random, to test how much variance MZ difference scores in school performance at age 16 were predicted by MZ difference scores at ages 12 and 14.

To test if MZ differences in learning-related measures were systematic, we computed correlations for the relative difference scores for all learning-related variables. We then tested, fitting a linear regression model with FIML using lavaan, to what extent MZ differences in learning-related variables explained MZ differences in school performance at age 16.

Results

Table 1 reports sample sizes, means and SDs for all study variables, as well as the corresponding difference scores. Relative difference scores were by and large normally distributed, with skew and kurtosis values ranging from -1.5 to 1.5. One exception was a twin pair whose difference in school performance at age 14 was an extreme outlier with a value of 8, suggesting a data recording error. This pair was excluded from all further analyses.

In absolute terms, MZ twins differed on average .43 (SD = .40, range 0 to 3.17) in school performance at age 16 (N = 2,177), which is equivalent to just under half a grade level. This average grade difference may appear small but is comparable to MZ twins' average difference in IQ (i.e., 6 points), with both approximating a third of a standard deviation. By comparison, DZ twins had a mean difference of .87 (SD = .74, range 0 to 5.6; N = 3,900) in school performance at age 16, twice the magnitude of the MZ differences. For 631 randomly matched pairs from the TEDS sample, the mean difference in school performance was 1.33 (SD = 1.03, range 0 to 5.78), three times the magnitude of the MZ differences. These values match previous estimates (Plomin & DeFries, 1980) and illustrate the greater similarity of MZ twins in school performance compared to DZ twins and unrelated individuals.

Table 1: Numbers of MZ twins, means and SDs for study variables and corresponding relative difference scores

	Twin 1			*Twin 2*			*Relative difference scores*		
	N	**Mean**	**SD**	**N**	**Mean**	**SD**	**N**	**Mean**	**SD**
School performance age 16	2201	8.89	1.21	2200	8.90	1.21	2177	-0.01	0.58
Intelligence	868	-0.06	0.97	859	-0.03	0.99	796	-0.02	0.90
Emotional stability	848	2.59	0.69	848	2.53	0.68	778	-0.06	0.81
Extraversion	847	3.66	0.59	848	3.67	0.64	777	-0.01	0.69
Openness	846	3.58	0.59	847	3.57	0.59	775	0.02	0.66
Agreeableness	844	3.70	0.57	845	3.70	0.58	771	0.00	0.69
Conscientiousness	845	3.76	0.62	845	3.75	0.60	772	0.01	0.68
Ambition	855	3.94	0.68	860	3.92	0.67	794	0.03	0.72
Optimism	868	3.24	0.69	873	3.27	0.73	815	-0.02	0.80
Curious	969	4.75	0.88	965	4.74	0.89	895	0.01	0.96
Grit	867	3.32	0.58	874	3.33	0.59	815	-0.01	0.66
Academic self-concept	856	3.56	0.64	861	3.57	0.61	796	-0.01	0.61
School resources	927	2.96	0.62	909	2.95	0.65	841	0.00	0.71
Classroom	877	1.65	0.34	873	1.66	0.37	761	-0.01	0.41
Homework	926	1.91	0.46	910	1.91	0.46	842	0.02	0.49
School attitude	928	3.26	0.57	916	3.26	0.59	850	-0.01	0.64
School engagement	922	3.01	0.69	910	2.99	0.70	839	0.01	0.91

MZ relative difference scores in school performance at age 12 correlated .21 with scores at age 14 (N = 425 pairs) and .26 with scores at age 16 (N = 842 pairs). MZ relative difference scores in school performance at age 14 correlated .37 with scores at age 16 (N = 896 pairs). Together, MZ difference scores in school performance at age 12 and 14 accounted for 20% of the variance in MZ differences in school performance at age 16 (N = 2,406 pairs), suggesting modest stability over time.

