Testing distributional assumptions in psychometric measurement models with substantive applications in psychology

Molenaar, D.

Publication date
2012

7

Heteroscedasticity due to Genotype by Environment Interactions: An application to Cognitive Ability Data from 14 Different Studies.

Cognitive ability is known to be considerably heritable. Researchers have tried to identify environmental variables that influence the heritability of cognitive ability, indicating a genotype by environment interaction. To date, environment variables that were found of interest include measured variables like income and socioeconomic status. The present Chapter focuses on genotype by environment interaction in cognitive ability where the environment variable is the unmeasured unique environmental factor, E, from the ACE-decomposition. We tested this in the GHCA-database (Haworth et al, 2010), which comprises data of 14 different cognition studies from 4 different countries including subjects of different ages. Results indicate that for younger subjects (4-13 years), the strength of E decreases across the additive genetic factor A, but that this effect reverts for older subjects (17-34 years). Interestingly, substantial heterogeneity is observed between the individual studies with respect to environmental en genetic influences on cognitive ability. Possible sources of this heterogeneity are discussed.

7.1 Introduction

Genetic and environmental influences on individual differences in cognitive ability enjoyed extensive investigation (see Plomin & Spinath, 2004, for an overview). Using family-based designs, in which phenotypic variance is decomposed into additive genetic, unique environmental, and shared environmental effects (denoted A, E, and C, respectively), heritability estimates have been reported between roughly .5 and .7 for adolescents and adults (see McGue, 1997). For young children the heritability is somewhat lower than .5 (see Haworth et al., 2010), and for infants as young as 10 months, the heritability is appreciably lower (heritability approaches 0; Tucker-Drob et al., 2011).

* This chapter is based on: Molenaar, D., van der Sluis, S., Boomsma, D.I., Haworth, C.M.A, Hewitt, J.K., Plomin, R., Wright, M.J., & Dolan, C.V. (2011). *Genotype by Environment Interactions in Cognitive Ability Tested in 14 Different Studies.* Manuscript in preparation.

Having established that, at least beyond childhood, genetic factors explain a substantial part of the phenotypic variance in cognitive abilities, question rises whether the heritability of cognitive ability is constant across the range of environmental effects. We consider variation in heritability as a function of an environmental variable to be a potential manifestation of Genotype by Environment interaction (GxE). Measures that have shown to influence or moderate the heritability of cognitive ability include parental income (Harden, 2007), socioeconomic status (Tucker-Drob et al. 2011; Turkheimer et al., 2003), parental education (Grant et al., 2010; van der Sluis et al., 2008), and educational attainment (Johnson, Deary & Iacono, 2009). It is clear that the moderators in these studies are not strict environmental measures, although they may have a strong influence on the effective environment (Plomin & Daniels, 1987).

The present Chapter focuses on genotype by environment interaction in cognitive ability where the environment variable is the unmeasured unique environmental factor, E, rather than some measured aspect of the environment (e.g., SES). Testing for GxE with E unobserved is of interest in the light of the ability differentiation hypothesis (Spearman, 1927). This hypothesis states that the magnitude of the correlations among the subtests of various cognitive abilities varies with intelligence itself: the correlations are supposed to be relatively higher in samples of subjects who score relatively low on general intelligence (*g;* Jensen, 1998). Given that *g* is substantially heritable, this implies that twins with a high *g* level (and thus a high level on A) should have a smaller intra twin correlation between a given IQ measure as compared to twins with a lower level on *g*. One way in which this may come about is when the unique environmental variance is lower at higher levels of A. To our knowledge, this specific implication of the differentiation hypothesis has not been addressed yet.

In addition, testing for GxE with unmeasured E is of interest in the context of Genome-Wide Association Studies (GWAS). Several authors have argued that the failure of GWAS to detect genetic variants that explain a substantial part of the phenotypic heritability might at least partly be due to GxE (see e.g. Eichler et al., 2010; Maher, 2008; Manolio et al., 2009).

In the present article, we analyze data on cognitive ability from the GHCA database (Genetics of High Cognitive Abilities Consortium; Haworth, 2010). This database comprises IQ scores from 14 individual studies conduced within 4 different countries: US, UK, Australia, and the Netherlands. In this database, Haworth et al. showed that there is a linear relation between heritability and age. However, data was collapsed over all countries, not taking into account differences between the individual studies in heritability and average age. In this article we will analyze the GHCA database, taking into account all 14 individual studies. Primarily, we test for GxE in these data, but we will shortly check whether the Haworth et al findings can be

replicated on the level of the individual studies. To test for GxE, we use the method recently proposed by Molenaar, van der Sluis, Boomsma, & Dolan (2011). This method is related to the Jinks and Fulker Test (1970; see also van der Sluis et al, 2009), but has the advantage of including data of both MZ and DZ twins. Outline is as follows: We first present the GxE-model, next, we describe the data in the GHCA database and we apply the GxE model to these data. We end with a discussion.

