Behav Genet (2012) 42:722–731
DOI 10.1007/s10519-012-9540-3

ORIGINAL RESEARCH

Genetic and Environmental Influences on Analogical and Categorical Verbal and Spatial Reasoning in 12-Year Old Twins

Miriam A. Mosing · Jane Mellanby · Nicholas G. Martin · Margaret J. Wright

Received: 13 December 2011 / Accepted: 12 April 2012 / Published online: 3 May 2012

Abstract Research on the genetic influences on different abstract reasoning skills (fluid intelligence) and their interrelation (especially in childhood/adolescence) has been sparse. A novel cognitive test battery, the Verbal and Spatial Reasoning test for Children (VESPARCH 1), consisting of four matched (in terms of test-procedure and design) subtests assessing verbal [analogical (VA) and categorical (VC)] and spatial [analogical (SA) and categorical (SC)] reasoning, was administered to a population based sample of 12-year old twins (169 pairs). Multivariate analysis was conducted to explore the genetic relationship between the four cognitive sub-domains. Heritabilities were 0.62 (VA), 0.49 (VC), 0.52 (SA), and 0.20 (SC). Genetic influences were due to one common factor with no specific genetic influences. This shared genetic factor also explained almost the entire covariance between the domains, as environmental variance was largely specific to each subtest. The finding of no genetic influences specific to each subtest may be due to the uniquely matched design of the VESPARCH 1, reducing confoundment of different test modalities used in conventional tests. For future research or when interpreting previous studies, our findings highlight the importance of taking such potential artefacts (i.e. different test modalities for different sub-domains) into account when exploring the relationship between cognitive sub-domains.

Keywords VESPARCH 1 · Abstract reasoning · Genetic influences · Heritability · Twin study · Fluid intelligence

Edited by Chandra Reynolds.

M. A. Mosing · N. G. Martin · M. J. Wright (✉)
Genetic Epidemiology Unit, Queensland Institute of Medical Research, 300 Herston Road, Herston, Brisbane, QLD 4006, Australia
e-mail: Margie.Wright@qimr.edu.au

J. Mellanby
University of Oxford, Oxford, UK

N. G. Martin · M. J. Wright
School of Psychology, University of Queensland, Brisbane, QLD, Australia

Introduction

Over the last century empirical research on intelligence has largely focused on the identification of one underlying second-order general factor on which all the performance tests of mental abilities load, a factor for general cognitive intelligence, often referred to as *g* (Posthuma et al. 2009; Spearman 1904). Although Horn and Cattell (1966) already suggested an independent fluid cognitive construct a few decades ago, the idea received attention again lately (Blair 2006). Fluid cognitive functions are cognitive processes, not necessarily associated with any specific domain, involving the effortful or active maintenance of information (verbal or visual-spatial) in working memory for purposes of executing goal-directed behaviour and planning (Kane and Engle 2002). Here we are specifically interested in abstract reasoning skills (part of fluid intelligence) in adolescents, tapping different cognitive domains such as spatial, verbal, categorical, and analogical reasoning. Abstract reasoning is an important component of higher order cognitive development in children (Richland et al. 2006). Analogical reasoning (a relational reasoning process involving the mapping of similarities between

concepts that are otherwise dissimilar) and categorical reasoning (the ability to process and recognise categorical relationships between words or objects) have been shown to be closely related (Blair 2006; Green et al. 2006a; Richland and Morrison 2010). Both entail relational knowledge and, especially analogy requires the maintenance, inhibition and manipulation of mental (verbal or spatial) representations and their relationships. However, analogical reasoning involves the mapping of higher-order commonalties (e.g. cause-effect relationships) between concepts, rather than simple recognition of common attributes or simpler relations between two items required for categorical reasoning (i.e. both yellow, or both fruit; Goswami 1989). Despite there being a wealth of studies on different aspects of intelligence and their interrelation (Posthuma et al. 2001; Rijsdijk et al. 2002), there is very little research that has focussed on abstract reasoning skills. In particular the underlying mechanism of different abstract reasoning skills and their interrelation are not yet well understood (Green et al. 2006a; Richland et al. 2006).

Although numerous twin studies have explored the heritability of *g* across the life-span, reporting increasing heritability throughout life-time with 30 % in early childhood to up to 91 % in late adulthood (e.g. Haworth et al. 2010; Hoekstra et al. 2009; Posthuma et al. 2009; Spinath et al. 2003; van Leeuwen et al. 2009; van Soelen et al. 2011), only a few studies have explored the genetic relationship between verbal and spatial abilities in children and adolescents (Alarcon et al. 1999; Casto et al. 1995; Petrill 1997; van Leeuwen et al. 2009), and none have specifically investigated reasoning skills. In addition, the vast majority of studies have used a composite measure or summed scores of several different subtests (i.e. verbal and spatial tests) often tapping slightly different capacities, making the tested abilities rather broad. The difference in format and content between traditional cognitive tests tapping different cognitive domains, e.g. spatial and verbal ability, can have confounding effects. For example, stimuli in the performance subtests of the Wechsler Intelligence Scale for Children (WISC; Wechsler 1992) are presented visually, require manual responses, and are often timed, as opposed to the stimuli of the verbal subtests which are aurally presented and require verbal responses (written or spoken). The same is true for the Wechsler Adult Intelligence Scale (WAIS; Wechsler 1955). This makes it difficult, if not impossible, to differentiate true performance levels on a specific ability from the artefacts of different presentation and response modalities.