MZ relative difference scores in school performance and the 16 learning-related variables were by and large positively inter-correlated with the twin who scored higher on one measure also scoring higher on others. Furthermore, relative difference scores in traits that are phenotypically more closely related, for example Conscientiousness and grit, were more strongly correlated (r = .33) than those for more distant constructs, for example differences in using school resources and grit (r = .04). Figure 1 illustrates the correlations. The highest association was for MZ differences in ambition and those in academic self-concept (r = .38). A first factor accounted for 15% of the variance in MZ relative difference scores in learning-related variables, suggesting that MZ differences shared common variance across learning-related traits.

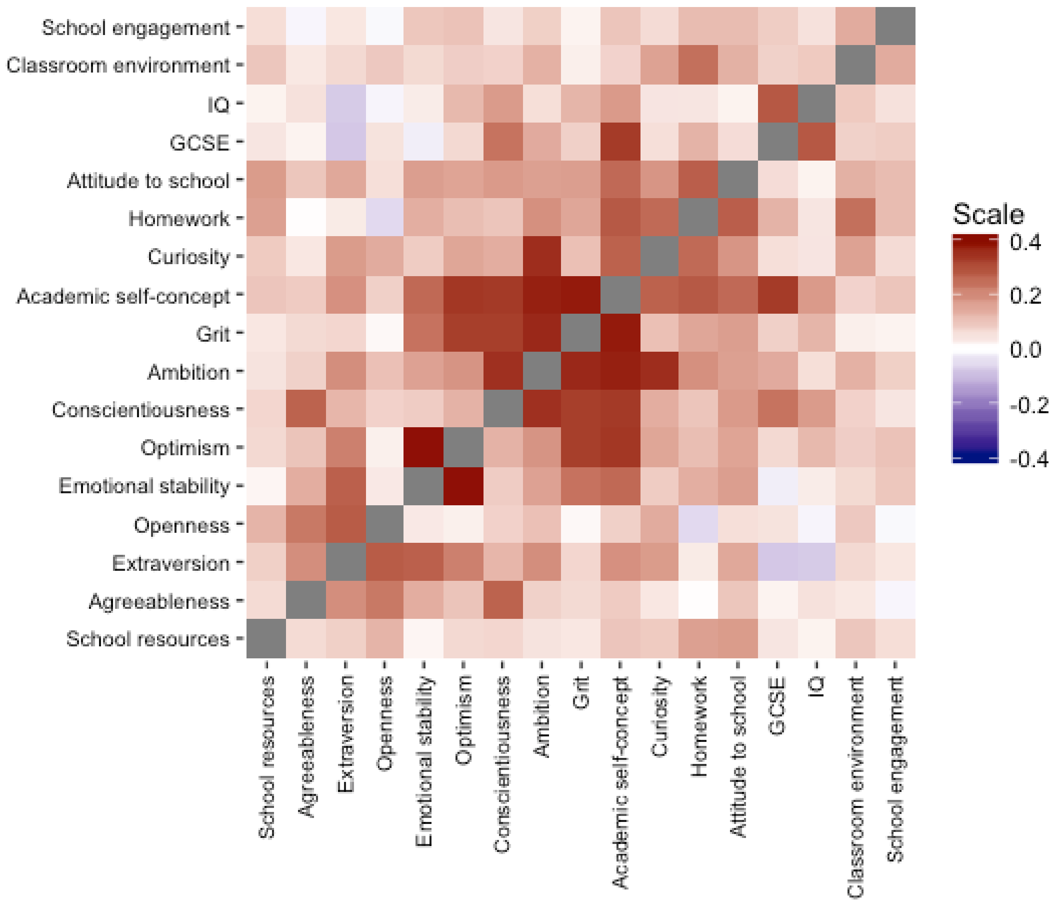

Figure 1: Correlations of MZ relative differences across learning related traits

Note. Variables have been scored in the same direction, so that higher scores indicate better outcomes. The correlation values in numbers are shown in Table S1 in the Supplementary Materials.