7. 2 The GxE model

Let Y_{ij} denote the phenotypic score of the *j*-th twin (j = 1, 2) of the *i*-th twin pair (i = 1, …, N). In the standard ACE-model, Y_{ij} is submitted to the following linear model

$$Y_{ij} = \nu + A_{ij} + C_{ij} + E_{ij,} \quad (1)$$

where

$$\begin{aligned} &\text{cor}(C_{i1,}\, C_{i2}) = 1, \\ &\text{cor}(E_{i1,}\, E_{i2}) = 0, \\ &\text{cor}(A_{i1},\, A_{i2}) = 1, && \text{in MZ twins,} \\ &\text{cor}(A_{i1},\, A_{i2}) = 0.5, && \text{in DZ twins,} \end{aligned}$$

i.e., the phenotypic score consists of a summation of an intercept (ν) and the scores on an additive genetic factor (A_{ij}), a common environmental factor (C_{ij}), and a unique environmental factor (E_{ij}). Under the assumption that A_{ij}, C_{ij}, and E_{ij} are uncorrelated, the phenotypic variance σ_Y^2 can be decomposed as follows:

$$\sigma_Y^2 = \sigma_A^2 + \sigma_C^2 + \sigma_E^2. \quad (2)$$

Within the ACE-model, Molenaar et al. (2011) distinguished between two possible GxE interactions, i.e., the interaction between A and C, and the interaction between A and E. To model these two interactions within the ACE-model, the phenotypic variance is conditioned on A, i.e.,

$$\sigma_{Y|A}^2 = \sigma_{C|A}^2 + \sigma_{E|A}^2. \quad (3)$$

Now, GxE can be modeled by allowing the variance of C and E to depend on the level of A. For instance, the variance of E can be decreasing across A, which would mean that standardized heritability increases as environmental influences diminish. To this end, a mathematical function should be specified between A and σ_E^2, and between A

and σ_C^2. Molenaar et al., (2010), van der Sluis et al (2009), and Hessen & Dolan (2009) proposed to use the exponential function, as this function precludes negative values for the variance components. Thus, the ACE-model can be extended to include GxE in the following way:

$$\sigma_{Y|A}^2 = \sigma_{C|A}^2 + \sigma_{E|A}^2 = \exp(\gamma_0 + \gamma_1 A) + \exp(\beta_0 + \beta_1 A), \quad (4)$$

where γ_0 and β_0 are baseline parameters which account for the part of the variance of C and E that does not depend on A, while γ_1 and β_1 are linear interaction parameters, which model the possible dependency of respectively the variance of C and E on A. Testing for an AxE interaction, for example, involves testing whether β_1 departs significantly from 0. Molenaar et al (2011) showed that the power to detect an AxC interaction is relatively poor. However, the power to detect AxE is generally good. In this Chapter, we will thus mainly focus on AxE, but we do take into account the possibility of AxC.

Equation 4 could be extended to include curvilinear effects, see van der Sluis et al (2009), i.e.,

$$\sigma_{Y|A}^2 = \sigma_{C|A}^2 + \sigma_{E|A}^2 = \exp(\gamma_0 + \gamma_1 A + \gamma_2 A^2) + \exp(\beta_0 + \beta_1 A + \beta_2 A^2), \quad (5)$$

where β_2 and γ_2 are curvilinear interaction parameters.

Molenaar et al. (2011) showed that this GxE model can be fitted to data of both MZ and DZ twins using Marginal Maximum Likelihood (MML; Bock & Aitkin, 1981). In this procedure, the observed data conditional on A is assumed to follow a normal distribution. Note that the unconditional data can be non-normal, as the presence of GxE will generally result in non-normality. Using MML, the model could be fitted to data in the freely available software package Mx (Neale, Boker, Xie, & Maes, 2006). Both AxC and AxE can be combined in a single model, i.e., β_1, β_2, γ_1 and γ_2 can be estimated together (example script are available from www.dylanmolenaar.nl).

7.3 Application to GHCA data

7.3.1 Description of the data

The database comprises univariate IQ scores of 14 studies conducted in four countries: US (Colorado, 3 studies; Minnesota, 2 studies; Ohio, 1 study), UK (1 study), Australia (1 study), and the Netherlands (6 studies). Aggregating all of these data yields a total of 10897 twin pairs (4911 MZ pairs and 5986 DZ pairs). The age of the subjects varies from 4 to 71. For each twin in the database, an IQ measure is available. However,

across the different studies, different test batteries have been used to obtain IQ score. For instance, in one of the US studies, the short form of the Stanfort-Binet Intelligence Scale was used; in the UK study, 4 subtests of the WISC-III were administered; and in another study from the US, the full WISC-III was used. We refer to Haworth, et al. (2010) and the references therein, for a detailed overview of the IQ test batteries used in the different studies.

7.3.2 Plan of analysis

As the aggregated data is heterogeneous with respect to age, we follow Haworth et al. (2010), and perform the analysis within more age-homogenous subgroups. Haworth et al considered 3 age categories: 4-10, 11-13, and 14-34. All subjects above age 34 are omitted as too few subjects were in this age range to construct a reasonably homogenous subgroup with respect to age. In our analysis, we modified the age categories of Haworth et al because we considered it possible that the nature of GxE differs between adolescents (who presumably are still in school and live at home) and young adults. Specifically, we created the categories: 4-10, 11-13, 14-16, and 17-34. We note that the general pattern of results, as presented below, does not depend on the exact age categories that are used: results are generally the same for the Haworth et al age categories. However, our categorization does provide a clearer picture of how GxE changes across age. See Table 7.1 for the sample sizes within the age groups of both Haworth et al. and our age categorization.

