ENVIRONMENTAL-RISK (E-RISK) LONGITUDINAL TWIN STUDY
CONCEPT PAPER FORM

Proposing Author: DW Belsky, C Odgers

Author's affiliation, phone, and e-mail address: Duke University

Sponsoring Investigator (if the proposing author is a student, a post-doc or a colleague): C Odgers

Proposed co-authors: TE Moffitt, A Caspi, L Arseneault, D Corcoran, J Prinz, K Sugden, B Wiliams

Provisional Paper Title: Sociogenomic analysis of neighborhood compositional effects

Date: 1/17/2017

Objective of the study and its significance:

The objective of the study is to test if distributions of genetic risk vary across neighborhoods. Specifically, we are interested in testing evidence for so-called "compositional effects." Compositional effects arise from selection of different types of people into different types of neighborhoods (Leventhal and Brooks-Gunn, 2000; Sampson et al., 2002) and may act as confounders of associations between neighborhood characteristics and residents' outcomes. For example, compositional effects have been explored as potential causes of the higher prevalence of schizophrenia in urban as compared to rural areas (Newbury et al., 2016; Sariaslan et al., 2016).

We aim to test for genetic evidence of compositional effects in domains of social achievement, mental health, and physical health using polygenic score methods combined with state-of-the-art neighborhood ecotyping using a range of data resources available in the E-Risk Study.

Compositional effects are hotly debated in the social sciences. The advent of molecular genetic measurements to quantify inherited risk for various health and social problems opens the opportunity to directly evaluate compositional effects. However, few genetic data are linked with detailed information about place. We propose to bring detailed information about human genomes together with detailed information about the geo-social context in which those genomes reside. Findings will inform a next-generation of neighborhood research.

1

Statistical analyses:

We will quantify genetic risk for low educational attainment, schizophrenia, asthma, and obesity based on data from published genome-wide association studies according to our established method (Okbay et al., 2016; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Moffatt et al., 2010; Locke et al., 2015; Shungin et al., 2015; Belsky et al., 2016).

We will quantify neighborhood disadvantage using the ACORN consumer classification tool and ecological risk evaluated from administrative data, surveys of neighborhood residents, and systematic social observation using Google Streetview (Odgers et al., 2009, 2012).

We will measure social attainment based on the highest education achieved by the E-Risk Study members at the time of the age-18 assessment. We will measure mental health based on the composite 'P'-factor mental health measure. We will measure obesity from anthropometric data collected at the age-18 assessment. We will measure asthma from life-history calendar interviews.

We will test associations among these 3 sets of variables as follows:

1) We will test associations between characteristics of the neighborhood where twins live and the twins' educational and NEET outcomes, asthma history, and obesity.

2) We will test associations between each polygenic score (PGS) and its respective outcome of interest (educational attainment PGS and twins' educational and NEET outcomes; schizophrenia PGS and twins' mental health outcomes; asthma PGS and twins' asthma histories; BMI and waist-hip ratio PGSs and twins' obesity outcomes). We will also test these associations after controlling for family history.

3) We will test associations between characteristics of the neighborhoods where twins live and the twins' polygenic scores for educational attainment, schizophrenia, asthma, BMI, and waist-hip ratio.

4) We will test if any gene-environment correlations observed in (3) reflect passive gene-environment correlation arising from parental selection into neighborhoods on genetic factors. Specifically, we will test if (i) mothers polygenic scores predict neighborhood characteristics; and (ii) mother's polygenic scores account for associations between children's polygenic scores and neighborhood characteristics.

Variables Needed at Which Ages (names and labels):

Study: E-Risk

From the Odgers Lab [no data are needed from London – data are at Duke]
Neighborhood data

Deprivation
sum_weightinc5
sum_weightinc7
sum_weightinc10
sum_weightinc12

Dangerousness
s2safe
SCIIH4, SCIIJ1, SCIIJ2, SCIIJ3
S2victim
SCIIL1, SCIIL2, SCIIL3
S2fear
SCIIF5, SCIIF6, SCIIF7
dprobs
SCIIC14, SCIIC15, SCIIC16, SCIIC17, SCIIC20, SCIIC21, SCIIC22
SSO_UNSAFEneigh

Dilaptitated
SSO_Pdis
SSO_Pdec
s2nprob
(SCIIC1, SCIIC2, SCIIC3, SCIIC4, SCIIC5, SCIIC6, SCIIC7, SCIIC8, SCIIC9, SCIIC10, SCIIC14, SCIIC15, SCIIC16, SCIIC17, SCIIC18, SCIIC19, SCIIC20, SCIIC21, SCIIC22)

Disconnected
s2coef10
(SCIID1, SCIID2, SCIID3, SCIID4, SCIID5, SCIIE1, SCIIE2, SCIIE3, SCIIE4, SCIIE5)

SCIIG1
SCIIG2
SCIIG3
SCIIG4
SCIIG5
SCIIG6
SCIIG7
SCIIG8

From the E-Risk Team [no data are needed from London – data are at Duke]
Age-18 Education
Age-18 BMI, Waist-Hip Ratio
Age-18 P-factor
Asthma History (from age 5, 10 LHCs)

Mother's BMI (from age 12 data)
Mother's family history of psychiatric problems (age 12)
Mother's & Father's education (age 5)
Mother's age at first birth (screener)

From the Duke Biobank Team
Polygenic scores for Educational Attainment, Schizophrenia, BMI, and Waist-Hip Ratio for the twins and their mothers

References cited:

Belsky, D.W., Moffitt, T.E., Corcoran, D.L., Domingue, B., Harrington, H., Hogan, S., Houts, R., Ramrakha, S., Sugden, K., Williams, B.S., et al. (2016). The Genetics of Success: How Single-Nucleotide Polymorphisms Associated With Educational Attainment Relate to Life-Course Development. Psychol. Sci. *27*, 957–972.