Here we employed a novel computerized test of reasoning ability, the first version of the Verbal and Spatial Reasoning Test for Children (VESPARCH 1; Mellanby and Langdon 2010) that allows for direct comparison in the verbal and spatial domains. Within each domain both analogical and categorical reasoning are assessed, with a total of four subtests. The VESPARCH 1, which has been adapted from the Verbal and Spatial Reasoning Test for adults (VESPAR; Langdon and Warrington 1995), was designed to (1) reduce the confounding effects of differences in format and content between traditional tests by using parallel forms of tests, (2) to minimize the effect of limited reading skills or vocabulary caused by educational deprivation on the verbal reasoning test, (3) to identify children underachieving in school relative to their potential by comparing their fluid reasoning test scores to their school attainment, and finally, (4) to identify children with high potential in one or both of the verbal and spatial test domains to potentially adapt their teaching accordingly.

Using the VESPARCH 1, our aim was to estimate for the first time the genetic and environmental influences on reasoning ability in early adolescence, and to disentangle the genetic and environmental covariance between the four subtests. There is continuing debate as to whether specific cognitive processes are related or distinct from one another, with a molar system representing a general process, functioning across a wide variety of cognitive tasks, while a modular system is based on numerous distinct cognitive processing units responsible for the different cognitive tasks (independent and specific; Alarcon et al. 1999; Petrill 1997). Modularity is supported by studies showing that damage to particular brain areas affects some cognitive domains but not others, while studies showing that shortcomings in a specific cognitive domain result in low performance on several different cognitive tasks, favouring the molarity hypothesis. Further, we explored the extent of genetic and environmental covariation on categorical and analogical reasoning. It has been suggested that the same underlying brain network is involved in both categorical and analogical reasoning tasks, with additional activation in analogical reasoning (Green et al. 2006b), and this has been supported by functional magnetic resonance imaging studies, revealing additional activation in the frontopolar cortex in analogical reasoning (Bunge et al. 2005; Green et al. 2006b, 2010).

In the current study we use a population sample of 12-year old twins that are taking part in the Brisbane Adolescent Longitudinal Twin Studies (Wright and Martin 2004). Utilizing the VESPARCH 1, which allows for matched testing procedures of verbal, spatial, analogical, and categorical reasoning, we examined the heritability of reasoning ability in early adolescence, and the extent to which different reasoning skills are influenced by the same set of genes and environmental factors.

Methods

Participants

The sample consisted of 338 monozygotic (MZ) and dizygotic (DZ) twin individuals, all 12-years of age. This included 169 twin pairs (28 MZ female, 29 MZ male, 46 DZ female, 28 DZ male, and 38 DZ opposite-sex twin pairs). The twins were recruited from schools in south-east Queensland and are participants in the Brisbane Longitudinal Twin Studies (BLTS; Wright and Martin 2004). Zygosity status was based on self-report of the twins' parent(s) or guardian as well as the judgment of the Research Nurse. As part of future work zygosity will be confirmed by zygosity typing using DNA, but from previous experience we expect an error rate of less than 1 % (Kasriel and Eaves 1976). Written informed consent was obtained from all participants and a parent or guardian. The study was approved by Queensland Institute of Medical Research Human Research Ethics Committee.

Measures

VESPARCH 1 (Mellanby and Langdon 2010). The VESPARCH 1 is a computerised test of verbal and spatial reasoning skills, consisting of four subtests (verbal category, VC; verbal analogy, VA; spatial category, SC; spatial analogy, SA) with 25 test-items each. Each of the subtests is preceded by five practice items that the participant has to complete before starting the assessment to make sure the participant correctly understood the test procedure. The VESPARCH 1 has been designed for 9–13-year olds and aims to minimize the influences of school attainment or social background on the test scores. The four subtests and their instructions are presented on the computer screen as well as aurally through headphones (instructions as well as the verbal subtests) to reduce the reliance on reading skills. Each test-item of the verbal and spatial categorical reasoning subtests consists of four words (verbal subtest) or abstract geometrical shapes (spatial subtest) and require identifying the item which does not belong to the other three items. The verbal and spatial analogical reasoning subtests, involving recognition of relationships between things, consist of a pair of somehow related (e.g. cause-effect, part-whole, etc.) words (verbal subtest) or geometrical shapes (spatial subtests) followed by a single item and four alternative items, of which one can be paired with the single item. Example items for each of the four subtests are shown in Fig. 1. In each of the subtests the participant is required to choose one of four response options by pointing and clicking the mouse. There is no time constraint to the test and this allows the

