DeFries-Fulker regression vs. multilevel models for differential heritability and environmentality

Rebecca Pillinger

DeFries-Fulker regression

- Technique for estimating the
  - genetic,
  - shared environmental, and
  - non-shared environmental

  components of the variance between individuals in some trait.
- Uses data on pairs of identical (MZ) and non-identical (DZ) twins
- One twin's score is regressed on the other
- Proportion of shared environmental variation and genetic variation given by coefficients in the regression

$$\mathbf{score}_{1j} = \beta_0 + \beta_1 \mathbf{score}_{2j} + \beta_2 \mathbf{rel}_j + \beta_3 \mathbf{rel.score}_{2j} + e_j$$

DeFries & Fulker (1985)

DeFries-Fulker regression: the details

$$\textbf{score}_{1j} = \beta_0 + \beta_1 \textbf{score}_{2j} + \beta_2 \textbf{rel}_j + \beta_3 \textbf{rel.score}_{2j} + e_j$$

- i indexes twins and j twin pairs
- **rel** is the relatedness of the twins. $\textbf{rel} = \begin{cases} 1 & \text{MZ twins} \\ 0.5 & \text{DZ twins} \end{cases}$
- β_3 gives the proportion of variance in scores due to genetics
- β_1 gives the proportion due to the shared environment
- $1 - \beta_1 - \beta_3$ gives the proportion due to the non-shared environment

Assumptions

- Shared environment the same for MZ and DZ twins
- No assortative mating

DeFries & Fulker (1985)

Differential heritability and environmentality

- The amount of variation due to each component may change over some background variable
- e.g. Turkheimer et. al (2003) find genetic component of variation in IQ increases with SES while environmental components decrease
- DeFries-Fulker regression can be extended to handle this

$$\begin{aligned}\mathbf{score}_{1j} = \beta_0 + \beta_1 \mathbf{score}_{2j} + \beta_2 \mathbf{rel}_j + \beta_3 \mathbf{rel.score}_{2j} \\ + \beta_4 \mathbf{score.SES}_{2j} + \beta_5 \mathbf{rel.score.SES}_{2j} + e_j\end{aligned}$$

- $\beta_3 + \beta_5 \mathbf{SES}$ gives the proportion of variance due to genetics
- $\beta_1 + \beta_4 \mathbf{SES}$ gives the proportion due to shared environment
- $1 - (\beta_1 + \beta_3) - (\beta_4 + \beta_5)\mathbf{SES}$ gives the proportion due to the non-shared environment

LaBuda et al. (1986); Cherny et al. (1992)

Why differential heritability?

Gene-environment interactions

- Some genes have a different effect depending on the environmental context
- Genes for ability may require beneficial environments to affect development
  - Bioecological model, Bronfenbrenner & Ceci (1994)
- Genes may enable good performance despite poor environments
  - Diathesis-stress model, e.g. Gottesman (1991), Paris (1999)

Differential genetic distributions

- The distribution of genes in the population may change with the environmental variable
- e.g. if certain genes enable individuals to reach high SES, those genes will be more prevalent at high SES
- There may be more gene variants at low (or at high) SES

Why differential heritability?

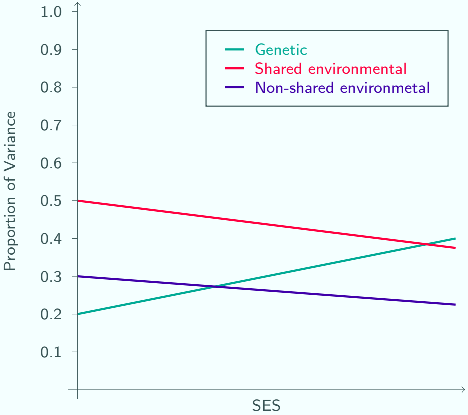
Environment-environment interactions

- Some environmental conditions may have a different effect depending on the environmental context
- They may be able to aid performance only under beneficial environments (combination of beneficial factors needed)
  - Environmental version of the bioecological model
- They may provide resilience against poor environments (combination of adverse factors needed) (Jenkins et al, 2008)
  - Diathesis-stress model

