General cognitive ability and the interplay between genes and environment

Vinkhuyzen, A.A.E.

2010

Chapter 2

Theoretical background, sample characteristics, and data collection

OUTLINE OF THIS CHAPTER

In the first part of this chapter, a theoretical background is provided of twin-family modeling, together with a description of complex processes such as gene-environment interaction and correlation, assortative mating, and cultural transmission that may underlie individual differences in general cognitive ability. We will describe how these mechanisms may affect the estimates of genetic and environmental influences such as obtained in classical twin studies, in which those effects cannot be modeled. In addition we will describe how these processes can be modeled within an extended twin family design.

In the second part of this chapter, a description of the sample, on which most of the studies in this thesis are based, and data collection on which studies in chapters three to eight are based, is given. The actual measures of cognitive ability and environmental indices that were reported in this PhD thesis are described in the third part of this chapter.

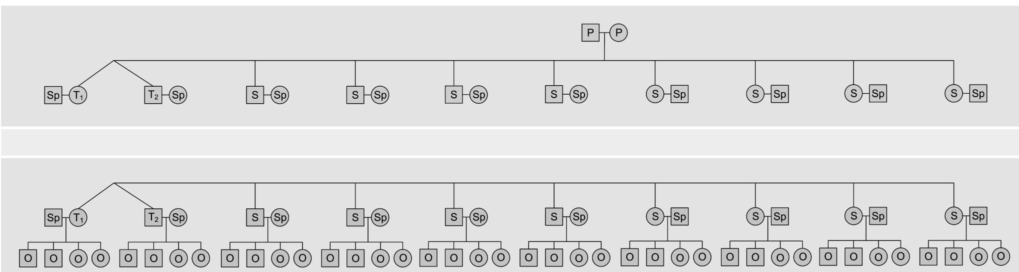
THEORETICAL BACKGROUND

The extended twin family design

In classical twin studies, data from monozygotic (MZ) and dizygotic (DZ) twins are used to decompose the total variance of a trait, also called the 'phenotypic or observed variance', into two main sources of variation: genetic variation and environmental variation. Based on the observed MZ and DZ twin correlations, as well as their known genetic and environmental relatedness, the relative proportions of genetic and environmental factors to the overall variation in a trait can be estimated (Falconer & Mackay, 1989).

To deal with the limitations of the classical twin design as discussed in Chapter 1, we extended the classical twin design by including siblings of twins, spouses of twins and of their siblings, and either the parents or the offspring of the twins and siblings (depending on age). Figure 2.1 shows two pedigrees for the extended twin family design as used in this PhD project.

Figure 2.1 Pedigrees for the two extended twin family designs as applied in this thesis



Notes: P = parent; T = twin; S = non-twin sibling, Sp = spouse; O = offspring of twins and their siblings; squares denote men; circles denote women. Please note that this figure illustrates the idealized pedigree: the maximum number of parents observed in one sample is two, the maximum number of twins is two, the maximum number of siblings is eight and the maximum number of offspring of twins or siblings is four.

Whereas in the classical twin design, only two crucial relationships are distinguished (i.e., MZ and DZ twin pairs), the extended twin family design, as used in this thesis, holds fifteen familially different relationships that can be increased to 22 if relationships involving DZ twins are distinguished from relationships involving siblings. Moreover, a total number of 102 relationships can be distinguished if the gender of the relatives is taken into account (See Appendix II for an overview of all possible relationships within the extended twin family design). When assuming the absence of sex differences and DZ-sibling differences, this large number of familial relationships can be reduced to six unique relations that provide information to the estimates of genetic and environmental factors. In Table 2.1, coefficients of the additive genetic covariance, genetic dominance covariance, and shared and non-shared environmental covariance are provided for all six distinctive pairs of relatives within the present study.

Table 2.1 Coefficients for the additive genetic and genetic dominance components and shared and non-shared environmental components of the covariance between relatives for an equilibrium population under random mating (extended from Lynch and Walsh, 1998, Table 7.3, page 148).