Table 2 shows the outcomes of the multiple regression that tested how much variance relative difference scores in learning-related variables explained in MZ's differences in GCSE attainment. The model included 16 predictors that together accounted for 22% of the variance in MZ difference scores in GCSEs (N = 2,280 pairs; R^2 adjusted for attenuation). Predictors that were significantly associated independently of all other predictors with MZ differences in GCSE performance included differences in IQ, conscientiousness, academic self-concept and grit (p < .05 and CI (95%) excluding 0 in all cases). By comparison, the largest beta weight was associated with MZ differences in academic self-concept, which alone accounted for 10.2% of the variance in MZ differences in school performance at age 16 (i.e. almost half of the total variance explained), while the other significant predictors accounted individually for 0.8% (i.e. grit) up to 4.8% (i.e. intelligence) of the variance. This finding suggests that MZ differences in school performance are reflected by MZ differences in cognitive and non-cognitive predictors of academic achievement.

Table 2: Regression coefficients to predict MZ relative differences in GCSEs at age 16 from earlier MZ differences in learning-related variables.

		B	**S.E.**	**β**	***p***	**CI (95%)**	
1	**Intelligence**	**.14**	**.02**	**.22**	**.000**	**.10**	**.19**
2	Neuroticism	-.05	.03	-.07	.103	-.10	.01
3	**Extraversion**	**-.10**	**.03**	**-.12**	**.002**	**-.16**	**-.04**
4	Openness	.05	.03	.06	.111	-.01	.12
5	Agreeableness	-.02	.03	-.03	.500	-.08	.04
6	**Conscientiousness**	**.12**	**.03**	**.14**	**.001**	**.05**	**.18**
7	**Academic self-concept**	**.30**	**.04**	**.32**	**.000**	**.22**	**.38**
8	Ambition	.04	.03	.05	.214	-.02	.10
9	Optimism	-.01	.03	-.02	.703	-.07	.05
10	Curious	-.03	.02	-.06	.125	-.08	.01
11	**Grit**	**-.08**	**.04**	**-.09**	**.024**	**-.15**	**-.01**
12	School resources	.01	.03	.01	.730	-.05	.07
13	Classroom	.01	.06	.01	.825	-.10	.12
14	Homework	.01	.01	.05	.159	-.01	.03
15	School attitude	-.02	.03	-.02	.522	-.09	.04
16	School engagement	.03	.02	.04	.244	-.02	.07

Note. Predictors are MZ's relative difference scores for the respective measure. Significant predictors with *p*-values below .05 and CI (95%) excluding 0, are shown in bold. B = unstandardized estimates; β = standardized estimates; S.E. = Standard Error. Model estimates are based on FIML with N = 2,280 MZ pairs.

DISCUSSION

The effects of non-shared environmental experiences were previously thought to be idiosyncratic in the sense that they involve chance and are age- and trait-specific (Plomin & Daniels, 1987; Plomin, 2011; Turkheimer & Waldron, 2000). The current findings challenge this assumption for the domain of school performance, however, because we show here for the first time that MZ differences are -- at least partly -- temporally stable and systematic across a broad nexus of learning-related traits.

Only 10% of the variance in academic performance can be attributed to non-shared environmental influences in classic twin models (Krapohl et al., 2014). Previous studies of non-shared environmental experiences typically focused on differential parenting and its association with later psychological and behavioral outcomes, revealing several meaningful associations of very small effect size (e.g. Asbury et al., 2006a; Caspi et al., 2004; Pike et al., 1996) but also non-significant relationships (e.g. Burt et al., 2006). For the current study, we investigated MZ differences in school performance in a large British twin sample representative of the U.K. population. We found moderate temporal stability of MZ differences in school performance from age 12 through 16, with earlier MZ differences accounting for 20% of the MZ differences in school performance at age 16. In addition, we showed that the twin within a pair who scored higher on school performance also scored higher on other learning related traits (i.e. MZ differences in learning-related traits were positively inter-correlated), suggesting generalist effects of non-shared environmental influences across many traits.