Table 7.1
Age categories as used by Haworth et al and as used in the present study

	Haworth et al			**Present**		
Group no	*Age*	N_{MZ}	N_{DZ}	*Age*	N_{MZ}	N_{DZ}
1	4-10	1140	1613	4-10	1140	1613
2	11-13	2195	2722	11-13	2195	2722
3	14-34	1507	1582	14-16	807	966
4	-	-	-	17-34	700	616
total N	-	4842	5917	-	4842	5917

As can be seen in Table 7.1, in our alternative categorization, the third category of Haworth et al is split into one relatively homogenous age group (14-16), and an additional group (17-34). We also considered a second alternative in which this additional group was made more homogenous (i.e., subjects of ages 17-20), but this categorization did not alter the results as compared to the results obtained using our categorization in Table 7.1.

The aggregated data are potentially heterogeneous as they originated from 14 different studies. To establish whether the results within the age groups hold across the different studies, we conduct the analyses on both the aggregated data and on the data within each study. Table 7.2 provides the distribution of subjects over age categories and studies. Some studies include only subjects in one age category (e.g., Netherlands #1), while other studies include subjects in multiple age categories (e.g., Colorado #1). In some cases, a study contained too few subjects in a given age category to enable model fitting on this subsample (e.g., Netherlands #4 only contains 13 MZ twins and 14 DZ twins between 14 and 16 years old). Therefore, for our analyses, we only select 17 subsamples that we considered large enough to provide stable parameter estimates. These samples are shown in boldface print in Table 7.2. Note that we completely omitted Netherlands #5 because we considered the total sample size to be too small, and we omitted Netherlands #6, as this study only contains subjects of age 40-70. MZ and DZ twin correlations of the IQ measures in the 17 subsamples are presented in Table 7.3, together with the skewness, kurtosis, and the Shapiro-Wilks test on normality.

Table 7.2

Number of MZ and DZ twins (N_{MZ}; N_{DZ}) within each age category for each study in the GHCA database

Study	Total N	Age Group			
		4-10	*11-13*	*14-16*	*17-34*
Netherlands #1	125; 112	**125; 112**	0; 0	0; 0	0; 0
Netherlands #2	49; 63	**49; 63**	0; 0	0; 0	0; 0
Netherlands #3	79; 111	0; 0	**79; 111**	0; 0	0; 0
Netherlands #4	79; 108	0; 0	0; 0	13; 14	**66; 94**
Netherlands #5	36; 58	0; 0	0; 0	0; 0	36; 58
Netherlands #6	69; 69	0; 0	0; 0	0; 0	0; 0
Colorado #1	752; 1025	**391; 534**	**204; 280**	**126; 173**	31; 38
Colorado #2	215; 175	17; 6	13; 14	**153; 125**	32; 30
Colorado #3	332; 364	0; 0	30; 44	**99; 140**	**203; 180**
Minnesota #1	777; 469	43; 16	**734; 453**	0; 0	0; 0
Minnesota #2	410; 214	0; 0	0; 0	**92; 31**	**318; 183**
Australia	338; 515	0; 0	0; 0	**324; 483**	14; 32
Ohio	121; 171	**121; 171**	0; 0	0; 0	0; 0
UK	1529; 2532	**394; 714**	**1135; 1818**	0; 0	0; 0

Note. Samples used for the analysis of the individual studies are in boldface.

Table 7.3

MZ and DZ correlations, skewness, kurtosis and a test on normality of the IQ measures in the 17 subsamples within each age category.

Study	Cor		Skewness		Kurtosis		Norm. test	
	MZ	*DZ*	*T1*	*T2*	*T1*	*T2*	*T1*	*T2*
			Age Category 4-10					
Netherlands #1	0.70	0.49	-0.41	-0.56	2.99	3.56	0.01	0.00
Netherlands #2	0.78	0.51	-0.16	0.21	3.73	2.99	0.23	0.27
Colorado #1	0.81	0.52	0.09	-0.12	3.12	3.19	0.05	0.03
Ohio	0.76	0.55	0.10	-0.06	2.85	3.03	0.69	0.69
UK	0.66	0.50	-0.33	-0.35	2.94	2.88	0.00	0.00
Aggregated	0.74	0.54	-0.12	-0.21	3.14	3.12	0.00	0.00
			Age Category 11-13					
Netherlands #3	0.86	0.53	-0.02	-0.06	3.26	2.84	0.37	0.39
Colorado #1	0.87	0.55	0.10	0.17	2.74	2.75	0.29	0.09
Minnesota #1	0.76	0.51	0.17	0.05	3.10	3.35	0.01	0.06
UK	0.68	0.44	-0.58	-0.67	3.39	3.52	0.00	0.00
Aggregated	0.73	0.47	-0.31	-0.38	3.20	3.29	0.00	0.00
			Age Category 14-16					
Colorado #1	0.84	0.52	0.29	0.20	3.27	3.25	0.05	0.38
Colorado #2	0.88	0.46	-0.12	-0.11	3.33	2.91	0.05	0.19
Colorado #3	0.77	0.49	0.21	-0.05	3.48	3.73	0.03	0.03
Minnesota #2	0.81	0.42	0.27	0.16	2.39	2.50	0.05	0.48
Australia	0.82	0.45	-0.13	-0.09	2.46	2.46	0.00	0.00
Aggregated	0.83	0.46	-0.02	-0.04	2.94	2.93	0.22	0.17
			Age Category 17-34					
Netherlands #4	0.86	0.34	-0.21	-0.21	2.48	2.27	0.08	0.05
Colorado #3	0.86	0.57	-0.03	-0.03	2.74	2.89	0.44	0.82
Minnesota #2	0.80	0.52	0.53	0.52	3.23	3.07	0.00	0.00
Aggregated	0.82	0.49	0.13	0.13	3.00	3.08	0.16	0.05

Note. 'cor' denotes twin correlation. T1 and T2 denote respectively the twin 1 and twin 2 scores. 'norm. test' denotes the p-value from a Shapiro-Wilks test on normality.

