Discovering the genetic architecture of the mind

Karlsson Linnér, R.

2019

document version
Publisher's PDF, also known as Version of record

Meta-analysis of the serotonin transporter promoter variant (*5-HTTLPR*) in relation to adverse environment and antisocial behavior

I don't think aggression works like thirst or sleep. I think aggression is more elicited by particular situations. I think it can be mitigated.

Steven Pinker

It's been proven by quite a few studies that plants are good for our psychological development. If you green an area, the rate of crime goes down. Torture victims begin to recover when they spend time outside in a garden with flowers. So we need them, in some deep psychological sense, which I don't suppose anybody really understands yet.

Jane Goodall

5

Based on Tielbeek, Karlsson Linnér, Beers, Posthuma, Popma, & Polderman (2016). *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics.*

Abstract

Several studies have suggested an association between antisocial, aggressive and delinquent behavior and the short variant of the serotonin transporter gene polymorphism (*5-HTTLPR).* Yet, genome-wide and candidate-gene studies in humans have not convincingly shown an association between these behaviors and *5-HTTLPR*. Moreover, individual studies examining the effect of *5-HTTLPR* in the presence or absence of adverse environmental factors revealed inconsistent results. We therefore performed a meta-analysis to test for the robustness of the potential interaction effect of the "long-short" variant of the *5-HTTLPR* genotype and environmental adversities, on antisocial behavior. Eight studies, comprising of 12 reasonably independent samples, totaling 7,680 subjects with an effective sample size of 6,724, were included in the meta-analysis. Although our extensive meta-analysis resulted in a significant interaction effect between the *5-HTTLPR* genotype and environmental adversities on antisocial behavior, the methodological constraints of the included studies hampered a confident interpretation of our results, and firm conclusions regarding the direction of effect. Future studies that aim to examine biosocial mechanisms that influence the etiology of antisocial behavior should make use of larger samples, extend to genome-wide genetic risk scores and properly control for covariate interaction terms, ensuring valid and well-powered research designs.

Introduction

Antisocial behaviors, such as aggression, have a destructive effect on the lives of victims, can disrupt communities and cause societal instability and fear. Next to these physical and psychological costs, antisocial behavior inevitably brings enormous monetary costs. The average costs per murder in the United States was estimated between $17.25 and $24 million and the costs of criminal careers of the most violent and prolific offenders was estimated to be $150–160 million[1,2]. Given its tremendous societal impact, the study of antisocial behavior has been an important focus of research throughout history. Scientifically informed early intervention strategies could potentially interfere with risk factors for antisocial development, thereby effectively reducing the burden of antisocial behavior for society. However, despite substantial research efforts, the complex etiology of antisocial behavior remains only partly understood.

Environmental Adversities

Criminological research has revealed extensive and invaluable information regarding the environmental adversities influencing antisocial behavior. An important environmental adversity that often arises from these studies is the negative effect of exposure to stressful or traumatic events early in life. These include physical, emotional or sexual abuse, neglect, and a range of household dysfunctions, such as witnessing violence against mothers or incarceration of a family member. Exposure to environmental adversities can have an immense impact on the cognitive and emotional development of children and is associated with developing serious mental disorders later in life and health problems throughout the lifespan, such as alcohol-seeking behavior in early and mid-adolescence[3], early initiation of illicit drugs and sexual risk behaviors[4]. Environmental adversities also impose an increased risk of developing antisocial outcomes, such as conduct disorder and adult criminality[5,6]. In addition to childhood environmental adversities, other environmental variables such as a lower socioeconomic status (SES) and delinquent peer affiliation are also associated with antisocial development. A recent meta-analysis, summarizing the data of 133 studies, demonstrated that lower family SES was associated with higher levels of antisocial behavior in children and adolescents[7]. Another study showed that socialization effects of peer influences on antisocial behavior were particularly important at the age of 16 to 20 years, after which the impact of peers seems to disappear[8]. This work illustrates the importance of taking developmental periods into account when studying the relationship between environmental adversities and antisocial behavior. Even though environmental adversities have negative overall effects, the degree to which they affect individual development differs.

5-HTTLPR and antisocial behavior

In addition to the influence of environmental risk factors, twin and adoption studies have shown an important contribution of genetic factors to antisocial behavior with about 50% of the individual differences in antisocial behavior being explained by genetic variation[9–11]. Together with the *MAOA* and *COMT* gene, the serotonin transporter gene (*SERT* or *5-HTTLPR*) is one of the most widely studied genes in relation to the development of antisocial behavior. *5-HTTLPR* is a polymorphism of the serotonin transporter gene that has a long (L) or a short (S) variant. The S variant has been found to affect the efficiency of the transcription rate of the gene, thereby regulating the serotonergic availability in the brain[12]. Previous studies in monkeys have shown that S-carriage of *5-HTTLPR* leads to an altered neural stress and thread circuitry as well as an increased hypothalamic–pituitary–adrenal (HPA) axis response to stress[13,14]. A study in mice demonstrated increased vulnerability to psychosocial stress in heterozygous *5-*

HTT knockout mice[15]. Also human studies have reported the involvement of *5-HTTLPR* in mood and emotion regulation[16]. Despite the mixed findings in (generally) small samples, *5-HTTLPR* has been suggested as an important candidate gene influencing the development of antisocial behavior. Recently however, Vassos et al. (2014)[17] demonstrated the limitations of candidate-gene studies with a systematic review of all published genetic association studies of aggression and violence. Their meta-analysis did not find significant associations between any candidate gene and aggression outcomes[17]. Regarding *5-HTTLPR*, 19 studies were meta-analyzed, yielding a non-significant odds ratio of 0.90. Nevertheless, their meta-analysis did not include gene-environment interactions, so it might be that moderation effects are missed out since the interaction effects of these genes with the environment were not modeled.

Gene-environment interaction (G×E) studies

Animal studies have reported individual differences in behavioral and biological responses to environmental challenges. Suomi and others[14] examined individual differences in resilience of rhesus monkeys that were reared in different environments during infancy: with peers only (PO) and with both mother and peers (MP). Their study revealed that rhesus monkeys carrying the S variant of *5-HTTLPR* only showed excessive aggression and other neurodevelopmental impairments when they had been PO-reared, but not when they were MP-reared. Similarly, another study in monkeys demonstrated that the S allele was associated with increased reactivity to repeated, chronic stress, inclining a higher risk of developing affective psychopathology[18]. In humans, a twin study found that the effect of maltreatment on conduct problems was more severe in children at high genetic risk (i.e., having a co-twin with conduct disorder), increasing the probability of conduct disorder with 24% compared to 2% at low genetic risk[19]. Similarly, a twin study on antisocial behavior found higher estimates of heritability in adolescents in socioeconomically more advantaged environments, compared to adolescents in socioeconomically less advantaged environments, indicating a gene-environment interaction[20].

Scholars have suggested two distinct, but not mutually exclusive, theoretical frameworks when interpreting these G×E effects. The diathesis-stress model proposes that the impact of negative life experiences on psychopathology depends on the individual's pre-dispositional vulnerability or diathesis[21–23]. An alternative explanation branching from the diathesis-stress model is the differential susceptibility theory (DST) which advocates that, along with variation in vulnerability to negative experiences, individuals also differ in their susceptibility to positive environmental conditions[24].

A landmark study by Caspi et al. (2003)[25] reported that individuals possessing two S alleles of *5-HTTLPR*, were most adversely affected by stressful life events, but in the absence of these events, they showed the lowest scores on depression and suicidality. Even though these findings fit within the DST framework, individual studies examining this interaction have shown inconsistent effects. Moreover, meta-analyses have been inconclusive with two studies[26,27] failing to support the original interaction, and one more inclusive meta-analysis finding support for an interaction effect[28]. The latter study however, yielded liberal inclusion criteria incorporating many indirect replications resulting in a higher risk of publication bias[29].