These results have two important implications. For one, they suggest that the effects of non-shared environmental experiences are not -- as gloomily forecast before -- idiosyncratic in the sense of chance (Plomin & Daniels, 1987; Plomin, 2011), at least not when it comes to school performance. Instead, non-shared environmental experiences contribute to the stability of individual differences in school performance, as we have shown here. For the other, studies that reported associations between MZ differences in two selected cognitive traits, for example in reading ability and intelligence (Ritchie et al., 2014), obscured a more principal phenomenon: Non-shared environments appear to have generalist effects on cognitive outcomes, with MZ differences in learning-related traits being positively inter-correlated across a broad nexus of variables. We also found that the strength of these positive inter-correlations among the MZ differences mirrored that of the trait correlations observed in unrelated individuals. Future research will clarify if this mirror effect only holds for learning-related traits or if the magnitude of association between MZ differences reflects that of unrelated individuals' correlations across the entire phenome.

We showed that MZ differences in 16 traits that are phenotypically associated with school performance, ranging from conscientiousness and grit to academic self-concept and school attitudes, accounted together for 22% of the MZ differences in school performance at age 16. Out of these 16 traits, MZ differences in academic self-concept were associated with the largest effect size in predicting MZ differences in school performance, accounting alone for almost half of the variance explained (i.e. 10%). Previous studies reported substantial effects of non-shared environmental influences in self-concept and related measures, for example 40% for educational efficacy at age 16 (Krapohl et al., 2014), 45% for self-esteem at age 17 (Raevouri et al., 2007), and 34% for self-estimated intelligence (Bratko, Vukasovic, Chamorro-Premuzic, & von Stumm, 2012). Because our study design only allows speculating about the direction of causality, future research will have to test if academic self-concept enhances school performance and thus, constitutes a suitable target for interventions to improve learning success, or if it is the school performance that drives differences in academic self-concept.

Strength and limitations

Our study includes a large twin sample that is reasonably representative of the U.K. (Kovas et al., 2007) and that was assessed on a wide range of relevant traits with established, reliable measures. It also has a key weakness. Our study's measures reflect subjective non-shared experiences or perceptions of experiences, but not actual events that occurred only for one twin but not the other within a pair. Assessing subjective non-shared environments includes twins' potentially different responses to shared environmental events and thus, is more comprehensive and accurate for capturing non-shared experiences. However, this approach does not allow identifying specific non-shared environmental events that may have caused MZ differences in psychological and behavioral outcomes.

What causes MZ differences in subjective non-shared experience? It seems likely that typical MZ differences across the phenome emerge as a result of many trifling differential experiences, rather than of extreme non-shared events (Asbury, Moran & Plomin, 2017). For example, one twin may have had a desk neighbor on his first school day, who shared his pencils for taking notes, while the co-twin had a less benevolent desk neighbor. This slight difference in experience may translate, over time and in conjunction with many other experiences, into more substantial character differences, for example with one twin developing a healthier academic self-concept.

Conclusions

The current study illustrates that the effects of non-shared environmental experiences on school performance are temporally stable and systematic across learning-related traits. Thus, non-shared environmental experiences appear to have generalist effects for cognitive outcomes, rather than being idiosyncratic in the sense of being due to chance. A challenge for future research will be to identify the many small events that underpin the effects of non-shared environmental experiences, akin to the ongoing challenge of identifying the genetic variants that account for the substantial heritabilities of psychological and behavioral outcomes.

ACKNOWLEDGEMENTS

We gratefully acknowledge the ongoing contribution of the participants in the Twins Early Development Study (TEDS) and their families. TEDS is supported by a program grant to RP from the UK Medical Research Council (MR/M021475/1 and previously G0901245), with additional support from the US National Institutes of Health (AG046938). The research leading to these results has also received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ grant agreement n° 602768 and ERC grant agreement n° 295366. SvS is supported by a Jacobs Foundation Fellowship. RP is supported by a Medical Research Council Professorship award (G19/2).

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.

Asbury, K., Dunn, J., & Plomin, R. (2006a). Birthweight‐discordance and differences in early parenting relate to monozygotic twin differences in behaviour problems and academic achievement at age 7. *Developmental Science*, *9*, 107-114.

Asbury, K., Dunn, J., & Plomin, R. (2006b). The use of discordant MZ twins to generate hypotheses regarding non-shared environmental influence on anxiety in middle childhood. *Social Development*, *15*, 564-70.