Differential environmental distributions

- The distribution of other environmental factors in the population may change with the environmental variable
- "Tolstoy hypothesis" (my name; also for differential genetic distributions)

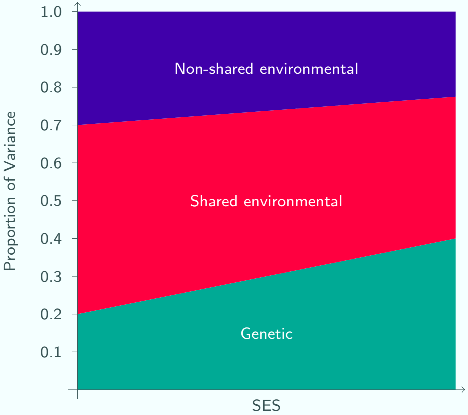
Problem with modelling proportions



Might conclude:

- genetic variation increases: gene-environment interaction
- environmental variation thus forms smaller proportion

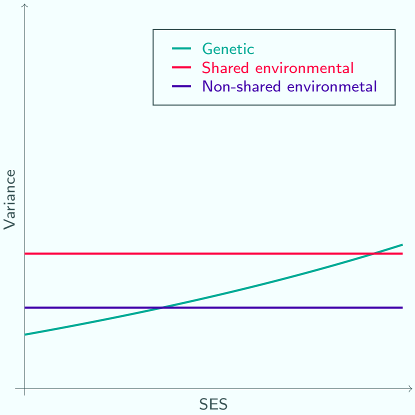
Problem with modelling proportions



Might conclude:

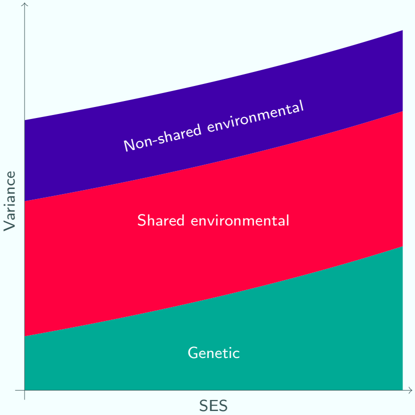
- genetic variation increases: gene-environment interaction
- environmental variation thus forms smaller proportion