Several studies in humans have examined the relationship between *5-HTTLPR*, environmental adversities and their effect on "antisocial behaviors". Again, the findings of these candidate gene-environment interaction studies have generally been inconclusive and are typically characterized by underpowered samples[29–31]. Hence, the aim of the current study is to meta-analyze all reported studies in humans that investigated a *5-HTTLPR*×environmental adversities interaction effect on antisocial behavior, to test for an overall effect. We defined

environmental adversity as exposure to psychosocial risk factors associated with antisocial conduct[32], such as childhood maltreatment, lower family SES and delinquent peer affiliation. Antisocial behavior was broadly defined and encompassed both reactive aggression as well as instrumental forms of aggression, such as psychopathy. To assess the methodological quality of the studies included in our meta-analysis, we performed a systematic review to guide the interpretation of results in terms of potential for biases and sources of heterogeneity.

Materials and Methods

Literature search and selection strategy

A systematic literature search was performed to collect relevant studies. Two independent researchers (JT, KB) conducted the search using four online databases (Google Scholar, PubMed, Web of Science, and Psychinfo). We performed a search on the gene-environment interaction between the serotonin transporter gene ("*5-HTTLPR"*, "*5-HTT*", "serotonin transporter gene", SERT) and environmental adversity ("abuse", "adverse childhood events" (or "ACE"), "maltreatment", "SES", "socioeconomic status", "environmental adversity", "parenting quality", "peer behavior" and "peer network), on antisocial outcome measures ("antisocial", "aggression", "conduct", "criminal", "delinquency", "crime"). Studies were first evaluated on their relevance based on title and abstract. Secondly, we checked for the presence of an antisocial outcome measure, environmental adversity measure, and *5-HTTLPR* measure (for criteria, see **Textbox I**). In addition, we searched the reference lists of relevant papers to find additional studies. **Figure 5.1** reflects the flowchart of our search. Of the 17 potentially suitable studies, nine did not meet our inclusion criteria and were excluded after full text screening. No additional papers, based on reference lists, were identified. Thus, the search resulted in the inclusion of eight suitable studies.

Inclusion

- Antisocial outcome measure
- Environmental adversity measure
- *5-HTTLPR* measure
- Association test of interaction between 'short-long' *5-HTTLPR* variant and environmental adversity on antisocial outcome measure

Exclusion

- Subjects with severe handicaps
- Medical intervention studies
- Focus on comorbidity with a particular disorder (e.g., ADHD, anxiety)

Textbox I. Inclusion and exclusion criteria for meta-analysis

Quality assessment

To evaluate the methodological quality of the included studies we used a short version of the quality assessment as proposed by Hayden et al. (2006)[33] including the items as shown in **Textbox II**. Every item was rated *positive* (+), *average* (+ -), or *negative* (-), by two independent researchers (JJT, KB). In case of disagreement, an additional third researcher (TJCP) rated that particular study. Studies were ranked based on the total sum score of all items, with "+" being one point, "+-" being a half point, and "-" being zero points. In line with previous studies, we did not weigh studies in the meta-analysis by their quality scores. However, to test whether low scoring studies affected our results, we repeated the overall meta-analysis utilizing a descending and ascending step-wise removal of the studies based on their quality score[34].

Study Participation
(1) Clear description of the key characteristics of the study population (distribution by age, gender and ethnicity).
(2) The sampling frame and recruitment are described, including characteristics of the place of recruitment and response rate.
(3) Inclusion and exclusion criteria are described.

Measures
(4) A clear definition of the measure of antisocial behavior and environmental adversity is provided.
(5) The measure of antisocial behavior and environmental adversity is adequately valid to limit misclassification (Includes when applicable: report of the reliability of the assessment).

Covariates
(6) Age and gender are accounted for in the analysis.
(7) IQ or SES or Race are accounted for in the analysis.
(8) Covariate interaction terms are accounted for in the analysis.

Textbox II. Items of Quality Assessment, as derived from Hayden et al. (2006)

Descriptives

Data from the eight studies that met criteria for inclusion were extracted and summarized in **Table 5.1**, based on the design of Munafo et al. (2009)[26]. Briefly, we provide the characteristics of all samples, as well as study designs, measurements, interaction and statistical models.

Meta-analysis method

The literature search for this study resulted in eight eligible studies, comprising of 12 reasonably independent samples. The low number of studies found in the literature search motivated the combination of the non-independent samples; and this concerns seven of the included samples from three of the studies. In these cases, the extracted analyses were non-independent analyses divided per-sex or per-ethnicity from the same population. The present meta-analysis includes studies with a variety of designs (e.g. some being case-control studies, others being extreme-group sampling), statistical models (e.g. count data, logistic regression and factor analysis) and subtypes of analyses (e.g. within-sex and pooled-sex analysis). In addition, the degree of scaling regarding both the phenotype and the genotype varies across studies. These different sources of between-study variability raise an important issue regarding the assumption of heterogeneity of the effect sizes available for meta-analysis[35]. Study heterogeneity introduces difficulties of drawing a directional conclusion based on the overall meta-analysis result and this is further discussed in the **Supplementary Material**. Furthermore, only five out of the 12 samples included sufficient information on the effect size and within-study error variance for the parameter estimate to be converted into a common effect-size metric and perform a random-effects meta-analysis[36]. Based on these issues we performed a meta-analysis on the level of significance, i.e. the *P* values, rather than the effect sizes. The motivation for this approach and potential issues when trying to draw a directional conclusion are further discussed in the **Supplementary Material**.

P value extraction

All of the included samples employed two-tailed, and hence non-directional tests of parameter significance; this is displayed in **Table 5.3**. The **Supplementary Material** explains in depth how we extracted and converted *P* values from the different non-directional tests of parameter significance into harmonized *Z* statistics. In summary, relevant *P* values were first independently extracted from the samples by three investigators (JT, KB & RKL), and any inconsistencies were settled by discussion. Secondly, the estimated direction of effect was harmonized over the studies to make sure that the variable coding of the effect allele, the coding of the environmental variable and the coding of the outcome measure had the same direction of effect in all studies. This is explained in depth in the **Supplementary Material**, and the result of this harmonization is shown in **Table 5.3**, column "Harmonized direction of G×E" and column "Harmonized *Z* statistic". Further, the meta-analysis was performed following the often employed methodology of combining significance levels referred to here as the non-weighted Stouffer's method and the weighted Liptak-Stouffer's method described in e.g., refs. [37–39]. In summary, all extracted two-tailed *P* values were harmonized into *Z* statistics that all have the same direction of effect in terms of the effect allele, the environmental adversity, and the outcome measure of antisocial behavior.

Stouffer's method and the weighted Liptak-Stouffer's method

The meta-analysis was first performed on the *Z*-transformed *P* values employing the non-weighted Stouffer's method implemented in the following way

$$Z_{Stouffer} = \frac{\sum_{i=1}^{k} Z_i}{\sqrt{k}}$$

where k is the number of included *Z* statistics. This combined metric tests the null hypothesis of no effect in all samples and each sample is assigned the same weight (Borenstein, 2009, p. 328; Becker, 1994, p. 218-219)[35,40]. Positive and negative *z*-statistics cancel out, so that the combined metric can be seen as a reasonable approximation of an average effect size under the assumption that all transformed *P* values are derived from the same parameter estimate, as discussed in the **Supplementary Material**. A clear caveat of the method is that no difference in weight is given to the studies based on their measure of precision. At the same time, the *P* value for a given effect size is a function of sample size, and hence the metric does incorporate information on the precision of a study (Becker, 1994, p. 227)[35].

Even though the non-weighted Stouffer's method includes information of per-sample precision via the *P* value, we also performed a meta-analysis with the weighted Liptak-Stouffer's method. This was carried out on the *Z*-transformed *P* values in the following way

$$Z_{Liptak-Stouffer} = \frac{\sum_{i=1}^{k} w_i Z_i}{\sqrt{\sum_{i=1}^{k} w_i^2}}$$

where

$$w_i = \sqrt{N_{eff,i}}$$

where $N_{eff,i}$ is the effective sample-size of study i which is the full sample size except in the case of a sample being of case-control design. In these cases, the effective sample-size was calculated according to Willer et al. (2010)[39] and rounded to closest integer.

$$N_{eff,i} = \frac{4}{\left(\frac{1}{N_{cases}} + \frac{1}{N_{controls}}\right)}$$

If the effective sample-size would be the same in all samples, then $Z_{Liptak-Stouffer}$ converges into $Z_{Stouffer}$ (Whitlock, 2005, p. 1369)[38]. We chose not to perform the meta-analysis with the Fisher's combined probability test, often referred to as "the sum of logs method", since there is a bias towards low compared to high *P* values with this method, as explained in Whitlock (2005)[38]. Two-tailed *P* values of the overall *Z* statistic from the non-weighted Stouffer's method and the weighted Liptak-Stouffer's method were calculated as $P = 2\Phi(-|Z|)$, as in Stock and Watson (2012)[41].