Asbury, K., Moran, N., & Plomin, R. (2017). Do MZ twins have discordant experiences of friendship? A qualitative hypothesis-generating MZ twin differences study. *PLOS ONE*, *https://doi.org/10.1371/journal.pone.0180521*

Bratko, D., Butkovic, A., Vukasovic, T., Chamorro-Premuzic, T., & von Stumm, S. (2012). Cognitive Ability, Self-assessed Intelligence and Personality: Common Genetic but Independent Environmental Aetiologies. *Intelligence*, *40*, 91-99.

Bartels, M., Rietveld, M. J., van Baal, G. C., & Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics*, *32*, 237–249.

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.

Burt, S. A., McGue, M., Iacono, W. G., Krueger, F. (2006). Differential parent–child relationships and adolescent externalizing symptoms: Cross-lagged analyses within a monozygotic twin differences design. *Developmental Psychology, 42(6),* 1289–1298.

Calvin, C.M., Deary, I.J., Webbink, D., Smith, P., Fernandes, C., Lee, S.H., Luciano, M., & Visscher, P.M. (2012). Multivariate genetic analyses of cognition and school achievement from two population samples of 174,000 and 166,000 school children. *Behavioral Genetics*, *42,* 699–710.

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.

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.

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.

Krapohl E., Rimfeld K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.B., Asbury, K., Harlaar., N., Kovas, Y., Dale, P.S., & Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. Proceedings of the National Academy of Science*, 111*, 15273-8.

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.

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.

Pike, A., Reiss, D., Hetherington, E. M., & Plomin, R. (1996). Using MZ differences in the search for nonshared environmental effects. *Journal of Child Psychology and Psychiatry, 37,* 695-704.

Plomin, R. (2011). Commentary: Why are children in the same family so different? Non-shared environment three decades later. *International Journal of Epidemiology*. *40*, 582–92.

Plomin, R., & Daniels, D. (1987). Why are children in the same family so different from one another? *Behavioral and Brain Sciences*, *10*, 1–16.

Plomin, R., & DeFries, J. C. (1980). Genetics and intelligence: Recent data. *Intelligence, 4,* 15-24.

Plomin, R., & Kovas, Y. (2005). Generalist genes and learning disabilities. *Psychological Bulletin, 131, 592–617.*

Raevuori, A., Dick, D. M., Keski-Rahkonen, A., Pulkkinen, L., Rose, R. J., Rissanen, A., Kaprio, J., Viken, R. J. & Silventoinen, K. (2007). Genetic and environmental factors affecting self-esteem from age 14 to 17: a longitudinal study of Finnish twins. *Psychological Medicine, 37,* 1625-1633.

Raven, J., Raven, J., & Court, J. (1998). *Mill Hill vocabulary scale*. Oxford: OPP.

Raven, J.C., Court, J., & Raven, J. (1998). *Manual for Raven's Progressive Matrices*. London: HK Lewis.

Ritchie, S., Bates, T. C., & Plomin, R. (2014). Does Learning to Read Improve Intelligence? A Longitudinal Multivariate Analysis in Identical Twins From Age 7 to 16. *Child Development*, *86*, 23–36.

Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2), 1-36.

Strong Genetic Influence on a UK Nationwide Test of Educational Achievement at the End of Compulsory Education at Age 16

Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Rimfeld, K., Krapohl, E., Haworth, C. M. A., Dale, S., & Plomin, R. (2013). Strong genetic influence on a UK nationwide test of educational achievement at the end of compulsory education at age 16. *PLOS ONE*, https://doi.org/10.1371/journal.pone.0080341

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.

von Stumm, S. (2017). Socioeconomic status amplifies the achievement gap throughout compulsory education independent of intelligence. *Intelligence*, *60*, 57–62.

Trafimow, D. (2015). A defense against the alleged unreliability of difference scores. *Cogent Mathematics*, *2:* 1064626.

Tucker-Drob, E. M., Briley, D. A., Engelhardt, L. E., Mann, F. D. & Harden, K. P. (2016). Genetically-mediated associations between measures of childhood character and academic achievement. *Journal of Personality and Social Psychology*, *111*, 790-815.

Turkheimer, E., & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin, 126,* 78–108.