7.3.3 Results

The data are standardized within each age group and within each study. In the aggregated data, we found no evidence for curvilinear effects (i.e., the parameters β_2 and γ_2 were never significant). In addition, we found no AxC interaction in any of the age groups (i.e., the parameter γ_1 was never significant). We therefore limit our results to the model with a linear AxE interaction. See Table 7.4 for the parameter estimates within each age group for each individual study and the aggregated data.

We first investigate whether the Haworth et al. (2010) hold across the 17 subsamples. We do this by plotting heritability estimates within each of the 17 subsamples against the average age of the subsamples. The resulting plot is in Figure 7.1.

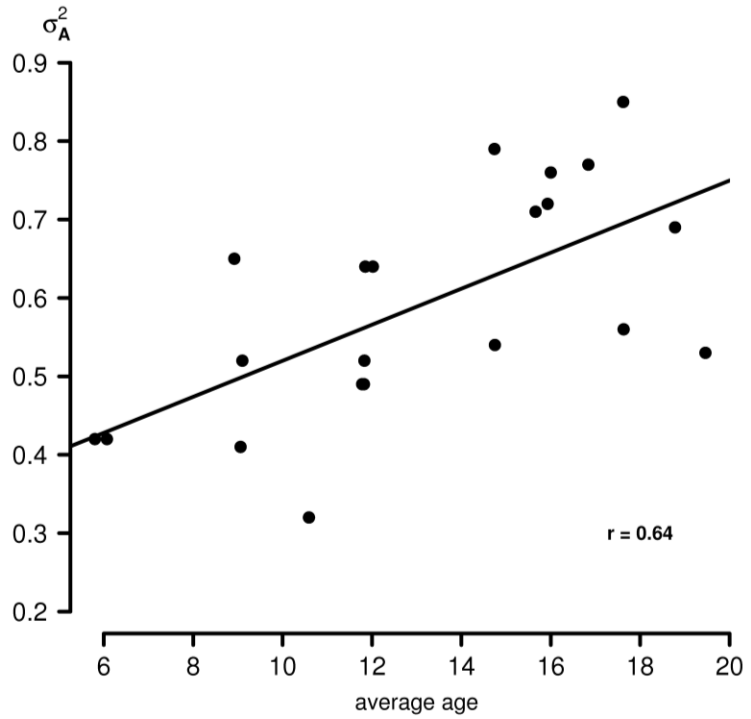

Figure 7.1. Heritability (σ_A^2), as estimated in the 17 samples from the 14 individual studies, plotted against the average age of the sample.

As it appears, the plot confirms the Haworth et al finding that heritability increases approximately linearly with age. However, there appears to be a large amount of heterogeneity in the heritability estimates. We will return to this point later.

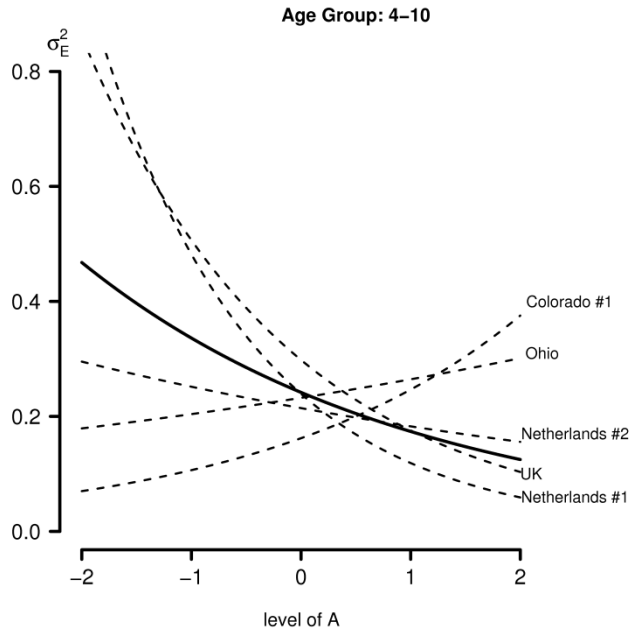

Figure 7.2. Variance of E (σ_E^2) as a function of A in age group 4-10. Striped lines represent the results of the individual studies, the solid bold line represents the results of the aggregated data.