Environmental adversity and sex specific meta-analysis

To perform stratified analyses, we divided the samples into two subsets: adverse childhood environment ($n = 8$) as well as SES and peer rejection ($n = 4$). Furthermore, we conducted both pooled-sex and within-sex analyses. Since all stratified analyses include a low number of samples, i.e. ≤ 8, possible outliers might drive these subset analyses.

Robustness checks

As a robustness check we tested the impact of each study on the overall significance by removing them individually from the analysis. Additionally, we analyzed two subsets where only studies with an effective sample-size above the median or the average were included. All robustness checks were performed using the $Z_{Liptak-Stouffer}$ meta-analytic procedure, in order to control for extreme values from small samples given the low number of available studies.

There is a possibility that the overall meta-analysis result is driven by false-positive findings in samples with small sample size, this is referred to as the small-study effect (Borenstein et al., 2009, p. 291)[42]. We therefore reran the meta-analysis while excluding samples with an effective sample-size lower than the average effective sample-size ($average\ N_{eff} = 560$) or lower than the median effective sample-size ($median\ N_{eff} = 389$) and observed the effect on the overall *P* value. Moreover, through step-wise removal of studies with the smallest sample size, we tested the number of studies that could be removed before the overall *P* value would change to non-significant (i.e., two-tailed *P* value > 0.05 for the weighted Liptak-Stouffer method). We also performed both descending and ascending step-wise removal of the studies according to their rated quality score, as explained in section **Quality Assessment**.

P-curve

As an additional robustness check we utilized the meta-analytical procedure, the *P*-curve as recently introduced by Simonsohn, Nelson, & Simmons (2014)[43] (see **Supplementary Material** for a detailed description of this method). We employed the *P*-curve in comparison with the non-weighted Stouffer's method and the weighted Liptak-Stouffer's since these methods rely on the availability of null results in the published body of literature. However, if null results are missing from the published body of literature because of publication bias then the results of the Stouffer's and Liptak-Stouffer's methods are biased towards showing a larger average effect than if the non-published null results would have been included[44]. We utilized two *P*-curves, the first testing the hypothesis of a positive interaction effect between the S variant and environmental adversity on antisocial behavior by including the subset of the two-tailed statistically *P* values from significant estimates with a positive interaction effect, which were then divided by two. The second *P*-curve tests the two-sided hypothesis of an interaction

effect that can take both positive and negative values in the terms of the S variant and this was performed on all of the statistically significant two-tailed *P* values in our sample.

Since the publication of this study, the reliability of the *P*-curve for analyses of non-experimental data has been criticized (Bruns & Ioannidis, 2016)[67].

Rosenthal's failsafe N

The potential for publication bias was further assessed through calculation of a failsafe number (N_{fs}) using the formula $N_{fs} = \frac{N_0}{z_c^2}(N_0 Z_0^2 - Z_c^2)$, where N_0 is the total number of samples in the meta-analysis, Z_c is the two-tailed critical value of Z with a type I error rate, $\alpha = 0.05$, and Z_0 is either the $Z_{Stouffer}$ or the $Z_{Liptak-Stouffer}$ obtained for the total sample[45]. The fail-safe N, first described by[46], is the number of additional "negative" statistical analyses (yielding an average *Z*-statistic of zero) that would need to be added to our meta-analysis to make the combined effect size statistically non-significant. Noteworthy, this method of publication-bias testing is heavily criticized in the recent publication bias literature regarding issues of interpretability under heterogeneity of included samples, the lack of underlying statistical model and therefore varying results depending on which version of the test is employed. For further discussion, we refer the interested reader to Becker (2005)[47] and Borenstein et al., (2009, p. 284-285)[42]. Based on their discussion, we interpret the results from this publication-bias testing moderately. Nonetheless, we chose to test for publication bias with this method because most other common publication-bias testing procedures require effect sizes and are thus beyond the scope of this paper (Rothstein et al., 2005)[48] because most of the included studies do not report these metrics.

Results

Quality assessment

Table 5.2 displays the methodological quality of the studies included in our meta-analysis, listed in descending order. Importantly, none of the studies attained the highest possible quality score and they all failed to control for the mandatory covariate×gene and covariate×environment interaction terms that are required for unambiguous interpretation of the estimated interaction effect[49].

Overall meta-analysis

The meta-analysis was performed over 12 samples, most of which were independent, adding up to an effective sample-size of 6,724 subjects (**Table 5.3**). Analyses were conducted using the non-weighted Stouffer as well as the weighted Liptak-Stouffer method, however we generally restrict the report to the weighted Liptak-Stouffer method unless there is discordance between the methods. We found support for an interaction effect of the *5-HTTLPR* genotype and environmental adversities on antisocial outcomes, with a *P* value of the non-weighted Stouffer method equal to 0.0029 and the weighted Liptak-Stouffer *P* value of 0.0006 (**Figure 5.2** and **Table 5.4**).

In the **Supplementary Material** we discuss the restrictions regarding directional inference based on our meta-analytical approach. Since directional conclusions are to be avoided, our results point toward a significant, non-zero interaction effect in at least one of the included samples. Thus, the meta-analysis supports the alternative hypothesis that there is an interaction effect between the *5-HTTPLR* gene region and environmental adversities on antisocial

behavior, however the interpretation concerning which allele does actually increases sensitivity remains ambiguous. This is further illustrated by the analysis in ref. [50] which estimated a u-shaped relationship with regards to the homozygotes.

Environmental adversity and sex specific meta-analysis

The weighted Liptak-Stouffer *P* value became less significant when focusing exclusively on the subset of eight samples that used adverse childhood environment as more specific parameter (*P* value = 0.020), and the choice of meta-analytic procedure gave noticeably different results when compared to the non-weighted Stouffer *P* value (*p*-value = 0.0017). However, both methods still resulted in statistical significance and the number of samples in the subset was relatively low. Outliers with below-average sample size such as the female African American sample of Douglas et al., 2011[51] (Z statistic = 5.00, n = 112), could therefore explain the difference in results between the methods. When focusing on the samples with SES and peer affiliation as the environmental adversity, the two-tailed Liptak-Stouffer *P* value became non-significant (*P* value = 0.055) and the two-tailed non-weighted Stouffer *P* value became very non-significant (*P* value = 0.47). In this case the weights given to the two large samples with positive effect, i.e. Aslund et al. (2013)[52], and Kretschmer et al. (2014)[53] clearly affects the results towards statistical significance, even though both methods gave non-significant results.

In the meta-analyses stratified by sex, we analyzed the combined male-female samples and the exclusive male and female samples separately. The pooled sex meta-analysis included six samples and the Liptak-Stouffer *P* value was significant (*P* value = 0.0085), however the non-weighted Stouffer *P* value was non-significant (*P* value = 0.063). The sex-specific analyses both comprised of three samples and the analysis revealed a significant overall effect for females (Liptak-Stouffer *P* value = 0.032, Stouffer *P* value = 0.0019) but a non-significant interaction effect for males (Liptak-Stouffer *P* value = 0.27, Stouffer *P* value = 0.45).

Robustness checks

The robustness check showed that the results were robust against the removal of each sample individually, as the meta-analytic two-tailed Liptak-Stouffer *P* value did not change to being non-significant (i.e. *P* value > 0.05). The result of this robustness check is displayed in Table IV. However, when removing Aslund et al. (2013)[52], then the two-tailed Liptak-Stouffer *P* value changed to be marginally significant, i.e. in the range of 0.01–0.05. This shows that the overall result is sensitive to the exclusion of either large samples or samples with large effects such as Aslund et al. (2013) ($Z_i = 3.20, N_{eff} = 752$).