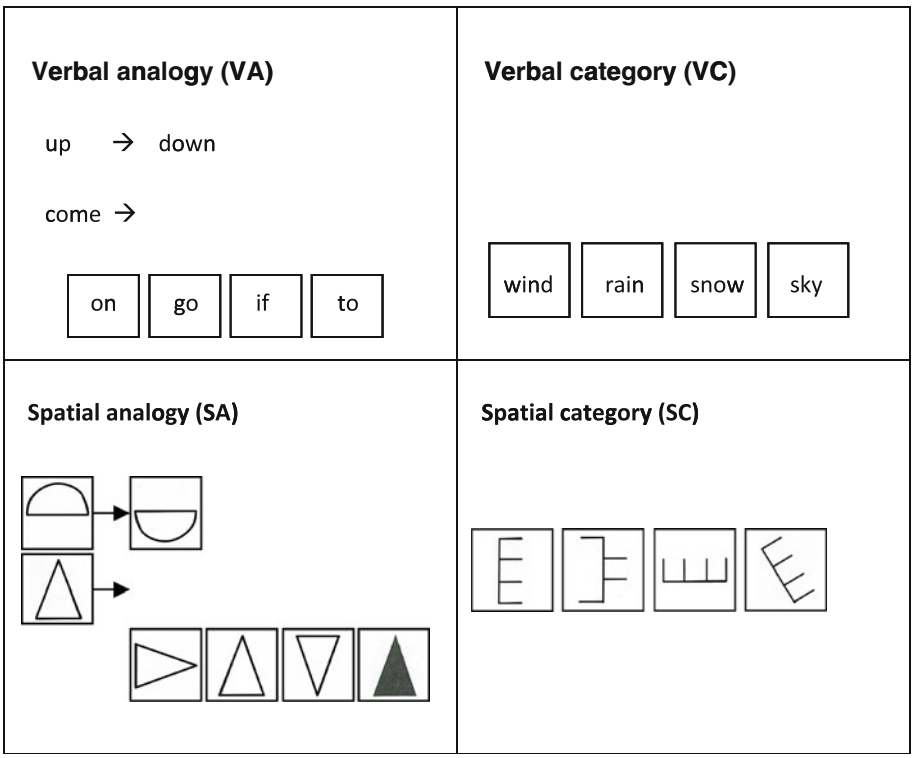

Fig. 1 Example items for each of the four subtests of the VESPARCH 1

Table 1 Means (standard deviations), phenotypic correlations and twin correlations for each zygosity and each of the four variables

	Means (SD), Cronbach's alpha, and phenotypic correlations			
	VA	SA	VC	SC
Sex (N)				
Females (186)	18.92 (3.62)	17.94 (3.25)	15.71 (2.92)	16.57 (2.83)
Males (152)	18.56 (3.78)	17.48 (4.21)	15.51 (3.09)	16.04 (2.84)
Cronbach's alpha	0.76	0.74	0.53	0.53
SA	0.58 (0.50; 0.65)			
VC	0.58 (0.50; 0.65)	0.50 (0.42; 0.59)		
SC	0.45 (0.36; 0.54)	0.49 (0.40; 0.57)	0.32 (0.22; 0.42)	
Zygosity (N_{pairs})	Twin correlations (95 % confidence intervals)			
MZ (57)	0.77 (0.64; 0.86)	0.64 (0.45; 0.78)	0.59 (0.39; 0.74)	0.23 (−0.2; 0.46)
DZ (112)	0.50 (0.35; 0.63)	0.34 (0.17; 0.50)	0.11 (−0.07; 0.29)	0.30 (0.12; 0.46)
MZF (28)	0.71 (0.48; 0.85)	0.50 (0.17; 0.74)	0.64 (0.36; 0.81)	0.36 (0.00; 0.64)
MZM (29)	0.82 (0.65; 0.91)	0.73 (0.51; 0.86)	0.55 (0.25; 0.76)	0.14 (−0.22; 0.47)
DZF (46)	0.58 (0.35; 0.74)	0.24 (−0.05; 0.49)	0.31 (0.02; 0.54)	0.22 (−0.07; 0.48)
DZM (28)	0.13 (−0.24; 0.47)	0.34 (−0.02; 0.63)	0.00 (−0.52; 0.17)	0.27 (−0.10; 0.58)
DZOS (38)	0.58 (0.33; 0.75)	0.48 (0.19; 0.69)	0.19 (−0.12; 0.48)	0.37 (0.06; 0.61)

MZ monozygotic, *DZ* dizygotic, *SA* spatial analogy, *SC* spatial category, *VA* verbal analogy, *VC* verbal category

participant to spend as much time as needed on each item, minimising stress, attention demand and memory load (Edwards et al. 2011; Langdon and Warrington 2000). Each of the test-items requires a response before proceeding to the next test-item. The vocabulary of the verbal subtests is based on words frequently appearing in literature for 9-year olds (younger than the children the test is aimed for) to reduce the impact of aphasia, and the spatial subtests are based on simple, clear and distinct shapes to minimize the demand on perceptual and spatial abilities. The internal consistency of the VESPARCH has been tested in more than 2,000 children revealing a Cronbach's alpha of 0.7 and a 1-year test–retest correlation of 0.6 (Edwards et al. 2011; Mellanby and Langdon 2010). For further details on the VESPAR and VESPARCH 1 refer to Langdon and Warrington (2000) and Mellanby and Langdon (2010). The final test score on each of the four subtests was derived by adding up the items answered correctly (i.e. a maximum score of 25 per subtest). In the present sample, the internal consistency of the whole test (Cronbach's alpha 0.86) as well as the spatial and verbal analogy subtests was relatively high (0.74 and 0.76), while the two category subtests were somewhat lower at 0.53 (Table 1). When combining always two of the subtests to a summary Verbal (e.g. verbal analogical and verbal categorical reasoning), Spatial, Category, or Analogy measure the internal consistency was 0.79, 0.77, 0.64, and 0.84, respectively.

Statistical analysis

For genetic modelling based on the classical twin design, the trait of interest is assumed to be continuous and normal distributed, receiving contributions from independent normally distributed genetic and environmental influences. Genetic influences (A) reflect the additive effects of alleles at multiple genes. Environmental effects can be partitioned into shared/common influences (C) such as their home, their parents, and their social environment (influences which contribute to twin similarity) and non-shared influences (E) is the variance in a trait caused by influences that make the twins different (i.e. unique experiences) as well as variance due to measurement error. Using structural equation modelling, we can determine the relative contribution of genetic (A) and environmental (C and E) influences on the variance of and the covariance between traits. This can be done for variance in a single trait as well as for the covariance between multiple traits and is possible as A, C, and E influences predict different patterns of MZ and DZ twin correlations. The most important assumptions of the classical twin design are that (i) the only difference between MZ and DZ twins is that MZ twins share all their segregating genes, while DZ twins on average only share half their genes and that (ii) the trait relevant environmental influences are similar for MZ and DZ twin pairs (Neale and Cardon 1992).