Relationship	$\hat{\sigma}_A^2$	$\hat{\sigma}_D^2$	$\hat{\sigma}_C^2$	$\hat{\sigma}_E^2$
MZ	1	1	1	0
DZ/sibs	½	¼	1	0
PO / AVMZ	½	0	0	0
AVDZ / COMZ	¼	0	0	0
CODZ	⅛	0	0	0
IN-LAWS	0	0	0	0

Notes: Notation follows Lynch and Walsh (1998): $\hat{\sigma}_A^2$ = standardized additive genetic variance; $\hat{\sigma}_D^2$ = standardized genetic dominance variance; $\hat{\sigma}_C^2$ = standardized shared environmental variance, $\hat{\sigma}_E^2$ = standardized non-shared environmental variance. Assortative mating is assumed to be absent. Correlations are assumed equal across twins and regular siblings and across sex; MZ=twin-twin MZ; DZ/sibs=twin-twin DZ/sibling; PO=parent-offspring; AVMZ=cousins avuncular through MZ; AVDZ=cousins avuncular through DZ/sibling; COMZ=nieces/nephews through MZ; CODZ=niece/nephews through DZ/sibling; IN-LAWS represent sister/brother in law through MZ, sister/brother in law through DZ/sibling, spouse-spouse through MZ, spouse-spouse through DZ/sibling, aunt/uncle-cousin in law through MZ, aunt/uncle-cousin in law through DZ/sibling, and parent-offspring in law.

Including parental information in the twin family design allows the assessment of additive genetic influences, genetic dominance deviation, effects of assortative mating, shared- and non-shared environmental influences, parental influences (i.e., cultural transmission), and $r_{(GE)}$ that is induced by the co-occurrence of additive genetic influences and cultural transmission. Additionally, including information on the spouses of the twins allows the resolution of the mechanism underlying assortative mating (e.g., phenotypic assortment and social homogamy) and the test whether the extent and type of assortative mating is equal over generations. Including information on the children of the twins allows one to test whether effects of cultural transmission remain equal over generations. Siblings of the twins, as well as the spouses and children of the siblings, provide information on whether twins and siblings differ with respect to (individual differences in) general cognitive ability, and these additional data increase statistical power to detect shared environmental and genetic influences (Posthuma & Boomsma, 2000).

COMPLEX GENE-ENVIRONMENT MECHANISMS

Gene environment correlation: $r_{(GE)}$

$r_{(GE)}$ refers to a situation where genes control an individuals' exposure to environmental factors, or a situation in which gene frequencies differ in different environments. In this state, environmental factors that influence an individuals' phenotype are not a random sample of the entire range of possible environments, but are correlated with, or a function of, the genotype of an individual. Usually, three different types of gene-environment correlation are distinguished (Plomin et al., 1977): passive, evocative and active $r_{(GE)}$. Passive $r_{(GE)}$ occurs when parents transmit both their genotypes and for the trait relevant environmental factors. For example, intellectual gifted parents may transmit genetic variants that are related to high cognitive ability and also provide their children with 'intelligence boosting' experiences. Evocative $r_{(GE)}$ occurs when people's behavior towards an individual is a reaction to the genetic predisposition of the individual. For example, different genetic makeup may induce different treatment by teachers in the classroom. Active $r_{(GE)}$ occurs when individuals actively shape or seek out their own environment, based on their genetic predisposition. That is, individuals will select environments that fit their genetic predisposition and construct an environment in which they can thrive.

When any of these forms of $r_{(GE)}$ occur, the effects of genes can no longer be considered independent of the effects caused by environmental factors. Ignoring the $r_{(GE)}$ in statistical genetic models may lead to biased estimates of the relative importance of both genetic and environmental factors (Purcell, 2002; Eaves et al., 1977). If $r_{(GE)}$ concerns the correlation between genes and *non*-shared environmental factors, ignoring it's presence will result in overestimation of the effects of the genetic factors, while ignoring $r_{(GE)}$, if it concerns the correlation between genes and *shared* environmental factors, will result in overestimation of the effects of the shared environmental factors (Purcell, 2002). Given the high estimates of genetic factors and absence of shared environmental factors for general cognitive ability, potential $r_{(GE)}$ is expected to concern the correlation between genes and non-shared environmental factors.

Gene environment interaction (GEI)

GEI refers to a situation where genes control an individual's sensitivity to environmental factors, or a situation in which the environment controls the (level of) expression of genes. For example, individual differences in genetic makeup related to general cognitive ability may become more pronounced and visible when educational quality is low, compared to when educational quality is high.