After exclusion of the seven samples with lower than average effective sample-size the overall Liptak-Stouffer *P* value remained significant (*P* value < 0.0001). Likewise, after exclusion of the six samples with lower than median effective sample-size the overall Liptak-Stouffer *P* value remained significant (*P* value ≈ 0.0001). Similarly, a step-wise removal of the studies with the smallest sample size revealed that the 10 smallest of the 12 samples could be removed before the overall two-tailed Liptak-Stouffer *P* value would change to non-significance (*P* value > 0.05).

Step-wise removal of the samples based on the quality assessment score showed that the results might be driven by the studies with the lowest quality. When we first removed the samples with lowest quality the overall weighted Liptak-Stouffer *P* value remained significant until only the two highest rated studies remained, i.e. Sadeh et al. (2012)[54] and Cicchetti et al. (2012)[55] with a tied quality score of seven. When only these two studies remained the Liptak-Stouffer *P* value equals 0.1499. When reversing the step-wise removal, starting with the highest quality studies

first, the Liptak-Stouffer *P* value stayed significant until only the three studies with the lowest quality remained. We however emphasize that the low overall number of samples in our study makes all analyses based on exclusions very sensitive to outliers in our study.

P-curve

The result from the *P*-curve method regarding the *P* values with a positive direction of effect in terms of the S variant supported the conclusions of the other meta-analytical methods above. The right-skew *P* value (i.e. test for evidential value) for this subset of *P* values was <0.0001, see Figure 5.3 *P-curve – positive direction of effect*. When testing the two-tailed *P* values the right-skew *P* value was <0.0001, see Figure 5.4 *P-curve – positive and negative direction of effect*. The conclusion from the *P*-curve method is therefore in support of the above meta-analysis results, while at the same time avoiding assumptions regarding missing studies in our sample. Since both *P*-curves showed statistically significant signs of evidential value, we did not test if the observed distributions were flatter than what would be expected at a power of 33%. The results from the left-skew test of the *P*-curve (i.e. testing for signs of possible *P*-hacking) showed that neither of the *P*-curves were evident of possible *P*-hacking (both left-skew *P* values > 0.99).

Rosenthal's failsafe N

The Rosenthal's failsafe N_{fs} of 424.7 for the Liptak-Stouffer method indicated that a ratio of 35 (≈ 424.7 / 12) unpublished or undiscovered statistical tests with an average effect of zero (two-tailed *P* value = 1) and average effective sample-size (N_{eff} = 560) would be needed per included test in our analysis to make the Liptak-Stouffer *P* value non-significant. The N_{fs} for the Stouffer method resulted in an approximate ratio of 27 (≈ 321.4 / 12) and was hence in concordance with the N_{fs} for the Liptak-Stouffer method.

Discussion

To circumvent the issue of statistical power typical for single G×E studies, we conducted the first meta-analysis summarizing the results of individual studies that examined the potential interaction effect between *5-HTTLPR* and environmental adversity on the level of antisocial behavior. Despite the overall significant interaction effect found in this study, the results should be interpreted in the light of some important limitations. First, the use of the meta-analytical methods based on pooled *P* values limits the interpretation in terms of direction of the effect. Therefore, our analyses cannot demonstrate whether the significant interaction effect is driven by the S or L allele or a combination of both. Scholars have argued that significance of the Liptak-Stouffer test merely indicates that the G×E interaction is statistically different from zero in at least one of the included studies[37,56]. Therefore, once well-powered direct replication studies, employing hierarchical regression designs and reporting effect sizes, become available, more informative meta-analyses utilizing common effect-size metrics could further examine these results. Secondly, and most importantly, the moderate quality of the individual studies, and specifically the heterogeneity of their different tests and statistics undermines the reliability of our meta-analytical findings.

An important methodological shortcoming in most G×E studies is insufficient inclusion of potential confounding effects of covariates on the G×E interaction term. Even though most of the studies in our meta-analysis included confounding variables such as ethnicity, gender, age

in the regression equation to control for these covariates on the main effects of genotype and environment, they did not control for the effects of the covariates on the G×E interaction. Controlling for covariate interaction terms in the full model is preferred and yields more power than creating subsets of data based on the covariate and then examining the G×E term. Therefore, it is essential to include all relevant covariate×gene and covariate×environment interactions in the same model that tests the G×E term in order to properly control for the confounding effects of covariates[49]. Thus, G×E studies that improperly model covariates do not rule out alternative explanations for their findings.

Next, most studies included in our meta-analysis utilize ethnically heterogeneous samples (**Table 5.1**) and since allele frequencies of the *5-HTTLPR* polymorphism differ across ethnicities, stratification could be a potential confounding variable as an ethnicity-by-environment interaction offers an alternative explanation for the reported G×E effect[49,57]. None of the studies corrected for population structure at the genetic level, as genome-wide data were not available in these samples at the time of analysis. Besides cross-ethnic differences in allele frequencies, a study by Willams et al. (2003)[58] also demonstrated differential effects of *5-HTTLPR* on central nervous system serotonin function, with the short allele being associated with higher CSF *5-HIAA* levels in African Americans, but lower levels in Caucasians. These differential effects across ethnicity of *5-HTTLPR* on central nervous system serotonin function, could explain why the study of Douglas et al. (2011)[51] found a significant interaction effect in the African Americans but not in the European Americans (see **Table 5.3**). Hence, the use of admixed samples in the present study confines the validity and generalizability of our results and emphasizes the need of future G×E studies to employ ethnically homogenous samples or correct for population structure by using genome-wide data. The above concerns are supported by our systematic review and quality assessment revealing that all of the studies included had methodological shortcomings, such as absence or improper controlling for covariates. Lastly, given the heterogeneity in the type of measures, design, sex and age range across studies, the studies cannot be considered as direct replication attempts testing the same underlying hypothesis, necessitating further caution in the interpretation and generalization of the results[59].

Although the interpretation of our findings requires carefulness in the light of the aforementioned limitations, they are in line with previous animal studies showing a differentiated reactivity to environmental stressors between S and L carriers of *5-HTTLPR*[14,60]. The interaction effect on antisocial behavior found in this study offers several possible routes of explanation. An explanation is that L carriers of *5-HTTLPR* are somehow more resilient to environmental adversity than S carriers. This hypothesis is particularly interesting in view of research suggesting that the L allele might be a potential risk factor for psychopathy[61]. Psychopathy is associated with the more instrumental type of aggression and may have a different underlying etiology than reactive aggression[62]. Therefore, the L allele of *5-HTTLPR*, which is linked with hypo-responsivity to environmental factors, may thus predispose to instrumental aggression, whereas the S allele, associated with hyper-responsivity to the environment, may predispose to reactive aggression[61]. The study of Sadeh et al. (2010)[63], included in our meta-analysis, supports this line of reasoning. They reported a significant interaction between the L allele and callous-unemotional and narcissistic features of psychopathy in the presence of environmental adversity. This might explain why we did not find evidence for a directional effect in our meta-analysis, since we employed a broadly conceptualized outcome measure that included psychopathy.

Moreover, due to the liberal inclusion criteria because of the small number of studies identified in the literature, the meta-analysis comprises both population-based and clinical samples, which differ in severity and may thus have distinct neurobiological underpinnings. An alternative explanation for the interaction effect would be that a positive environment (i.e. secure

attachment relationships), somehow protects S or L carriers for their increased risk on developmental problems. Thus far, G×E research has predominantly focused on the genetic vulnerability towards *negative* environmental factors, such as maltreatment or harsh parenting. However, in potential, those individuals who are more malleable through their increased genetic plasticity, could also be more receptive to treatment and may benefit more from positive environments[64]. To what extent significant G×E effects fit into the framework of differential susceptibility or diathesis-stress remains to be understood as it cannot be inferred from the current meta-analysis whether the *5-HTTLPR* polymorphism resembles the characteristics of a differential susceptibility or risk gene. Further research in the exact etiological mechanisms underlying these differences in susceptibility or genetic risk is warranted, as more insights into the gene-environment interplay could improve prevention programs and could potentially lead to more tailored-based treatment[65]. Cohen & Piquero (2009)[65] estimated that preventing a high-risk youth from developing a criminal career could save society $2.6 million to $4.4 million. By increasing our knowledge on the etiological risk factors contributing to antisocial development, we could reduce the societal and financial burden of antisocial behavior.