The twin data were analyzed using maximum-likelihood (ML) modelling employing Mx, a statistical package for the analysis of family data (Neale and Cardon 1992; Neale et al. 2006). Initially, a saturated model was fitted estimating all parameters and then progressively more restricted models are compared to the preceding model. In ML procedures, specific hypotheses regarding the significance of those parameters can be tested statistically by comparing changes in the goodness-of-fit (Δ -2LL) of various models to the observed data (distributed as χ^2) against the change in degrees of freedom (Δdf). If the change in χ^2 ($\Delta\chi^2$) is not significant, generally, the more parsimonious model is regarded the one of choice.

Prior to genetic modelling each of the four variables was tested for equality of means (within twin pairs and across zygosity groups; MZ/DZ), variance and covariance, as well as for equality of correlations of the twin groups. Preliminary testing revealed no significant differences in these estimates. Five individuals scored more than three standard deviations below the mean (two on the SA and three on the SC test) and were therefore winsorized to 3 standard deviations. There were no within-twin pair outliers (i.e. twins being extremely discordant). We also tested for effects of age (in months, as all twins were 12-years old) and sex, as well as the number of months of schooling on the means. All three covariates showed no significant effect.

A Cholesky-decomposition was fitted exploring the genetic relationship between the four subtests of the VESPARCH 1, with the measures included in the following order: VA, SA, VC, and SC. After estimating the relative magnitude of all parameters, we compared the fit of sub-models and other models (a one-factor independent pathway model) to determine the most parsimonious model explaining the relationship between the four cognitive tests and to test the significance of specific parameters.

Results

Table 1 shows the mean scores (ranging between 15.5 and 18.9 out of 25), standard deviations and the phenotypic as well as the different MZ and DZ twin correlations for each of the four subtests. Only 6 of the twins (individuals) scored the maximum score (25) on one of the analogy tests, with none of the twins reaching the maximum score on either of the categorical sub-tests. As mentioned above, there were no significant covariate effects (age, sex, and months of schooling) on the means. The phenotypic correlations between the four subtests were moderate, ranging between 0.32 and 0.58. The correlation between the two category subtests (VC and SC) was 0.32 and significantly lower than the 0.58 correlation between the analogical subtests (VA and SA), and was also lower than the 0.58 correlation between the verbal subtests (VC and SC).

Genetic modelling

The twin correlations (Table 1) for SA, VC and VA were much higher (though not significantly for SA) for MZ than for DZ twins, indicating genetic influences on those variables, while the MZ and DZ twin correlations for SC were similar with very wide confidence intervals. There was no significant difference in twin correlations between the sexes in the five zygosity groups (i.e. between MZ females versus males and between three DZ groups) indicating no significant differences in the genetic architecture of abstract reasoning skills between males and females in the present sample. However, the confidence intervals were very wide indicating low power to detect such potential sex-differences.

An ACE Cholesky decomposition was fitted first given that three (VA, SA, SC) out of the four variables showed some indication of C-influences with the DZ correlations being more than half the MZ correlation. The significance of all C- and A-influences was tested by fitting an AE and a CE model, respectively. Subsequently, a full ACE independent pathway model with one common genetic path as well as a further reduced independent pathway model (without all non-significant paths) were fitted to the data and compared to the fit of the Cholesky decomposition (ACE). Table 2 shows the model fitting results for the different models. The reduced independent pathway model

Table 2 Model fitting results for the four variables with the best fitting model in bold

Multivariate model fitting results	AIC	-2LL	df	Δ -2LL	Δdf	*p* Value
Cholesky decomposition—ACE model	3879.85	6501.85	1311			
Cholesky decomposition—AE model	3867.00	6509.00	1321	7.15	10	0.71
Cholesky decomposition—CE model	3888.20	6522.20	1321	20.35	10	0.03
Independent pathway model—1 common genetic factor (ACE)	3875.94	6501.94	1313	0.09	2	0.96
Reduced independent pathway model—drop all non-sign. paths	**3854.09**	**6514.09**	**1330**	**12.15**	**17**	**0.79**

Model fit (-2LL) of the independent and common pathway model was compared to the Cholesky (ACE model) decomposition