In the presence of GEI, the relative contribution of genetic and environmental factors to individual differences in a particular phenotype fluctuates across different environmental and genetic conditions. Consequently, point estimates of genetic and environmental effects may not accurately reflect the total range of heritabilities across all levels of an environmental factor.

When GEI concerns interaction between genes and *non*-shared environmental factors, ignoring the influences of GEI in statistical models will result in overestimation of the effects of the non-shared environmental factors. If GEI interaction concerns the interaction between genes and *shared* environmental influences, ignoring its presence will result in overestimation of the effect of genetic factors (Jinks & Fulker, 1970; Eaves et al., 1977; Purcell, 2002). As estimates of genetic factors for general cognitive ability are generally high, interaction between genes and shared environmental factors rather than interaction between genes and non-shared environmental factors is expected. Explicit modeling of GEI effects is thus important if one wishes to distinguish and understand the factors that cause individual differences in general cognitive ability.

Assortative mating

Behavior geneticists speak of 'assortative mating' when mates show higher or lower resemblance than expected by chance. Positive assortative mating may increase while negative assortative mating may decrease the genetic and environmental correlations between mates. This, in turn, will respectively increase or decrease the genetic resemblance between e.g. DZ twin pairs, relative to MZ twin pairs. Three main processes of assortative mating can be distinguished: social homogamy, phenotypic assortment and social (or marital) interaction; these processes can occur simultaneously. Social homogamy refers to the situation in which individuals who grow up in a similar social background are more likely to mate with individuals from that same background. That is, under social homogamy, assortment takes place within groups that are differentiated environmentally (Falconer & Mackay, 1989). Phenotypic assortment refers to the situation in which mates select each other on the basis of similarities (positive phenotypic assortment) or dissimilarities (negative phenotypic assortment) in observable characteristics, such as general cognitive ability. That is, given positive phenotypic assortment, individuals with high levels of cognitive ability tend to mate individuals with high levels of cognitive ability, while individuals with lower levels of cognitive ability tend to mate individuals with lower levels of cognitive ability. Social interaction refers to the situation in which mates resemble each other more and more as a consequence of living together and influencing each others' behavior, such that the longer mates interact, the greater their resemblance. Since the data in the present thesis are not sufficient to model social interaction, only social homogamy and phenotypic assortment are considered in this thesis.

Social homogamy and phenotypic assortment can be studied by comparing spousal resemblance, as expressed with a phenotypic correlation, of different spouse combinations (e.g., *direct spouses*: a twin and his/her spouse and parental spouse pairs; *cross-sibling-spouse pairs*: a twin and the spouse of his/her co-twin; *spouse-spouse pairs*: the spouse of a twin and the spouse of the twins' co-twin) within a sample. Assortative mating due to social homogamy reflects a shared social environment and therefore the effect of sharing an environment on the mating process is similar for all relatives. No difference is expected between different spouse pairings. In the situation of pure social homogamy, the correlation between direct spouses is therefore expected to be similar to the correlation between cross-sibling-spouse pairs and spouse-spouse pairs.

In the situation of phenotypic assortment, however, mate selection is purely based on the observed phenotype, which may be influenced by both genetic and environmental factors. Under positive phenotypic assortment on a heritable trait, observed phenotypic correlations between spouses are expected to decline with the genetic distance of the relationship between family members; correlations of spouse-spouse pairs and cross-sibling-spouse pairs are expected to be lower than the correlation of direct spouse pairs. Since MZ twins have more genes in common compared to DZ twins, cross-sibling-spouse correlations and spouse-spouse correlations are expected to be different for MZ and DZ twin pairs, depending on the extent to which the phenotype under study reflects the genotype. Consequently, in the situation of positive phenotypic assortment on a heritable trait, such as cognitive ability, cross-sibling-spouse and spouse-spouse correlations are expected to be higher for MZ twins.

Possible consequences for the estimates of genetic and environmental factors in case assortative mating is not considered, depend on the process underlying the mate selection, i.e., social homogamy or phenotypic assortment. In the situation of pure social homogamy, spousal correlations resulting from pure social homogamy are purely environmental. Consequently, no genetic correlation between mates exists, and the contribution of genetic factors is not affected (Falconer & Mackay, 1989).