Concerning the external validity, since our meta-analysis result is mainly based on studies that used adverse childhood environment as a source of environmental adversity (five out of eight studies), it is too early to conclude that the impact of other environmental factors, such as SES or peer affiliation is dependent on the *5-HTTLPR* genotype. It is thus important to further examine the relative contribution of specific environmental stressors in the G×E interaction model on antisocial behavior. A recent study that made use of genome-wide meta-analysis results from the Psychiatric Genomics Consortium showed that the effect of polygenic risk scores on depression was increased in the presence of childhood trauma[66]. Using the same approach, future research should examine whether individuals with high polygenic vulnerability exposed to childhood environmental adversity are particularly at risk for developing antisocial behavior.

To conclude, taking advantage of the rapid technological advancements in the field of genomics, including the rise of genome-wide data and polygenic risk score analysis, future studies should focus on how a broader polygenic profile interacts with environmental adversity factors to achieve a more sophisticated understanding of how biosocial interactions impact on antisocial development. Naturally, these polygenic studies should acknowledge the statistical issues accompanying G×E research, and properly control for potential confounders using well-powered study designs.

References – Chapter 5

1. DeLisi, M. *et al.* Murder by numbers: monetary costs imposed by a sample of homicide offenders. *J. Forens. Psychiatry Psychol.* **21,** 501–513 (2010).
2. McCollister, K. E., French, M. T. & Fang, H. The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation. *Drug Alcohol Depend.* **108,** 98–109 (2010).
3. Anda, R. F. *et al.* Adverse childhood experiences, alcoholic parents, and later risk of alcoholism and depression. *Psychiatr. Serv.* **53,** 1001–1009 (2002).
4. Hillis, S. D., Anda, R. F., Felitti, V. J. & Marchbanks, P. A. Adverse childhood experiences and sexual risk behaviors in women: a retrospective cohort study. *Fam. Plann. Perspect.* **33,** 206–211 (2001).
5. Jonson-Reid, M. *et al.* Effects of Child Maltreatment and Inherited Liability on Antisocial Development: An Official Records Study. *J. Am. Acad. Child Adolesc. Psychiatry* **49,** 321–332 (2010).
6. Reavis, J. A., Looman, J., Franco, K. A. & Rojas, B. Adverse childhood experiences and adult criminality: how long must we live before we possess our own lives? *Perm. J.* **17,** 44–48 (2013).
7. Piotrowska, P. J., Stride, C. B., Croft, S. E. & Rowe, R. Socioeconomic status and antisocial behaviour among children and adolescents: a systematic review and meta-analysis. *Clin. Psychol. Rev.* **35,** 47–55 (2015).
8. Monahan, K. C., Steinberg, L. & Cauffman, E. Affiliation with antisocial peers, susceptibility to peer influence, and antisocial behavior during the transition to adulthood. *Dev. Psychol.* **45,** 1520–1530 (2009).
9. Burt, S. A. & Klump, K. L. Etiological distinctions between aggressive and non-aggressive antisocial behavior: results from a nuclear twin family model. *J. Abnorm. Child Psychol.* **40,** 1059–1071 (2012).
10. Polderman, T. J. C. *et al.* Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat. Genet.* **47,** 702–709 (2015).
11. Rhee, S. H. & Waldman, I. D. Genetic and environmental influences on antisocial behavior: a meta-analysis of twin and adoption studies. *Psychol. Bull.* **128,** 490–529 (2002).
12. Roth, B. L. *The serotonin receptors: from molecular pharmacology to human therapeutics*. (Humana Press, 2006).
13. Bennett, A. *et al.* Early experience and serotonin transporter gene variation interact to influence primate CNS function. *Mol. Psychiatry* **7,** 118–122 (2002).
14. Suomi, S. J. Risk, resilience, and gene x environment interactions in rhesus monkeys. *Ann. N. Y. Acad. Sci.* **1094,** 52–62 (2006).
15. Bartolomucci, A. *et al.* Increased vulnerability to psychosocial stress in heterozygous serotonin transporter knockout mice. *Dis. Model. Mech.* **3,** 459–470 (2010).
16. Canli, T. & Lesch, K.-P. Long story short: the serotonin transporter in emotion regulation and social cognition. *Nat. Neurosci.* **10,** 1103–1109 (2007).
17. Vassos, E., Collier, D. A. & Fazel, S. Systematic meta-analyses and field synopsis of

genetic association studies of violence and aggression. *Mol. Psychiatry* **19,** 471–477 (2014).

18. Spinelli, S. *et al.* The serotonin transporter gene linked polymorphic region is associated with the behavioral response to repeated stress exposure in infant rhesus macaques. *Dev. Psychopathol.* **24,** 157–165 (2012).
19. Jaffee, S. R. *et al.* Nature X nurture: genetic vulnerabilities interact with physical maltreatment to promote conduct problems. *Dev. Psychopathol.* **17,** 67–84 (2005).
20. Tuvblad, C., Grann, M. & Lichtenstein, P. Heritability for adolescent antisocial behavior differs with socioeconomic status: gene-environment interaction. *J. Child Psychol. Psychiatry.* **47,** 734–743 (2006).
21. Gottesman, I. I. & Shields, J. A polygenic theory of schizophrenia. *Proc. Natl. Acad. Sci. U. S. A.* **58,** 199 (1967).
22. Monroe, S. M. & Simons, A. D. Diathesis-stress theories in the context of life stress research: implications for the depressive disorders. *Psychol. Bull.* **110,** 406 (1991).
23. Zuckerman, M. *Diathesis-stress models*. *Vulnerability to psychopathology: A biosocial model* 3–23 (American Psychological Association, 1999).
24. Belsky, J. & Pluess, M. Beyond diathesis stress: differential susceptibility to environmental influences. *Psychol. Bull.* **135,** 885–908 (2009).
25. Caspi, A. *et al.* Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science (80-.).* **301,** 386–389 (2003).
26. Munafo, M., Durrant, C., Lewis, G. & Flint, J. Gene x environment interactions at the serotonin transporterlocus. *Biol. Psychiatry* **65,** 211–219 (2009).
27. Risch, N. *et al.* Interaction Between the Serotonin Transporter Gene (5-HTTLPR), Stressful Life Events, and Risk of Depression. *JAMA* **301,** 2462–2471 (2009).
28. Karg, K., Burmeister, M., Shedden, K. & Sen, S. The serotonin transporter promoter variant (5-HTTLPR), stress, and depression meta-analysis revisited. *Arch. Gen. Psychiatry* **68,** 444–454 (2011).
29. Duncan, L. E. & Keller, M. C. A Critical Review of the First 10 Years of Candidate Gene-by-Environment Interaction Research in Psychiatry. *Am. J. Psychiatry* **168,** 1041–1049 (2011).
30. Dick, D. M. *et al.* Candidate Gene–Environment Interaction Research Reflections and Recommendations. *Perspect. Psychol. Sci.* **10,** 37–59 (2015).
31. Okbay, A. & Rietveld, C. A. On Improving the Credibility of Candidate Gene Studies : A Review of Candidate Gene Studies Published in Emotion. *Emotion* **15,** 531–537 (2015).
32. Rutter, M. & Maughan, B. Psychosocial adversities in childhood and adult psychopathology. *J. Pers. Disord.* **11,** 4–18 (1997).
33. Hayden, J., Cote, P. & Bombardier, C. Evaluation of the quality of prognosis studies in systematic reviews. *Ann. Intern. Med.* **144,** 427–437 (2006).
34. Kavvoura, F. K. & Ioannidis, J. P. A. Methods for meta-analysis in genetic association studies: A review of their potential and pitfalls. *Hum. Genet.* **123,** 1–14 (2008).
35. Becker, B. J. Combining Significance Levels. in *The Handbook of Research Synthesis*

(eds. Cooper, H. & Wedges, L. V) 215–230 (Russell Sage Foundation, 1994).