Problem with modelling proportions



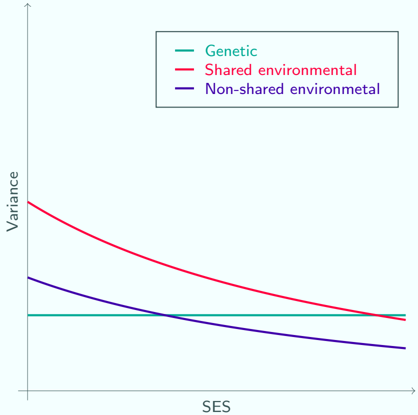
This is certainly one possibility

Problem with modelling proportions



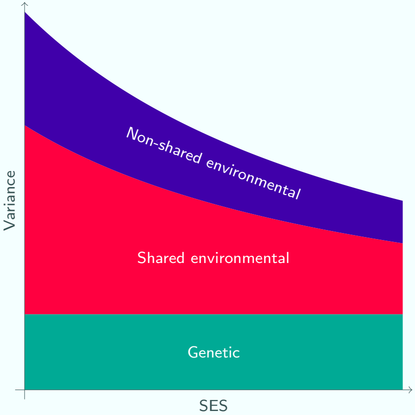
This is certainly one possibility

Problem with modelling proportions



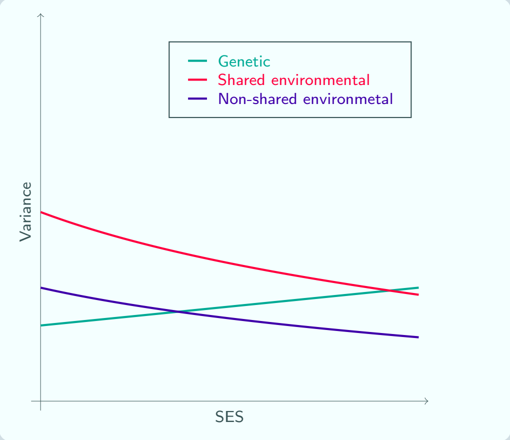
However, it could be the environmental variance decreases and the genetic variance remains constant

Problem with modelling proportions



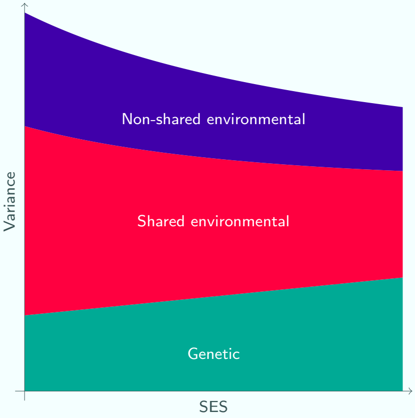
However, it could be the environmental variance decreases and the genetic variance remains constant

Problem with modelling proportions



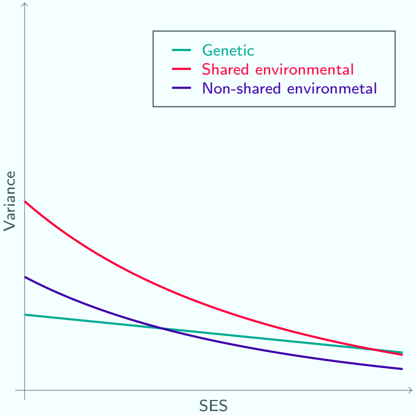
Or the genetic variance could increase and the environmental decrease

Problem with modelling proportions



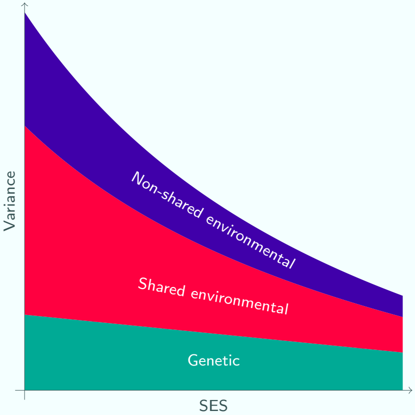
Or the genetic variance could increase and the environmental decrease

Problem with modelling proportions



The genetic variance could even decrease

Problem with modelling proportions



The genetic variance could even decrease

Solution: multilevel modelling!

The multilevel genetic model

$$\textbf{score}_{ij} = \beta_0 + \beta_1\textbf{SES}_j + u_{0j} + u_{1j}\textbf{SES}_j + g_{0ij} + g_{1ij}\textbf{SES}_j + e_{0ij} + e_{1ij}\textbf{SES}_j$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim \mathrm{N}\left(0, \begin{bmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{bmatrix}\right)$$

$$\begin{bmatrix} g_{0ij} \\ g_{1ij} \end{bmatrix} \sim \mathrm{N}\left(0, \begin{bmatrix} \sigma^2_{g0} & \\ \sigma_{g01} & \sigma^2_{g1} \end{bmatrix}\right)$$

$$\begin{bmatrix} e_{0ij} \\ e_{1ij} \end{bmatrix} \sim \mathrm{N}\left(0, \begin{bmatrix} \sigma^2_{e0} & \\ \sigma_{e01} & \sigma^2_{e1} \end{bmatrix}\right)$$

$$\mathrm{cov}(g_{01j}, g_{02j}) = \textbf{rel}\,\sigma^2_{g0}$$

$$\mathrm{cov}(g_{11j}, g_{12j}) = \textbf{rel}\,\sigma^2_{g1}$$

$$\textbf{rel} = \begin{cases} 1 & \text{MZ twins} \\ 0.5 & \text{DZ twins} \end{cases}$$