If positive phenotypic assortment is the underlying process for mate selection, additive genetic factors and dominance genetic factors will be correlated in spouse pairs (Fisher, 1918; Rice et al., 1978; Heath & Eaves, 1985). If increased genetic resemblance between relatives is not considered in a twin-family model, estimates of genetic and environmental parameters will be biased.

Cultural transmission

Parents may transmit not only their genetic material, but also their environment to their children. Cultural transmission refers to the transmission of environmental factors that are related to a trait (e.g., general cognitive ability) from the parental phenotype to the offspring's environment. Since children who are raised in the same home grow up within a common environment as created by the parents, cultural transmission is by definition part of the shared environment in the offspring generation.

When cultural transmission exists in the presence of genetic transmission, environmental influences become correlated with genetic influences ($r_{(GE)}$). Consequently, when cultural transmission and the resulting $r_{(GE)}$ are not modeled, parent-offspring correlations, but also twin/sibling correlations, will exceed correlations that are expected under genetic transmission alone, leading to inflated estimates of shared environmental factors.

SAMPLE CHARACTERISTICS AND DATA COLLECTION

This thesis reports on the results from a large ongoing project on the genetics of cognition (Posthuma et al., 2001a). The study was initiated in 1997 with the collection of anatomical, electrophysiological and behavioral indices of cognitive ability. Data on behavioral indices of cognitive ability from 788 twins and their non-twin siblings were collected between 1997 and 2001 (first wave of data collection), see Posthuma (2002) for an extensive description of sample characteristics and data collection.

To be able to model complex processes that may underlie individual differences in cognitive ability, the existing sample was extended with partners and either the parents or the children of the twins and siblings, as well as twins and siblings that were not measured in the first wave of data collection. All relatives were subjected to an extensive test protocol to measure general cognitive ability between 2007 and 2009 (second wave of data collection). In addition, questionnaire data on environmental indices were collected between 2007 and 2009 in all participants (i.e., twins, siblings, their spouses and either their parents or their children) using the Life Experiences List (LEL, developed within this project). Saliva samples were also collected as part of a companion PhD project (T. Sampaio Rizzi).

Within the first wave of data collection, data on cognitive ability were collected at the laboratory of the VU University Amsterdam. Within the second wave of data collection, data were collected either at the laboratory of the VU University Amsterdam or at the participants' home, depending on preference of the participants. The majority of the participants preferred testing at home (60%). Both waves of data collection were performed with understanding and written consent of each participant. The study was approved by the Central Committee on Research Involving Human Subjects in the Netherlands.

Twins and siblings that were part of the first wave of data collection were registered at the Netherlands Twin Register (NTR; Boomsma et al., 2006). To recruit participants for the second wave of data collection, non-registered family members of the twins and their non-twin siblings were contacted after permission from their registered family members. All participants were registered in the NTR at the time of measurement.

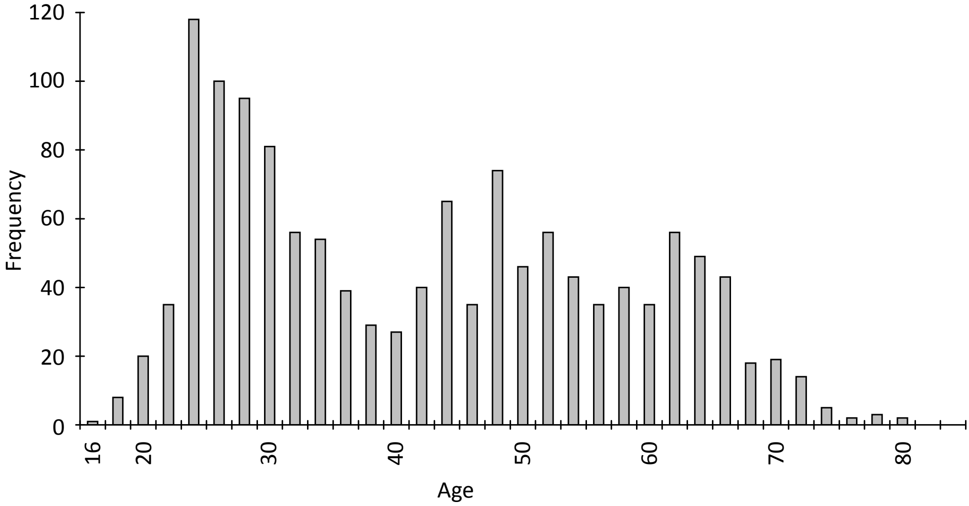
In total, data were available for 1419 participants. Data on cognitive ability were available for 1350 participants (317 families, 45.9% men), and questionnaire data were available for 1072 participants (260 families, 44.3% men). 69 participants did supply questionnaire data but data on cognitive ability were absent. 337 participants did have data on cognitive ability, but questionnaire data were missing. On average 4.46 (SD = 2.57) subjects per family participated (range 1-22, median = 4). In Table 2.2 the composition of the sample is summarized, together with frequencies of relatives with data on cognitive ability and on environmental indices. The relatively large number of 337 participants that did not have questionnaire data is partly due to the fact that 17.4% of the participants from the first wave of data collection were no longer registered in the NTR (137 participants) at the time of the second wave of data collection. The overall response rate for the LEL was 76%.