36. Borenstein, M., Hedges, L. V., Higgins, J. & Rothstein, H. R. *Front matter*. (Wiley Online Library, 2009).

37. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. Front Matter. in *Introduction to Meta-Analysis* i–xxix (John Wiley & Sons, Ltd, 2009).

38. Whitlock, M. C. Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *J. Evol. Biol.* **18,** 1368–1373 (2005).

39. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26,** 2190–2191 (2010).

40. Borenstein, M., Hedges, L. V, Higgins, J. P. T. & Rothstein, H. R. *Meta-Analysis Methods Based on Direction and p-Values*. *Introduction to Meta-analysis* (John Wiley & Sons, 2009). doi:10.1002/9780470743386.ch36

41. Stock, J. H. & Watson, M. W. *Introduction to Econometrics*. (Pearson Education Limited, 2012).

42. Borenstein, M., Hedges, L. V., Higgins, J. P. T. & Rothstein, H. R. *Introduction to Meta-Analysis*. (2009). doi:10.1002/9780470743386

43. Simonsohn, U., Nelson, L. D. & Simmons, J. P. P-curve: A key to the file-drawer. *J. Exp. Psychol. Gen.* **143,** 534–547 (2014).

44. Franco, A., Malhotra, N. & Simonovits, G. Publication bias in the social sciences: Unlocking the file drawer. *Science (80-.).* **345,** 1502–1505 (2014).

45. Orwin, R. G. A fail-safe N for effect size in meta-analysis. *J. Educ. Stat.* 157–159 (1983).

46. Rosenthal, R. The 'file drawer problem' and tolerance for null results. *Psychol. Bull.* **86,** 638–641 (1979).

47. Becker, B. J. Chapter 7 – Failsafe N or file-drawer number. in *Publication bias in meta-analysis: Prevention, assessment and adjustments* (eds. Rothstein, H. R., Sutton, A. J. & Borenstein, M.) 111–125 (Wiley & Sons, Ltd, 2005).

48. Rothstein, H. R., Sutton, A. J. & Borenstein, M. *Publication Bias in Meta-Analysis - Prevention, Assessment and Adjustments*. (2005).

49. Keller, M. C. Gene × Environment Interaction Studies Have Not Properly Controlled for Potential Confounders: The Problem and the (Simple) Solution. *Biol. Psychiatry* **75,** 18–24 (2014).

50. Aslund, C. *et al.* Impact of the interaction between the 5HTTLPR polymorphism and maltreatment on adolescent depression. A population-based study. *Behav. Genet.* **39,** 524–531 (2009).

51. Douglas, K. *et al.* 5-HTTLPR as a Potential Moderator of the Effects of Adverse Childhood Experiences on Risk of Antisocial Personality Disorder. *Psychiatr. Genet.* **21,** 240–248 (2011).

52. Åslund, C. *et al.* Self-Reported Family Socioeconomic Status, the 5-HTTLPR Genotype, and Delinquent Behavior in a Community-Based Adolescent Population. *Aggress. Behav.* **39,** 52–63 (2013).

53. Kretschmer, T., Sentse, M., Dijkstra, J. & Veenstra, R. The interplay between peer

rejection and acceptance in preadolescence and early adolescence, serotonin transporter gene, and antisocial behavior in late adolescence: The TRAILS Study. *Merrill. Palmer. Q.* **60,** 193–216 (2014).

54. Sadeh, N., Javdani, S. & Verona, E. Analysis of monoaminergic genes, childhood abuse, and dimensions of psychopathy. *J. Abnorm. Psychol.* **122,** 167–179 (2012).

55. Cicchetti, D., Rogosch, F. a. & Thibodeau, E. L. The effects of child maltreatment on early signs of antisocial behavior: Genetic moderation by tryptophan hydroxylase, serotonin transporter, and monoamine oxidase A genes. *Dev. Psychopathol.* **24,** 907–928 (2012).

56. Laoutidis, Z. G. & Luckhaus, C. The Liptak-Stouffer Test for Meta-Analyses. *Biol. Psychiatry* **77,** e1–e2 (2015).

57. Ioannidis, J. P. A., Ntzani, E. E. & Trikalinos, T. A. 'Racial' differences in genetic effects for complex diseases. *Nat. Genet.* **36,** 1312–1318 (2004).

58. Williams, R. B. *et al.* Serotonin-related gene polymorphisms and central nervous system serotonin function. *Neuropsychopharmacol. Off. Publ. Am. Coll. Neuropsychopharmacol.* **28,** 533–541 (2003).

59. Manuck, S. B. & McCaffery, J. M. Gene-environment interaction. *Annu. Rev. Psychol.* **65,** 41–70 (2014).

60. Champoux, M. *et al.* Serotonin transporter gene polymorphism, differential early rearing, and behavior in rhesus monkey neonates. *Mol Psychiatr* **7,** 1058–1063 (2002).

61. Glenn, A. L. The other allele: exploring the long allele of the serotonin transporter gene as a potential risk factor for psychopathy: a review of the parallels in findings. *Neurosci. Biobehav. Rev.* **35,** 612–620 (2011).

62. Glenn, A. L. & Raine, A. Psychopathy and instrumental aggression: Evolutionary, neurobiological, and legal perspectives. *Int. J. Law Psychiatry* **32,** 253–258 (2009).

63. Sadeh, N. *et al.* Serotonin transporter gene associations with psychopathic traits in youth vary as a function of socioeconomic resources. *J. Abnorm. Psychol.* **119,** 604–609 (2010).

64. Belsky, J. *et al.* Vulnerability genes or plasticity genes? *Mol. Psychiatry* **14,** 746–754 (2009).

65. Cohen, M. A. & Piquero, A. R. New evidence on the monetary value of saving a high risk youth. *J. Quant. Criminol.* **25,** 25–49 (2009).

66. Peyrot, W. J. *et al.* Effect of polygenic risk scores on depression in childhood trauma. *Br. J. Psychiatry J. Ment. Sci.* **205,** 113–119 (2014).

67. Bruns, S. B. & Ioannidis, J. P. A. P-curve and p-hacking in observational research. *PLoS One* **11,** 1–13 (2016).

Figures and tables – Chapter 5

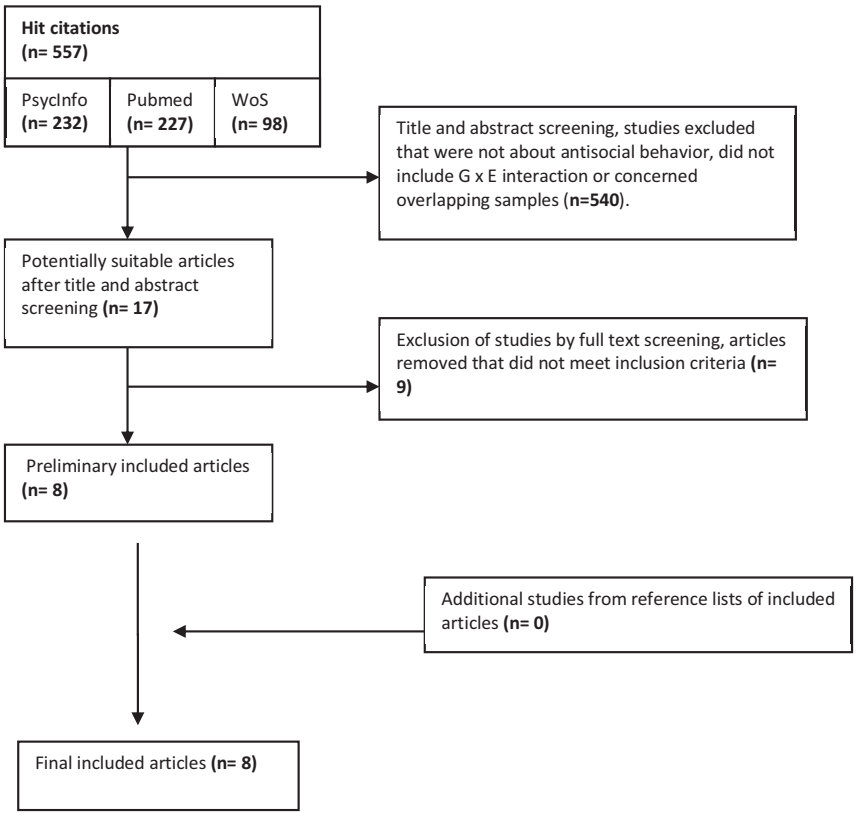

Figure 5.1. Flowchart of study selection.