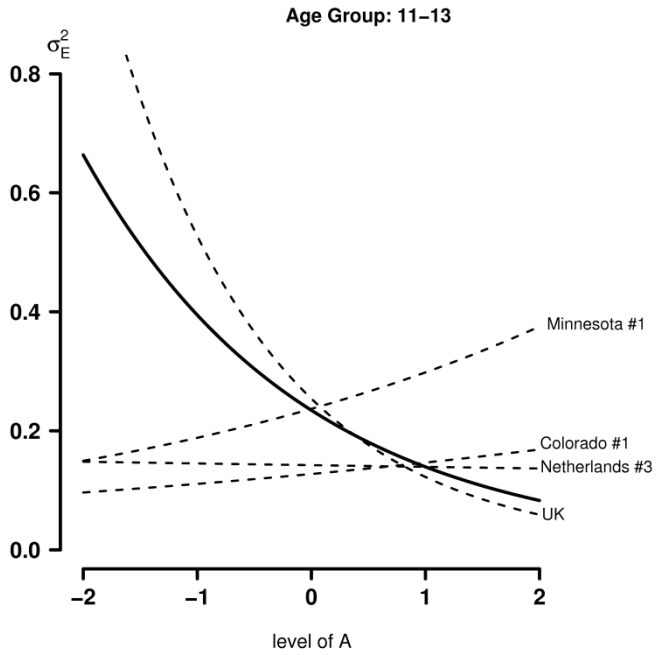

Figure 7.3. Variance of E (σ_E^2) as a function of A in age group 11-13. Striped lines represent the results of the individual studies, the solid bold line represents the results of the aggregated data.

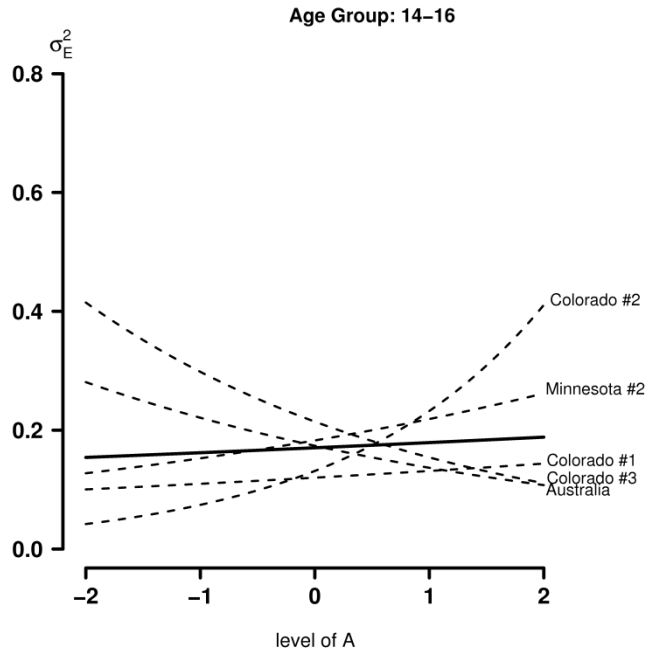

Figure 7.4. Variance of E (σ_E^2) as a function of A in age group 14-16. Striped lines represent the results of the individual studies, the solid bold line represents the results of the aggregated data.

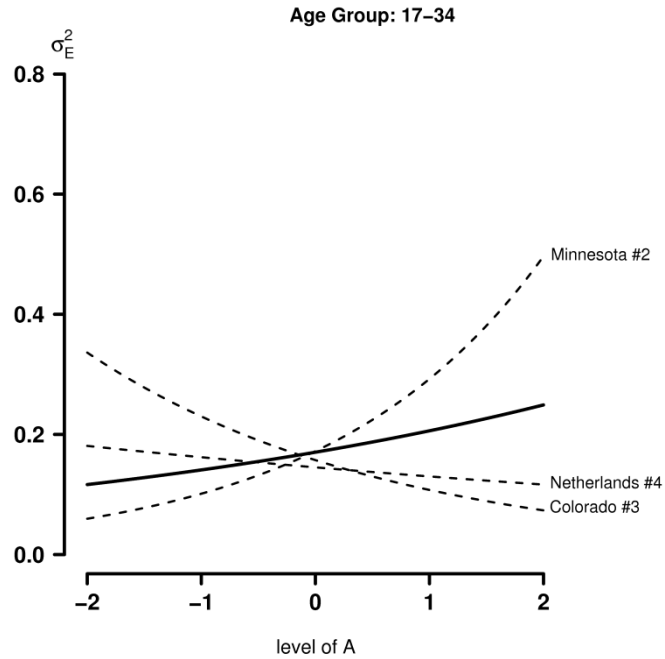

Figure 7.5. Variance of E (σ_E^2) as a function of A in age group 17-34. Striped lines represent the results of the individual studies, the solid bold line represents the results of the aggregated data.

The model implied AxE interaction for each study and for the aggregated data is depicted in Figures 2-5 for ages 4-10, ages 11-13, ages 14-16, and ages 17-34, respectively. From the results of the aggregated data in Figure 7.2 and Figure 7.3, it appears that for age categories 4-10 and 11-13 the variance of E is decreasing across levels of A. The decrease is significant at α=.05 according to both the likelihood ratio test and the 95% confidence interval[37] of β_1, see Table 7.4. Within the individual studies some heterogeneity is apparent, but individual study results tend to follow the trend from the aggregated data. Main departures are Colorado #1 in age group 4-10, and Minnesota #1 in age group 11-13. Both studies show a significant effect in the opposite direction as compared to the effect in the aggregated data, i.e., the variance of E is increasing across A for these studies. Other studies show no effect or an effect in congruence with the aggregated result.

In age category 14-16, no effect is observed in the aggregated data according to the likelihood ratio test and the confidence interval of β_1, see Table3. Within the individual studies some heterogeneity is apparent, but results tend to follow the results from the aggregated data. Some studies show an effect in the opposite direction (i.e., Colorado #1, Ohio, Colorado #2, and Minnesota #1), however, none of these effects are significant at $\alpha = 0.05$, see Figure 7.4 and Table 7.4.