Domingue, B.W., and Belsky, D.W. (2017). The social genome: Current findings and implications for the study of human genetics. PLOS Genet. *13*, e1006615.

Leventhal, T., and Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. Psychol. Bull. *126*, 309–337.

Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C., Vedantam, S., Buchkovich, M.L., Yang, J., et al. (2015). Genetic studies of body mass index yield new insights for obesity biology. Nature *518*, 197–206.

Moffatt, M.F., Gut, I.G., Demenais, F., Strachan, D.P., Bouzigon, E., Heath, S., von Mutius, E., Farrall, M., Lathrop, M., and Cookson, W.O.C.M. (2010). A large-scale, consortium-based genomewide association study of asthma. N. Engl. J. Med. *363*, 1211–21.

Newbury, J., Arseneault, L., Caspi, A., Moffitt, T.E., Odgers, C.L., and Fisher, H.L. (2016). Why Are Children in Urban Neighborhoods at Increased Risk for Psychotic Symptoms? Findings From a UK Longitudinal Cohort Study. Schizophr. Bull. *42*, 1372–1383.

Odgers, C.L., Moffitt, T.E., Tach, L.M., Sampson, R.J., Taylor, A., Matthews, C.L., and Caspi, A. (2009). The protective effects of neighborhood collective efficacy on British children growing up in deprivation: A developmental analysis. Dev. Psychol. *45*, 942–957.

Odgers, C.L., Caspi, A., Bates, C.J., Sampson, R.J., and Moffitt, T.E. (2012). Systematic social observation of children's neighborhoods using Google Street View: a reliable and cost-effective method. J. Child Psychol. Psychiatry *53*, 1009–17.

Okbay, A., Beauchamp, J.P., Fontana, M.A., Lee, J.J., Pers, T.H., Rietveld, C.A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S.F.W., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. Nature *533*, 539–542.

Sampson, R.J., Morenoff, J.D., and Gannon-Rowley, T. (2002). Assessing "Neighborhood Effects": Social Processes and New Directions in Research. Annu. Rev. Sociol. *28*, 443–478.

Sariaslan, A., Fazel, S., D'Onofrio, B.M., Långström, N., Larsson, H., Bergen, S.E., Sariaslan, A., D'Onofrio, B.M., Fazel, S., Lichtenstein, P., et al. (2016). Schizophrenia and subsequent neighborhood deprivation: revisiting the social drift hypothesis using population, twin and molecular genetic data. Transl. Psychiatry *6*, e796.

Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *advance online publication*.

Shungin, D., Winkler, T.W., Croteau-Chonka, D.C., Ferreira, T., Locke, A.E., Mägi, R., Strawbridge, R.J., Pers, T.H., Fischer, K., Justice, A.E., et al. (2015). New genetic loci link adipose and insulin biology to body fat distribution. Nature *518*, 187–196.

Data Security Agreement

Provisional Paper Title	Sociogenomic analysis of neighborhood compositional effects
Proposing Author	DW Belsky, C Odgers
Today's Date	1/17/2017

Please keep one copy for your records
(Please initial your agreement)

__DWB__   I am current on Human Subjects Training (CITI (www.citiprogram.org) or training in human subject protection through my post or courses.

__DWB__   My project is covered by Duke or King's IRB OR I have /will obtain IRB approval from my home institution.

__DWB__   I will treat all data as "restricted" and store in a secure fashion.

_DWB___   I will not share the data with anyone, including students or other collaborators not specifically listed on this concept paper.

_DWB___   I will not post data online or submit the data file to a journal for them to post. Some journals are now requesting the data file as part of the manuscript submission process. The E-Risk Study cannot be shared because the Study Members have not given informed consent for unrestricted open access. Speak to Terrie or Avshalom for strategies for dealing with data sharing requests from Journals.

__DWB__   Before submitting my paper to a journal, I will submit my draft manuscript and scripts for data checking, and my draft manuscript for co-author mock review, allowing three weeks.

__DWB__   I will submit analysis scripts and new variable documentation to project data manager after manuscript gets accepted for publication.

_DWB___   I will return all data files to the Data Manager after the project is complete. Collaborators and graduates of DPPP may not take a data file away from the DPPP office. The data remains the property of the Study and cannot be used for further analyses without express, written permission.

_DWB___   I will ensure geographical location information, including postcodes or geographical coordinates for the E-Risk study member's homes or schools, is <u>never</u> combined or stored with any other E-Risk data (family or twin-level data)

Signature: Daniel W. Belsky

CONCEPT PAPER RESPONSE FORM

A. To be completed by the proposing author

Proposing Author:

X I have read the E-Risk data-sharing policy guidelines and agree to follow them

Provisional Paper Title: Sociogenomic analysis of neighborhood compositional effects

Potential co-authors: TE Moffitt, A Caspi, L Arseneault, D Corcoran, J Prinz, K Sugden, B Wiliams

Intended Submission Date (month/year):   6/1/2017

Please keep one copy for your records and return one to Louise (louise.arseneault@kcl.ac.uk)

B.   To be completed by potential co-authors:

☐   Approved   ☐   Not Approved   ☐   Let's discuss, I have concerns

Comments:

Please check your contribution(s) for authorship:

☐   Conceptualizing and designing the longitudinal study

☐   Conceptualizing and collecting one or more variables

☐   Data collection

☐   Conceptualizing and designing this specific paper project

☐   Statistical analyses

☐   Writing

☐   Reviewing manuscript drafts

☐   Final approval before submission for publication

☐   Acknowledgment only, I will not be a co-author

Signature: ....