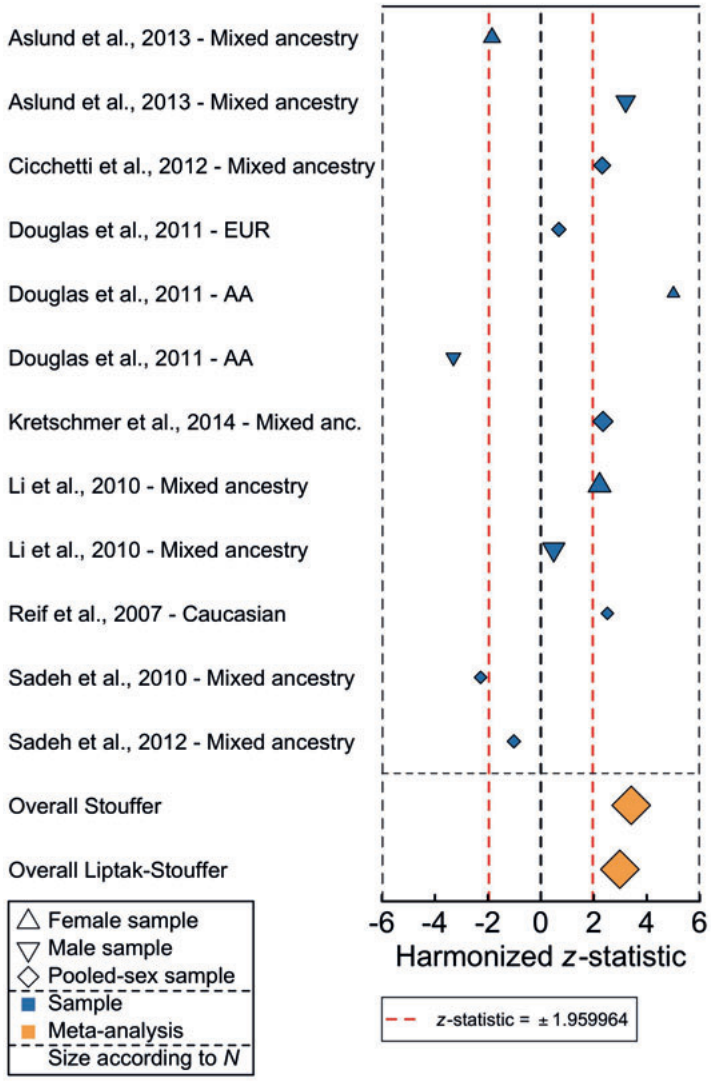

Figure 5.2. Forest plot – Harmonized *z*-statistics and meta-analysis estimates. Forest plot of the harmonized *z*-statistics, as well as the overall Stouffer and Liptak-Stouffer *z*-statistics.

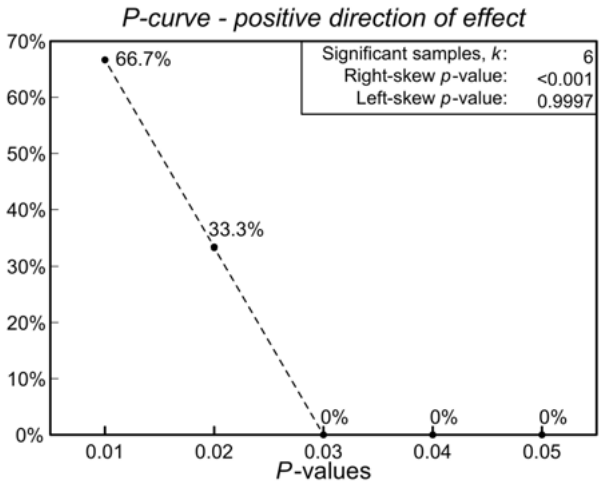

Figure 5.3. *P*-curve – positive direction of effect. *P*-curve testing the one-tailed hypothesis of a positive effect in terms of the S allele by including the subset of two-tailed statistically significant *P*-values from estimates with a positive interaction effect divided by two.

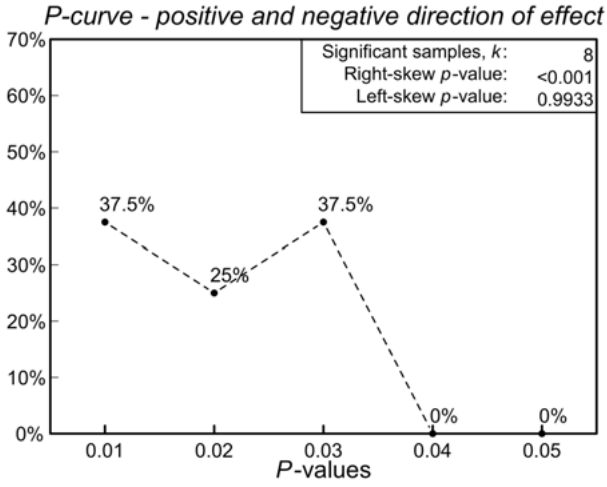

Figure 5.4. *P*-curve - positive and negative direction of effect. *P*-curve testing the two-tailed hypothesis of a positive or negative effect in terms of the S allele by including all of the statistically significant two-tailed *P*-values in our sample.

Table 5.1 | Baseline characteristics of all included studies

Study	Sampling	% Male	Mean Age (M), Age Range or SD	Total N	Ancestry	Antisocial Measure & Design	Environmental adversity (EA)	Grouping: Genotype	Statistical Model
Aslund et al. 2013	**Cross-sectional -** Survey of Adolescent Life Vestmanland	51.3	17-18	1467	Mixed	Sum score based on 15-item delinquency scale	SES, 7-point Likert scale; divided in low, medium and high SES	SS vs LS vs LL	General linear model
Cicchetti et al. 2012	**Cross-sectional -** Children participating in summer camp research programme for school-aged low-income children	49.8	M= 11.3 (10-12)	627	Mixed	Pittsburgh Youth Survey self-report; Peer ratings; teacher report form	MCS based on DHS records (physical abuse, neglect, emotional maltreatment)	SS vs LS vs LL	ANCOVA
Douglas et al. 2011	**Cross-sectional -** Subjects dependent on alcohol, cocaine or opioids at 4 sites	58.9	European American/African American M= 38.2/41.0 (37-42)	1381	European American and African American	44-item SSADDA, specific ASPD criteria; dichotomous score	ACE (violent crime, exposure, sexual abuse, physical abuse) Ordinal score 0-3	SS vs LS vs LL	Logistic generalized estimating equations regression
Kretschmer et al. 2014	**Longitudinal -**Prospective cohort Dutch adolescents-Tracking Adolescents' Individual Lives Survey (TRAILS)	Not stated	T1: M=11.1	1118	Mixed	31 items of ASBQ at T1	Peer rejection & acceptance; nomination assessment	SS vs LS vs LL	Regression models
Li et al. 2010	**Cross-sectional -** National Longitudinal Study of Adolescent Health - random sample of youth from US high schools	48.0	M=15.7 (12-20)	2488	Mixed	In-home interview on frequency of 7 overt ASB items, converted to summed score	Retrospective In-home interview on maltreatment (neglect, physical abuse, sexual abuse)	SS vs SL vs LL	Latent class multinomial logistic regression
Reif et al. 2007	**Cross-sectional -** Volunteers in forensic examination to the Institute of Forensic Psychiatry of the University of the Saarland for evaluation of legal responsibility or risk assessment	100	Violent group (N=72, M=32.1 ± 12.1) Non-violent group (112, M=35.4 ± 11.3)	184	Caucasian	Subjects assigned to a violent or nonviolent group by experts	Childhood Adverse Environmental Index-mean score calculated 0-2	S/S and S/L vs. L/L	Stepwise logistic regression
Sadeh et al. 2010	**Cross-sectional -** Youth from treatment or legal agencies and general community	42.0	M= 14.3, SD=1.5	118	Mixed	20-item self-report Antisocial Process Screening Device-	Family income level (three categories) and occupation (coded 1-7)	LL vs. SL vs. SS	Hierarchical linear regression