Genetic variance:
$\sigma^2_{g0} + 2\sigma_{g01}\textbf{SES}_j + \sigma^2_{g1}\textbf{SES}^2_j$

Shared env. variance:
$\sigma^2_{u0} + 2\sigma_{u01}\textbf{SES}_j + \sigma^2_{u1}\textbf{SES}^2_j$

Non-shared env. variance:
$\sigma^2_{e0} + 2\sigma_{e01}\textbf{SES}_j + \sigma^2_{e1}\textbf{SES}^2_j$

Summary

	Genetic variation	Environmental variation
Interaction	Gene-environment interaction	Environment-environment interaction
Tolstoy hyp.	Different distributions of genes	Different distributions of environments

Summary

	Genetic variation	Environmental variation
Interaction	Gene-environment interaction	Environment-environment interaction
Tolstoy hyp.	Different distributions of genes	Different distributions of environments

Summary

	Genetic variation	Environmental variation
Interaction	Gene-environment interaction	Environment-environment interaction
Tolstoy hyp.	Different distributions of genes	Different distributions of environments

Summary

	Genetic variation	Environmental variation
Interaction	Gene-environment interaction	Environment-environment interaction
Tolstoy hyp.	Different distributions of genes	Different distributions of environments

Further work

Simulation

- Simulate data from scenarios similar to these examples
- Fit DeFries-Fulker and multilevel genetic models and compare results

Real data

- Find and get access to real twin data
- Fit DeFries-Fulker and multilevel genetic models and compare results

Structural Equation Modelling

- I don't have much knowledge of SEM
- Turkheimer et al. (2003) appear to have used SEM and modelled the amounts not proportions of variance so looks like SEM could be another solution
- I need to study SEM to find out

References

- Asbury, K., Wachs, T.D. and Plomin, R. (2005) Environmental moderators of genetic influence on verbal and nonverbal abilities in early childhood. *Intelligence* **33**
- Bronfenbrenner, U. and Ceci, S.J. (1994) Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological Review* **101**
- Cherny, S.S., Cardon, L.R., Fulker, D.W. and DeFries, J.C. (1992) Differential heritability across levels of cognitive ability. *Behavior Genetics* **22**:2
- DeFries, J.C. and Fulker, D.W. (1985) Multiple regression analysis of twin data. *Behavior Genetics* **15**:5
- Gottesman, I.I. (1991) *Schizophrenia genesis: The origins of madness* New York: Freeman
- Guo, G. and Wang, J. (2002) The mixed or multilevel model for behavior genetic analysis. *Behavior Genetics* **32**:1 pp 37 - 49
- Jenkins, J., Biscelga, R., Cheung, C., Frampton, K., Gass, K., Gonzales, A., Simpson, A., Valencia, R. and Yu, B. (2008) Psychosocial adversity and resilience. In Rutter, M., Bishop, D., Pine, D., Scott, S.m Stevenson, J., Taylor, E. and Thapar, A. (eds) *Rutter's Child and Adolescent Psychiatry, 5th Edition* Blackwell, Malden, Massachussetts; Oxford; Carlton South, Victoria; Berlin
- LaBuda, M.C., DeFries, J.C. and Fulker, D.W. (1986) Multiple regression analysis of twin data obtained from selected samples. *Genetic Epidemiology* **3**
- Paris, J. (1999) *Genetics and psychopathology: Predisposition-stress interactions*. Washington DC: American Psychiatric Press
- Turkheimer, E., Haley, A., Waldron, M., D'Onofrio, B. and Gottesman, I.I. (2003) Socioeconomic status modifies heritability of IQ in young children *Psychological science* **14**:6