From Figure 7.5 it is apparent that in age category 17-34, the aggregated data show an increase of σ_E^2 across A. This effect is significant according to the likelihood ratio test and the confidence interval of β_1, see Table 7.4. The effect in the aggregated data is replicated in the Minnesota #3 study. However, the Colorado #3 study shows an effect in the opposite direction.

If we compare the results from Table 7.4 to those from Table 7.3, we can conclude that mainly, if normality is rejected (Table 7.3), GxE is present in the data (Table 7.4). However, for some cases this does not hold. For instance, the Colorado #3 study in age group 14-16 is not associated with a GxE interaction but normality is rejected (at $\alpha = 0.05$). The opposite is also observed, e.g., in the Colorado #3 study GxE is present but normality is not rejected.

[37] We considered 95% one-at-a-time confidence intervals as an exploratory tool to see which parameters are significantly different from 0 for α=0.05. We also conducted likelihood ratio tests on β_1 which can be used to provide a more strict test on the significance of this parameter for any α, see Table 6.3.

Table 7.4

Age descriptives, parameter estimates, and number of MZ and DZ twins within age categories, for aggregated data and data within study.

Cat	Study	N_{MZ}	N_{DZ}	Age		σ_A^2	σ_E^2		σ_C^2	$\chi^2(1)$
				mean	*sd*		β_0	β_1		
4-10	Netherlands #1	125	112	5.80	0.06	0.42	-1.43	-0.70	0.26	14.62
						(0.18; 0.68)	(-1.77; -1.13)	(-1.09; -0.34)	(0.03; 0.50)	
	Netherlands #2	49	63	9.10	0.10	0.52	-1.54	-0.16	0.24	0.53
						(0.18; 0.98)	(-1.92; -1.11)	(-0.64; 0.29)	(0.00; 0.60)	
	Colorado #1	391	534	8.92	0.80	0.65	-1.82	0.42	0.19	2.65
						(0.51; 0.84)	(-2.04; -1.59)	(-0.17; 0.68)	(0.05; 0.32)	
	Ohio	121	171	6.07	0.68	0.42	-1.46	0.13	0.34	0.62
						(0.21; 0.65)	(-1.71; -1.20)	(-0.19; 0.46	(0.12; 0.55)	
	UK	394	714	10.59	0.24	0.32	-1.21	-0.53	0.34	42.52
						(0.18; 0.46)	(-1.37; -1.06)	(-0.72; -0.37)	(0.22; 0.45)	
	Aggregated	1140	1616	9.06	1.79	0.41	-1.42	-0.33	0.34	33.98
						(0.34; 0.49)	(-1.52; -1.33)	(-0.45; -0.22)	(0.26; 0.41)	
11-13	Netherlands #3	79	111	12.02	0.08	0.64	-1.95	-0.02	0.21	0.01
						(0.41; 0.87)	(-2.25; -1.62)	(-0.49; 0.51)	(0.00; 0.47)	
	Colorado #1	204	280	11.85	0.81	0.64	-2.06	0.14	0.22	1.01
						(0.50; 0.80)	(-2.26; -1.85)	(-0.13; 0.43)	(0.08; 0.38)	
	Minnesota #1	734	453	11.82	0.40	0.49	-1.44	0.23	0.27	3.82
						(0.36; 0.62)	(-1.55; -1.33)	(-0.001; 0.42)	(0.14; 0.40)	
	UK	1135	1818	11.78	0.40	0.49	-1.37	-0.73	0.16	306.29
						(0.41; 0.57)	(-1.47; -1.28)	(-0.83; -0.64)	(0.09; 0.23)	
	Aggregated	2195	2720	11.83	0.47	0.52	-1.45	-0.52	0.21	203.10
						(0.46; 0.59)	(-1.52; -1.38)	(-0.60; -0.45)	(0.15; 0.27)	

(Table continues)

Table 7.4 *(continued)*

Cat	Study	N_{MZ}	N_{DZ}	Age		σ_A^2	σ_E^2		σ_C^2	$\chi^2(1)$
				mean	*sd*		β_0	β_1		
14-16	Colorado #1	126	173	14.74	0.77	0.79	-2.12	0.09	0.08	0.27
						(0.63; 0.93)	(-2.36; -1.86)	(-0.26; 0.46)	(0.02; 0.26)	
	Colorado #2	153	125	16.00	0.00	0.76	-2.03	0.57	0.10	11.58
						(0.49; 0.96)	(-2.31; -1.76)	(0.25; 0.93)	(0.02; 0.36)	
	Colorado #3	99	140	14.75	0.84	0.54	-1.54	-0.33	0.24	1.89
						(0.28; 0.81)	(-1.97; -1.21)	(-0.95; 0.12)	(0.00; 0.50)	
	Minnesota #2	92	31	16.84	0.10	0.77	-1.70	0.19	0.00	1.13
						(0.62; 0.93)	(-1.95; -1.42)	(-0.16; 0.52)	(0.00; 0.34)	
	Australia	324	483	15.93	0.26	0.72	-1.75	-0.24	0.09	5.12
						(0.59; 0.82)	(-1.91; -1.59)	(-0.45; -0.03)	(0.02; 0.22)	
	Aggregated	807	966	15.66	0.81	0.71	-1.77	-0.05	0.12	0.41
						(0.63; 0.78)	(-1.86; -1.67)	(-0.20; 0.10)	(0.06; 0.20)	
17-34	Netherlands #4	66	94	17.62	0.38	0.85	-1.93	-0.11	0.06	0.24
						(0.62; 1.00)	(-2.28; -1.55)	(-0.53; 0.35)	(0.00; 0.23)	
	Colorado #3	203	180	19.46	2.00	0.53	-1.85	-0.38	0.31	4.94
						(0.36; 0.73)	(-2.11; -1.61)	(-0.74; -0.05)	(0.11; 0.50)	
	Minnesota #2	318	183	17.63	0.36	0.56	-1.76	0.53	0.23	17.21
						(0.38; 0.76)	(-1.97; -1.57)	(0.27; 0.82)	(0.07; 0.43)	
	Aggregated	700	615	18.78	2.71	0.69	-1.77	0.19	0.14	4.39
						(0.57; 0.80)	(-1.88; -1.65)	(0.01; 0.37)	(0.05; 0.26)	