Table 2.2 Sample sizes (individuals) by type of relative for cognitive ability and environmental indices

	Cognitive ability	Environmental indices (LEL)
Twins	603 (including 3 triplets)	388 (including 1 triplet)
Siblings of twins	263	172
Parents of twins/sibs	152	167
Spouses of twins	139	138
Spouses of sibs	18	18
Children of twins	140	149
Children of sibs	34	39
Spouse of child of twin	1	1
Total	1350	1072

Notes: relatives are expressed in relation to the twins; LEL = Life Experiences List.

Figure 2.2 Age distribution of the total sample



Notes: the age displayed in the above figure is the age of the participants at the time of the cognitive assessment

The age distribution of the sample is depicted in Figure 2.2. Average age of the participants was 41.11 years (SD = 15.06; range: 15.71 – 79.87) at the time they completed the cognitive assessment.

Zygosity

Zygosity of same-sex twins was based on DNA polymorphisms (114 pairs, 83%) or, if information on DNA markers was not available, on questions about physical similarity and confusion of the twins by family members and strangers. Agreement between zygosity diagnoses from survey and DNA was 97% (Willemsen et al., 2005). All five zygosity groups were well represented: monozygotic males (MZM: 20.6%), monozygotic females (MZF: 25.4%), dizygotic males (DZM: 12.4%), dizygotic females (DZF: 22.4%) and dizygotic opposite sex (DOS: 19.2%).

Cognitive ability

The cognitive test protocol of the second wave of data collection lasted two and a half up to three hours, including break and formal precedures. The order of the tests is shown in Table 2.3. The Verbal Learning & Memory Task exists of two parts with an interval of twenty minutes between the end of the first part and the start of the second part. To avoid any verbal interference, no verbal or linguistic tests were administered within this interval. During the entire assessment, verbal conversation was restricted to a minimum.

Table 2.3 Cognitive test protocol as assessed within the second wave of data collection

Phenotype	Test
Psychometric IQ	WAIS-IIIR (Wechsler, 1997)
Verbal learning and memory	Verbal Learning & Memory Task (part 1) (Mulder et al., 1996)
Time perception	Time Perception Application (Barkley, 1998)
Visuo-spatial memory	Corsi Block Tapping Task (Corsi, 1972)
Executive functioning	Trail Making Test (Reitan, 1955)
Verbal learning and memory	Verbal Learning & Memory Task (part 2)
Inspection Time	Inspection Time Task (Luciano et al., 2001)
Linguistic ability	Non Words Test (reading/repeating)
Linguistic ability	Word Stress Task (Schiller, 2006)
Behavior during test and characteristics of test setting and participant.	Task Observation Form

During the assessment, a Task Observation Form was filled out by the test administrator. Characteristics of the test setting (i.e., test location, time of the day and particularities that may influence the assessment) and the participants (i.e., physical and mental conditions that may influence the assessment) were recorded. Participants were paid €25,- if they completed the cognitive test protocol. In case participants were tested at the laboratory of the VU University Amsterdam, traveling-expenses were refunded.

DESCRIPTION OF MEASURES OF COGNITIVE ABILITY AND ENVIRONMENTAL INDICES

In the last part of this chapter, the measures are described on which the studies reported in Chapters 3 to 8 are based. The cognitive measures that were collected during this PhD project but that were not reported on in this thesis, are described in appendix III. Statistical descriptives and phenotypic correlations between all measures of cognitive ability that were collected within this project are provided in appendix IV and appendix V, respectively.