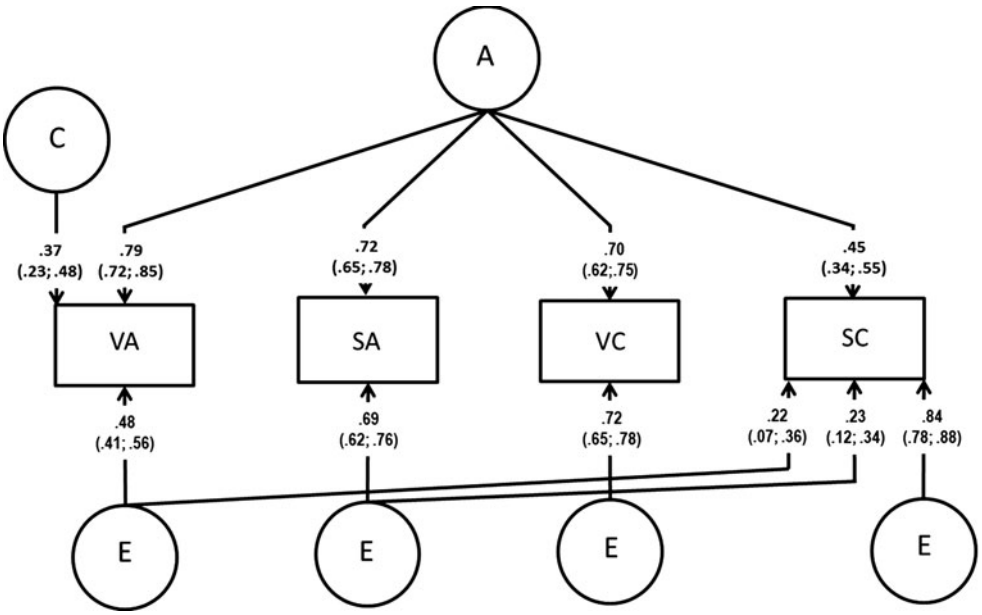

Fig. 2 The reduced independent pathway model with all non-significant pathways dropped. All estimates are standardized so that when all pathway estimates leading to one variable are squared they will add up to one

Table 3 Genetic correlations between the four subtests

	VA	SA	VC
SA	0.93		
VC	0.97	0.99	
SC	0.90	1.00	0.98

was very similar to the full Cholesky decomposition (shown in Appendix) and showed the most parsimonious fit and, therefore, will be presented and discussed here.

The full reduced independent pathway model is shown in Fig. 2. The heritability estimates were 0.62, 0.52, 0.49, and 0.20 for VA, SA, VC, and SC, respectively. A common genetic factor (A) influencing all four reasoning tests could explain all the covariation between VA, SA, and VC. None of the four reasoning tests showed significant specific genetic influences, indicating that all genetic influences on the variation in the tests was due to genes shared with the other variables. Part of the covariation between SC and the two analogical subtests was explained by environmental influences shared between SC with VA and VC, respectively; although, each of these only explained 5 % of the total variance in SC, respectively. In contrast to the non-significant specific genetic factors, the specific non-shared environmental influences (E) were significant and relatively large for each of the four reasoning tests explaining between 23 % (VA) and 71 % (SC) of the total variance. Table 3 shows the genetic correlations, which were very high, ranging between 0.90 (between SC and VA) and 1.00 between the two spatial subtests. Finally, as indicated by the twin correlations, VA was the only variable with a significant common-environmental (C) influence, explaining 14 % of the total variance.

In addition, Table 4 shows the heritability estimates for the sum scores for each of the variables (i.e. the two verbal, the spatial, the analogical, and the categorical subtests added up). The estimates indicate a much higher heritability for the verbal sum-score (75 %), although not significantly different, compared to the spatial one (24 %), while the heritability estimates for the analogy and category sum-scores were very similar.

Table 4 Additive genetic (A), shared (C) and non-shared (E) environmental influences on the four summary measures (i.e. always combining two of the subtests) as well as the Grand Total

	Heritabilities of summary measures (CIs)				
	Sum analogy	Sum category	Sum spatial	Sum verbal	Sum total
A	0.49 (0.20; 0.81)	0.56 (0.11; 0.68)	0.24 (0.00; 0.64)	0.75 (0.44; 0.86)	0.59 (0.27; 0.83)
C	0.29 (0.00; 0.53)	0.00 (0.00; 0.34)	0.30 (0.00; 0.56)	0.05 (0.00; 0.33)	0.17 (0.00; 0.44)
E	0.22 (0.15; 0.33)	0.44 (0.32; 0.62)	0.47 (0.33; 0.64)	0.20 (0.14; 0.30)	0.24 (0.16, 0.35)

Discussion

Here we explored the sources of variance in two verbal and two spatial (analogical and categorical) reasoning tests as assessed by the VESPARCH 1 in young adolescents (12-year old Australian twins). The VESPARCH 1 is a newly developed cognitive test matched in format and content (presentation and response modes) across verbal and spatial (analogical and categorical) reasoning tasks, enabling a direct comparison between those cognitive domains and the subtests. We also aimed to investigate sources of covariation between verbal and spatial reasoning in an effort to contribute to the discussion with respect to modularity versus molarity in cognitive abilities.

On average the twins scored slightly higher on all four subtests compared to a study by Edwards et al. (2011), exploring performance on the two VESPARCH 1 analogy subtests in children (hearing and hard-of-hearing) who were 8 years old. However, given that the present sample was somewhat older (12 years old) a slightly higher score was expected. The four subtests correlated relatively well, with the two categorical measures (VC and SC) correlating least at 0.32 and the two analogical (VA and SA) and verbal (VA and VC) measures correlating the most at 0.58. Interestingly, the lowest and the highest correlations were significantly different, which may be an artefact due to the low scale reliability indicated by the Cronbach's alpha of 0.53 of the two category subtests, i.e. VC and SC correlating the least.