Note. $\chi^2(1)$ denotes a likelihood ratio test on parameter β_1. 95% confidence intervals are in brackets.

*: In a few cases we experienced numerical difficulties in estimating the lower bound of γ_0 because the number of DZ twins was too small. 'Aggregated' is based on the aggregated data within the corresponding age category. As some samples are omitted in the analysis of the individual studies (see Table 6.2), N_{MZ} and N_{DZ} will not necessarily add up to the aggregated N_{MZ} and N_{DZ}.

7.3.4 Conclusion on aggregated data

In the aggregated data, we observed that for subjects between the age of 4 and 13, the influence of the unique environment factor, E, increases for increasing levels of the additive genetic factor, A. For subjects between the age of 14 and 16, no effect was detected. For the subjects between 17 and 34, the influence of factor E increased for increasing levels of A. This pattern of results suggests that the direction of the GxE on IQ reverts during adolescence and young adulthood.

7.3 5 Heterogeneity across studies

The pattern of results as presented in Table 7.4 reveal substantial differences between the individual studies. First, heritability is highly heterogeneous within each age group; For age 4-10, heritability differs between .32 to .65; for age 11-13, heritability differs between .49 to .64; for age 14-16, heritability differs between .54 to .79; and finally, for age 17-34, heritability differs between .53 to .85. Part of this heterogeneity can be explained by the small sample sizes of some of the studies, making the estimate of σ_A^2 highly vulnerable to sampling error. Still, some of the larger studies within the same age group show substantial differences in σ_A^2, e.g., in the age category of 4-10, the UK study (N = 1108) is associated with a heritability of 0.32, while the Colorado #1 study (N = 925) is associated with a heritability of 0.65 which is twice as large. In addition to the heterogeneity with respect to σ_A^2, we see some heterogeneity with respect to the effect of the common environment, σ_C^2. Estimates of σ_C^2 ranged between 0.19-0.34; 0.16-0.27; 0.00-0.24; and 0.06-0.31 within age group 4-10, 11-13, 14-16, and 17-34 respectively. Again, some of these differences can be due to sampling error, however still, large studies show different results, e.g., the UK study is associated with a σ_C^2 estimate of 0.34, while the Colorado #1 study is associated with a σ_C^2 estimate of 0.19. Finally, something similar can be seen with respect to the estimates of the AxE effect, i.e., β_1. Within the age groups, estimates vary widely, with opposite effects within the same age group. Large studies also show substantial differences, e.g., within age group 17-34, the Coloradio #3 study demonstrate a β_1 estimate of -0.38 while the Minnesota #2 study is associated with a β_1 estimate of 0.53, which indicates an effect in the opposite direction.

Given the heterogeneity we found with respect to σ_A^2, σ_C^2, and β_1 - even in the larger studies -, question rises which factors underlie these differences. Below we discuss some of these factors. First, studies differ substantial in the IQ measures that were used. Some studies used full test batteries (Colorado #1, Colorado #2, Netherlands #3, and Netherlands #4); others used only 2 subtests (Colorado #3), 4 subtests (Ohio, UK, Minnesota #1, Minnesota #2), 5 subtests (Australia), or 6 subtests (Netherlands #1

and Netherlands #2). These differences make the IQ measures to reflect different aspects of cognitive ability. For instance, some studies rely only on subtest from the Verbal Comprehension and Perceptual Organization domain (e.g., the Minnesota studies) while in other studies measures are taken from all domains (Working Memory, Perceptual Speed, etc; e.g., as in the Colorado #1 and #2 studies). This difference is important as heritability, unique environmental, and common environmental influences may differ across specific cognitive abilities (e.g., Finkel, Pedersen, McGue, & McClearn, 1995).

Another source of heterogeneity between the individual studies is the possible differences on background variables known to be related to heritability, e.g., parental income, educational attainment, or SES. These differences are likely to occur as samples are possibly biased due to demographical issues, e.g., in the Australia study, twins are mainly from the greater Brisbane area; and in the Minnesota studies, all twins lived reasonably close to the laboratory were testing took place. In addition, there are differences between the individual studies in the way in which subjects were recruited. Some studies relied on birth records (e.g.., Colorado #2) others relied on primary and secondary schools (e.g., Australia) or media advertisements (e.g., Ohio). These differences could also have contributed to heterogeneity with respect to important background variables.