General cognitive ability

General cognitive ability (or psychometric IQ) was assessed with the Dutch version of the WAIS-IIIR. Participants whose IQ was assessed in the first wave of data collection (785 participants: twins and siblings) completed eleven subtests of the WAIS-IIIR: Block design, Letter-number sequencing, Information, Matrix reasoning, Similarities, Picture completion, Arithmetic, Vocabulary, Digit symbol-coding, Digit-symbol pairing and Digit symbol-free recall. Participants whose IQ was assessed in the second wave of data collection (617 participants: twins, siblings, parents, offspring of twins and siblings and spouses of twins and siblings) completed seven subtests of the WAIS-IIIR: Block design, Letter-number sequencing, Information, Matrix reasoning, Arithmetic, Vocabulary and Digit symbol-coding. Correlation between full scale IQ assessed with eleven subtests and full scale IQ assessed with seven subtests was substantial (Pearson's r = .97, N = 785, p<.001). To measure test-retest reliability, 59 participants participated in both the first and the second wave of data collection. Test-retest reliability for full scale IQ was substantial (Pearson's r = .85, N = 59, p<.001), full scale IQ was based on seven subtests at both time points for computation of the test-retest reliability. For those 59 subjects whose IQ was assessed twice, data from the first wave data collection were used in the analyses. In total, sex and age corrected IQ scores were available for 1343 participants.

Environmental Indices

Environmental indices were measured with the Life Experiences List (LEL). The LEL was developed to measure environmental indices and individual traits and qualities that were expected to be related to cognitive ability and/or individual differences in cognitive ability. The LEL consisted of a combination of acknowledged questionnaires as well as questions that were specifically developed for this project. An overview of the subjects covered by the LEL is shown in Table 2.4. The combination of cognitive and environmental data allows us to investigate the complex interplay between genes and environment in the context of general cognitive ability.

Table 2.4 Indices of the Life Experiences List (LEL)

Demographics
Response date, first name, gender, birth date, and zip code
Family composition
Childhood living situation, partner selection, and current family size
Sport
Frequency of participation in specific sports: current and past (between ages 6-18 years). Way of travelling to work/school (e.g., walk, bike, scooter, car, train or bus), and average daily travel time
Music and other leisure time activity
Frequency of participation in specific music and other leisure time activities: current and past (between ages 6-18 years)
Leisure time activity
A list of leisure time activities used to quantify time spend on certain activities: current and past (between ages 6-18 years)
Parental rearing style and general childhood
Subjective assessment of parental rearing style Subjective assessment of parental attitude towards educational attainment Questions on reading behavior during childhood Information on breastfeeding Information on body composition compared to peers at elementary school and high school Information on teasing by peers at elementary school and high school
Life events
Information on influential Life events (past five years and over lifetime).
Family functioning
General Functioning subscale of the Family Assessment Device (Epstein et al., 1983) Conflict subscale of the Family Environment Scale (Moos & Moos, 1976)
Own rearing style and leisure activities with offspring
Subjective assessment of own rearing style Subjective assessment of own attitude towards educational attainment offspring Questions on reading behavior
Education and profession
Received education and current work status
Education and profession of parents, partner and close friend
Received education and current work status

Social support and social behavior
Social Support Questionnaire-6 (Sarason et al., 1987; Sarason et al., 1990). Two aspects of perceived social support: (i) number of people available in six specified problem situations (ii) degree of satisfaction with the perceived support
Decomposed Game Measure (Van Lange et al., 1997). Someone's general tendencies toward others.

Achievement motivation
Achievement Motivation Test (short version) (Hermans, 1970). This questionnaire assesses general and school related achievement motivation

Experience seeking
Experience Seeking scale of the Sensation Seeking List (Zuckerman, 1996). Degree of unconformable life style and desire to expand one's experience through the mind and senses.

Autistic traits
Autism Quotient (short version) (Baron-Cohen et al., 2001). This questionnaire quantifies autistic traits in the normal population

Self report
Young Adult Self Report (short version) (Achenbach, 1990; Achenbach, 1997; Achenbach, 1991). Only items related to the attention subscale were included