Multivariate genetic (ACE) modelling revealed moderate heritabilities for the four cognitive measures: 62 % (VA), 52 % (SA), 49 % (VC), and 20 % (SC). The confidence intervals of the pathways indicate that the heritability estimate for SC was significantly lower than the estimates for the other sub-tests. As no study to date has used the VESPARCH 1 or a similar test to explore the genetic architecture of verbal, spatial, analogical and categorical reasoning, we have to compare our findings to studies exploring similar concepts utilizing different tests. The heritability estimates of the four VESPARCH 1 subtests are in line with past studies on the heritability of general cognitive ability reporting an increase throughout life, with estimates of about 30 % in early childhood and up to 91 % in late adulthood (Haworth et al. 2010; Hoekstra et al. 2009; Posthuma et al. 2009; van Leeuwen et al. 2009; van Soelen et al. 2011). Our findings are also in line with other studies specifically exploring the heritability of verbal and spatial abilities in late childhood and early adolescence, even though the tests used [mainly utilizing verbal IQ and performance IQ measures of the WISC (Wechsler 1992)], and the reported estimates vary widely, ranging between 26 % and 82 % (Alarcon et al. 1998; Casto et al. 1995; Hoekstra et al. 2007, 2009; van Soelen et al. 2011). Note that performance IQ is based on some, but not only, subtests requiring spatial orientation and reasoning. Finally, although there is no consistent difference in heritability between verbal compared to spatial ability in the child literature, adult studies do report a higher heritability for verbal ability compared to other specific cognitive abilities (Posthuma et al. 2001, 2003; Rijsdijk et al. 2002), which also is in line with our findings of a trend for higher heritability estimates for the two verbal subtests—62 % for analogy and 49 % for category—and the verbal sum-score (75 %) as compared to the equivalent spatial subtests—52 % for the analogy and 20 % for the category—and the spatial sum-score (24 %). However, no study to date has explored the difference in heritability between analogical and categorical reasoning.

The reduced independent pathway model, showed that one common genetic factor explained the entire covariation between three of the four traits (VA, SA, and VC), while some common non-shared environmental influences (i.e. correlated measurement error) explained minor additional covariance (5 % each) between SC and VC and SC and VA, respectively. This suggests that the environmental influences are largely specific with shared genetic influences explaining the vast majority, if not all, of the covariation between the four reasoning tests. The multivariate model, as well as the very high genetic correlations (ranging between 0.90 and 1.00), strongly indicates that although each of the four subtests tap different cognitive domains, they are all influenced by the same set of genes, indicating that within-individual differences in performance on those tests would be due to specific environmental influences. This is a very interesting finding which only is possible due to the four sub-tests of the VESPARCH 1 being matched for format, content and response. However, within individual differences might also be due to gene-environment interactions and correlations which we did not explore.

Based on the modularity hypothesis we would have expected to find two largely genetically independent factors influencing the verbal and spatial tests. However, we find substantial genetic overlap between the different measures of those two domains suggesting molarity rather than modularity, i.e. one shared set of genes explains the variance in performance on different but matched tests tapping different cognitive domains. This is in line with a number of studies exploring genetic and environmental influences on the relationship between verbal and spatial abilities (Alarcon et al. 1999; Casto et al. 1995; Petrill

1997; van Leeuwen et al. 2009) which revealed large genetic overlap between the specific cognitive abilities suggesting one underlying common genetic factor. However, unlike those past studies (Alarcon et al. 1999; Casto et al. 1995; Petrill 1997; van Leeuwen et al. 2009), the present study does not reveal any genetic influence specific to either the verbal or spatial domain. This finding may be explained by the matched test settings of the VESPARCH 1 subtests, i.e. the unique genetic factors in the other studies could potentially be artefacts of different presentation and response modalities of the different subtests used in the WISC.

The finding of shared genetic influences on the analogical and categorical reasoning tests partly supports the findings by Green et al. (2006a) suggesting that both reasoning tests require similar skills (i.e. relational knowledge and categorisation) activating the same underlying brain networks. However, it also has been reported that there is additional activation in the frontopolar cortex in the analogical reasoning task (Bunge et al. 2005; Green et al. 2006b, 2010). In line with this theory, we would have expected significant specific genetic influences on the two analogical reasoning tasks in addition to the genetic overlap between the categorical and analogical subtests. As indicated by the generally large confidence intervals, a lack of power could be an explanation for the non-significant specific influences on the analogical reasoning.

Finally, previous studies also have reported a significant C-effect on the variation in cognitive ability (for a review see Deary et al. 2006) which decreases with increasing age. For example, Haworth et al. (2010) showed in a longitudinal study of 11,000 twin pairs (9, 12, and 17 years) of four subsamples that, while heritability of general cognitive ability increased (41 %, 55 %, and 66 %, respectively), C influences were decreasing from 33 % at 9 years of age to 18 % at 12 years and remained at this level to 17 years. This is in line with our findings of a significant C-effect (14 %) on VA. Also, comparison of the MZ versus DZ correlations in our sample shows some effects of C in SA and SC in addition to VA, as the DZ correlations (0.34 and 0.30 for SA and SC, respectively) were higher than half the MZ correlations for these two variables. However, as indicated by the wide confidence intervals of the twin correlations, the samples size (N = 169 twin pairs) provided insufficient power to detect such modest effects.

A limitation of the present study was the relatively small sample size, reflected in the wide confidence intervals for the twin correlations and the variance component estimates. However, we included a sample as large as possible using all available data from the VESPARCH 1. In the future we will have the opportunity to attain follow-up data in the same twins at age 16, which will enable us to investigate genetic and environmental influences in a longitudinal study design. Also, even though Mellanby and Langdon (2010) reported a relatively high internal consistency for the subscales of the VESPARCH 1, the present data showed a relatively high Cronbach's alpha (0.74 and 0.76) for the two analogy subtests, while the two category subtests were somewhat lower (0.53). A low Cronbach's alpha indicates a high measurement error which would be reflected in inflated E-estimates. Finally, we did not explore gene-environment interactions and correlations and therefore cannot rule out those effects influencing some of our findings.