With respect to the estimates of β_1, some interesting differences are observed that are vulnerable to additional explanations besides those above. From Table 7.4, it appears that the β_1 estimates vary with opposite effects across studies within the same age group. Some of the β_1 estimates are non-signficant ($\alpha = 0.05$) and are associated with wide confidence intervals, indicating that sample sizes are too small to draw a reliable conclusion about the β_1 parameter. In age group 4-10, no substantive variation is observed regarding β_1 estimates, as both effects that are significant are in the same direction (i.e., Netherlands #1 and UK). In age group 11-13 only one effect is significant (UK), variation in the remaining studies is mainly due to sampling error. In age group 14-16, a substantive difference is observed: the Colorado #2 study conceals a significant effect with $\beta_1 = 0.57$, while the Australia study is associated with a significant effect in the opposite direction, $\beta_1 = -0.24$. This result indicates that in Colorado, unique environmental effects are stronger for individuals with higher levels of A, while in Australia, unique environmental effects are weaker for individuals with higher levels of A. As both studies have participants of approximately the same age, some other factor should be responsible for this difference. Of course, much dissimilarity exists between the country Australia and the US state Colorado, so it is hard to point to one specific variable. However, some factors that could be thought of are: differences between Australia and Colorado in after school activities (e.g., intellectual or recreational), differences in involvement of the parents in the intellectual

development of the child, and differences in the age until which education is compulsory. With respect to the latter, it is known that in Australia (or more specifically in Brisbane where the Australia study was conducted), education is compulsory until the age of 15-16, while in Colorado, children have to attend school at least until 17. This is an interesting difference as it indicates that for age category 14-16 the Australian sample is characterized by subjects for which education was not compulsory anymore, while this is not true for the Colorado sample. How this may have affected the AxE interaction is open to discussion, however, it would be interesting to analyze additional samples of Australian subjects of different ages to see how AxE is expressed in these data.

In age category 17-34 two effects are significant, Colorado #3 with $\beta_1 = -0.38$, and Minnesota #2 with $\beta_1 = 0.53$, i.e., both effects are in opposite direction. As Colorado and Minnesota are both in the USA, it seems less plausible that differences in e.g., the educational system underlie this observation. However, a noticeable difference between the studies is the age range - as indicated by the standard deviation of the age variable (see Table 7.4) -. The Minnesota #2 study includes participants between the ages of 17-18.5, while the Colorado #3 study includes participants of ages 17-25. To see whether the difference in age range was responsible for the difference between the Minnesota and Colorado results, we analyzed the Colorado #3 data while limiting the age range in this sample to 17-20. Results showed a substantial difference for the AxE interaction as compared to the original results, i.e., β_1 was estimated to be $\beta_1 = 0.20$ as compared to $\beta_1 = -0.38$ in the original analysis. Although the new β_1 estimate was not significant [$\chi^2(1) = 0.80$], the estimate is more in line with the Minnesota results and with the aggregated results. In the new analysis, sample size decreased from N=383 to N=214, but we think that the change in β_1 is still an interesting observation which indicates that the discrepancy between Minnesota #2 and Colorado #3 could be due the large difference in age range.

7.4 Discussion

In the present Chapter, we showed that there is some evidence for AxE interaction in cognitive ability in the GHCA data (Haworth et al, 2010), which comprises 14 studies including subjects of different ages. Results across studies differed substantial with respect to the additive genetic, unique environment, and common environment factors, even within each age group. Taking these differences into account, we replicated the finding of Haworth et al. that heritability is (approximately) linearly increasing across age.

As we noted above, the present undertaking is of interest with respect to the ability differentiation hypothesis which claims that correlations among IQ subtests

decrease as a function of the general intelligence factor (Spearman, 1927). As we argued, from this hypothesis it follows that environmental influences should increase across the genetic factor, as general intelligence is substantially heritable. This genetic explanation of the differentiation hypothesis has – to our knowledge - not been addressed before. Present results only support this predication in the age group 17-34. For ages 4-13, an opposite effect is found. This result suggests that the differentiation effect depends on the age of the subjects. This postulation is supported by Facon (2006), who analyzed subjects between 7 and 15 and found support for ability differentiation only for the subjects of age 13 and older. However, as opposed to the present results, Facon did not find any effect for younger ages. Tucker-Drob (2009) analyzed subjects with ages between 4 and 101, and found a consistent ability differentiation effect in the predicted direction for all ages, which suggests that there is no age moderation. Our results, in which we find an opposite effect for ages below 13, could be due to the smaller heritability of cognitive ability for younger ages (Haworth et al, 2010). For younger ages, association between the general intelligence factor and the additive genetic factor, A, is weaker, and therefore a possible differentiation effect can be masked by effects related to the residual part of factor A (i.e., the part not shared with general intelligence). Facon and Tucker-Drob do not have this problem as these authors do not rely on factor A as a proxy for *g*.

We have limited our analyses to AxE and AxC. We did not consider CxE interaction, because with the present approach, it is very difficult to resolve AxE and CxE (Molenaar et al., 2011). As C is generally accepted to be an appreciable source of phenotypic variance in IQ in children, we cannot rule out that the AxE as detected in young children (Figure 7.2 and 3) is in fact CxE. Given that A and C are likely to be correlated in children (due to passive genotype-environment covariance; Plomin, DeFries, & Loehlin, 1977), one may conservatively interpret the AxE, as observed in children, as Familiarity x Environment interaction.