In summary, heritability estimates for the four abstract reasoning skills were low to moderate ranging between 20 % (SC) and 62 % (VA), with some additional shared-environmental (C) influences on VA. The present study demonstrates that variance in and covariance between performance on four abstract reasoning tests tapping different cognitive domains (i.e. spatial, verbal, categorical and analogical reasoning) are influenced by the same shared set of genes. This finding is strongly in favour of the molarity hypothesis stipulating that different cognitive abilities are driven by the same genetic influences. Unlike past studies, our findings showed no indication of additional genetic influences specific to each of the four reasoning measures. This may be because of the uniquely matched design of the VESPARCH 1, which controls for potential confounders of different presentation and test modality found in the majority of other cognitive tests, and emphasizes the importance of taking such potential artefacts into account.

Acknowledgments This research was supported by the following grant from the Australian National Health and Medical Research Council (NHMRC 552485) and the Australian Research Council (ARC DP1093900). We would like to thank Ann Eldridge and Marlene Grace for the test administration, assessment and data collection and David Smyth for IT support. Finally, this research would not be possible without the willing co-operation of the young twins and their families who participated in this study.

Appendix

See Fig. 3.

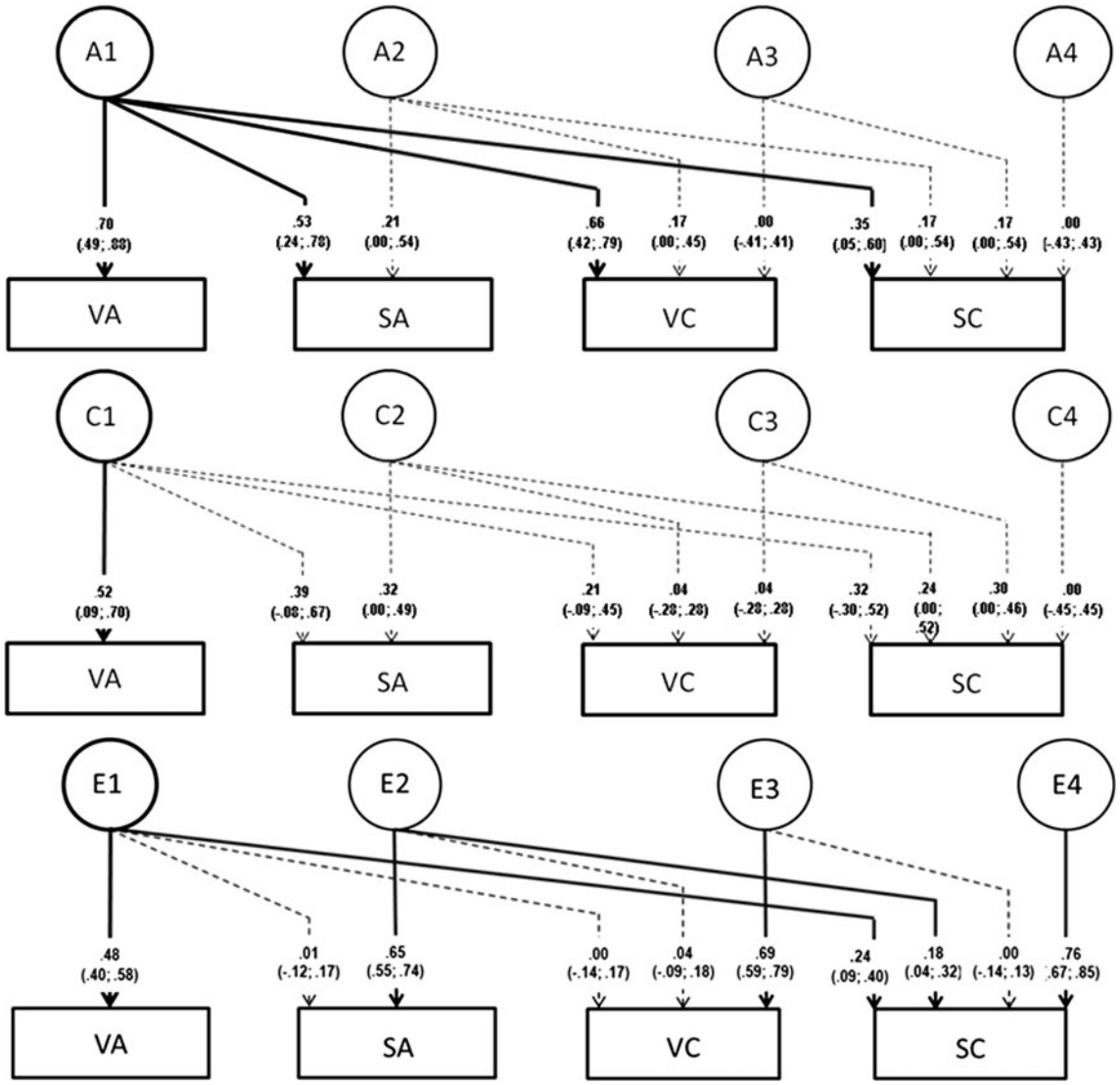

Fig. 3 Full Cholesky decomposition (ACE) showing genetic and environmental influences on the relationship between verbal analogy (VA), spatial analogy (SA), verbal category (VC), and spatial category (SC). Non-significant pathways in the model ($p > 0.05$) were retained for completeness and are shown as *dashed lines*. All estimates are standardized

References

Alarcon M, Plomin R, Fulker DW, Corley R, DeFries JC (1998) Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. Behav Genet 28(4):255–264

Alarcon M, Plomin R, Fulker DW, Corley R, DeFries JC (1999) Molarity not modularity: multivariate genetic analysis of specific cognitive abilities in parents and their 16-year-old children in the Colorado Adoption Project. Cogn Dev 14(1):175–193

Blair C (2006) How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. Behav Brain Sci 29(2):109–116

Bunge SA, Wendelken C, Badre D, Wagner AD (2005) Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. Cereb Cortex 15(3):239–249

Casto SD, Defries JC, Fulker DW (1995) Multivariate genetic-analysis of Wechsler Intelligence Scale for Children Revised (WISC-R) factors. Behav Genet 25(1):25–32

Deary IJ, Spinath FM, Bates TC (2006) Genetics of intelligence. Eur J Hum Genet 14(6):690–700

Edwards L, Figueras B, Mellanby J, Langdon D (2011) Verbal and spatial analogical reasoning in deaf and hearing children: the role of grammar and vocabulary. J Deaf Stud Deaf Educ 16(2): 189–197

Goswami U (1989) Relational complexity and the development of analogical reasoning. Cogn Dev 4:251–268

Green AE, Fugelsang JA, Dunbar KN (2006a) Automatic activation of categorical and abstract analogical relations in analogical reasoning. Memory Cogn 34(7):1414–1421

Green AE, Fugelsang JA, Kraemer DJM, Shamosh NA, Dunbar KN (2006b) Frontopolar cortex mediates abstract integration in analogy. Brain Res 1096:125–137

Green AE, Kraemer DJM, Fugelsang JA, Gray JR, Dunbar KN (2010) Connecting long distance: semantic distance in analogical reasoning modulates frontopolar cortex activity. Cereb Cortex 20(1):70–76

Haworth CMA, Wright MJ, Luciano M, Martin NG, de Geus EJC, van Beijsterveldt CEM et al (2010) The heritability of general cognitive ability increases linearly from childhood to young adulthood. Mol Psychiatry 15(11):1112–1120

Hoekstra RA, Bartels M, Boomsma DI (2007) Longitudinal genetic study of verbal and nonverbal IQ from early childhood to young adulthood. Learn Individ Differ 17(2):97–114

Hoekstra RA, Bartels M, van Leeuwen M, Boomsma DI (2009) Genetic architecture of verbal abilities in children and adolescents. Dev Sci 12(6):1041–1053

Horn JL, Cattell RB (1966) Refinement and test of the theory of fluid and crystallized general intelligences. J Educ Psychol 57(5): 253–270

Kane MJ, Engle RW (2002) The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: an individual-differences perspective. Psychon Bull Rev 9(4):637–671

Kasriel J, Eaves L (1976) Zygosity of twins—further evidence on agreement between diagnosis by blood-groups and written questionnaires. J Biosoc Sci 8(3):263–266

Langdon DW, Warrington EK (1995) The VESPAR: a verbal and spatial reasoning test. Lawrence Erlbaum Associates, Hove

Langdon D, Warrington EK (2000) The role of the left hemisphere in verbal and spatial reasoning tasks. Cortex 36(5):691–702

Mellanby J, Langdon D (2010) Verbal and spatial reasoning test for children. Cambridge Assessment, Cambridge

Neale M, Cardon LR (1992) Methodology for genetic studies of twins and families. Dordrecht, The Netherlands

Neale MC, Boker SM, Xie G, Maes HH (2006) Mx: statistical modeling, 7th edn. Department of Psychiatry, Richmond

Petrill SA (1997) Molarity versus modularity of cognitive functioning? A behavioral genetic perspective. Curr Dir Psychol Sci 6(4): 96–99

Posthuma D, de Geus EJC, Boomsma DI (2001) Perceptual speed and IQ are associated through common genetic factors. Behav Genet 31(6):593–602

Posthuma D, Baare WFC, Pol HEH, Kahn RS, Boomsma DI, De Geus EJC (2003) Genetic correlations between brain volumes and the WAIS-III dimensions of verbal comprehension, working memory, perceptual organization, and processing speed. Twin Res 6(2):131–139

Posthuma D, de Geus EJC, Deary IJ (2009) The genetics of intelligence. In: Goldberg TE, Weinberger DR (eds) The genetics of cognitive neuroscience. The MIT Press, Cambridge

Richland LE, Morrison RG (2010) Is analogical reasoning just another measure of executive functioning? Front Hum Neurosci 4:1–2

Richland LE, Morrison RG, Holyoak KJ (2006) Children's development of analogical reasoning: insights from scene analogy problems. J Exp Child Psychol 94(3):249–273

Rijsdijk FV, Vernon PA, Boomsma DI (2002) Application of hierarchical genetic models to Raven and WAIS subtests: a Dutch twin study. Behav Genet 32(3):199–210

Spearman C (1904) "General intelligence" objectively determined and measured. Am J Psychol 15:201–292

Spinath FM, Ronald A, Harlaar N, Price TS, Plomin R (2003) Phenotypic g in early life: on the etiology of general cognitive ability in a large population sample of twin children aged 2-4 years. Intelligence 31:195–210

van Leeuwen M, van den Berg SM, Hoekstra RA, Boomsma DI (2009) The genetic and environmental structure of verbal and visuospatial memory in young adults and children. Neuropsychology 23(6):792–802

van Soelen ILC, Brouwer RM, van Leeuwen M, Kahn RS, Pol HEH, Boomsma DI (2011) Heritability of verbal and performance intelligence in a pediatric longitudinal sample. Twin Res Hum Genet 14(2):119–128

Wechsler D (1955) Manual for the Wechsler adult intelligence scale. The Psychological Corporation, New York

Wechsler D (1992) Wechsler intelligence scale for children—third edition UK (WISC-IIIUK). The Psychological Corporation, London

Wright MJ, Martin NG (2004) Brisbane adolescent twin study: outline of study methods and research projects. Aust J Psychol 56(2):65–78