	recruited via agency referrals, newspaper advertisement and posted flyers					Callous-Unemotional; narcissism; impulsivity			
Sadeh et al. 2012	**Cross-sectional -** Selected on high rates of antisocial behaviour and high psychosocial adversity. Recruited via probation/parole agencies, county jail, and mandated treatment centers, as well as via newspaper advertisements targeting individuals with legal convictions	100	M=30.9 (18-61)	237	Mixed	Psychopathy Checklist: Screening Version - Composite score of two-factor and three-factor model	Childhood abuse, CTQ Composite scale, continuous	SS vs. LL	Hierarchical regression analyses

Note: ACE: adverse childhood events; ASB: antisocial behaviour; ASBQ: antisocial behaviour questionnaire, ASPD: antisocial personality disorder; CTQ: Childhood trauma questionnaire; DHS: department of human services; M: mean; MCS: maltreatment classification; SES: social economic status; SSADDA: Semi-structured Assessment for Drug Dependence and Alcoholism.

Table 5.2 | Quality assessment results of included studies

Domain Criterion	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**Total**
Cichetti et al., 2012	+	+	+	+	+	+	+	-	7
Sadeh et al., 2012	+	+	+	+	+	+	+	-	7
Aslund et al., 2013	+/-	+	+	+	+	+	+	-	6.5
Li et al., 2010	+	+	+/-	+	+/-	+	+	-	6
Sadeh et al., 2010	+	+	-	+	+	+	+	-	6
Douglas et al., 2011	+	+	-	+	+	+	+/-	-	5.5
Reif et al., 2007	+	+	+	+/-	-	+	+	-	5.5
Kretschmer et al., 2014	+/-	+/-	-	+	+	+/-	-	-	3.5

Note: A: Clear description of the key characteristics of the study population. B: The sampling frame and recruitment are described. C: Inclusion and exclusion criteria are described. D: Clear definition of the measure of antisocial behavior and EA is provided. E: The measure of antisocial behavior and EA is adequately valid to limit misclassification. F: Age and gender are accounted for in the analysis. G: IQ or SES or Race are accounted for in the analysis. H: Covariate interaction terms are accounted for in the analysis.

Table 5.3 | Extracted and harmonized values of the meta-analysis sample, in total 12 samples derived from eight studies

Study, Year	Effective N[i], Sex, Ethnicity	Effect allele in statistical test	Extracted test statistic of the GxE interaction	Reported *p*-value	Environmental adversity measure
Aslund et al., 2013	464, F, Mixed	NA	F(2, 661) = 2.74	0.065	Socio-economic status
Aslund et al., 2013	752, M, Mixed	NA	F(2,689) = 6.67	0.001	Socio-economic status
Cicchetti et al., 2012	627, M/F, Mixed	NA	F(2, 586) = 3.91	0.02	Maltreatment
Douglas et al., 2011	315, M/F, EA	NA	Chi-square statistic = 1.4, df = 2	0.5	Adverse Childhood events
Douglas et al., 2011	112, F, AA	S	Logit beta = 1.8181, SE = 0.3422. [Calculated from OR in Table 6, p. 20]	<0.001	Adverse Childhood events
Douglas et al., 2011	261, M, AA	S	Logit beta = 1.0296, SE = 0.3083 [Calculated from OR in Table 6, p. 20]	<0.001	Adverse Childhood events
Kretschmer et al., 2014	1118, M/F, Mixed	S	Standardized beta = 0.021 [Table 4, and *p*-value from correspondence with author]	0.019	Peer rejection
Li et al., 2010	1294, F, Mixed	L	Log-linear Beta = −1.82, SE = 0.82	<0.05	Maltreatment
Li et al., 2010	1194, M, Mixed	L	Log-linear Beta = −0.37, SE = 0.76	0.62	Maltreatment
Reif et al., 2007	172, M/F, Caucasian	L	Logit Beta = -1.65, SE = 0.65	0.011	Childhood adverse environment
Sadeh et al., 2010	178, M/F, Mixed	L	Linear regression Beta = −0.16	0.023	Socio-economic status
Sadeh et al., 2012	237, M/F, Mixed	L	Linear regression *t*-statistic = 1.02	0.31	CTQ Childhood abuse

Table 5.3 continues on next page.

[i] Effective sample-size calculated as reported in methods section.

Table 5.3 cont.

Study, Year	Direction of interaction effect [Reference to original study]	Harmonized direction of G x E[u] (S variant)	Calculated two-tailed *p*-value[v]	Harmonized *z*-statistic[2]
Aslund et al., 2013	L decreases ASB in females when interacted with E, however SES reverse coded [p. 59; Fig. 2, p. 58]	Negative	0.0653	-1.8432
Aslund et al., 2013	L increases ASB in males when interacted with E, however SES reverse coded [p. 59; Fig. 2, p. 58]	Positive	0.0014	3.2047
Cicchetti et al., 2012	L decreases ASB when interacted with E [Fig 8. p. 920]	Positive	0.0206	2.3159
Douglas et al., 2011	Not significant (L decreases ASB in EA when interacted with E [Table 5, p. 19]	Positive	0.4966	0.6799
Douglas et al., 2011	L decreases ASB in AA females when interacted with E. [Table 6, p. 20]	Positive	<0.0001	5.0064
Douglas et al., 2011	L increases ASB in AA males when interacted with E. [Table 6, p. 20]	Negative	0.0010	-3.3016
Kretschmer et al., 2014	L decreases ASB when interacted with E [Table 4, p. 205]	Positive	0.0190[4]	2.3455
Li et al., 2010	L decreases ASB in females when interacted with E [p. 795]	Positive	0.0266	2.2170
Li et al., 2010	Not significant (L decreases ASB in males when interacted with E) [p. 795]	Positive	0.6265	0.4867
Reif et al., 2007	L decreases ASB when interacted with E [p. 2379]	Positive	0.0120	2.5115
Sadeh et al., 2010	LL decreases ASB when interacted with E, which SL and SS does not, however SES reverse coded [p. 6]	Negative	0.0230[w]	-2.2734
Sadeh et al., 2012	Not significant (L increases ASB when interacted with E) [p. 7]	Negative	0.3088	-1.0178

Note: AA: African Americans; ASB: Antisocial Behavior; EA: European Americans; F: Female; M: Male; NA: Not Available

[u] Positive indicates greater sensitivity among S variant 5-HTTLPR subjects to environmental adversity, as presented in the original study report. Negative indicates a greater sensitivity among L variant 5-HTTLPR subjects to environmental adversity. For further discussion of the extraction and calculation of the harmonized *z*-statistics, see methods section.

[v] The values displayed in the table are rounded; exact *p*-values were provided in all calculations and analyses.

[w] Study does not report the required information in order to calculate an exact *p*-value; hence the reported *p*-value is included in the meta-analysis.

Table 5.4 | Results of Stouffer's z meta-analysis

Study, Year	**Effective sample-size, sex, ethnicity**	**Direction of effect***	**Harmonized *z*-statistic**	**Two-tailed *p*-value of weighted Liptak-Stouffer z-statistic after Study Exclusion**
Aslund et al., 2013	464 F	Negative	-1.8432	<0.0001
Aslund et al., 2013	752 M	Positive	3.2047	0.0130
Cicchetti et al., 2012	627 M/F	Positive	2.3159	0.0045
Douglas et al., 2011	315 M/F, EA	Positive	0.6799	0.0008
Douglas et al., 2011	112 F, AA	Positive	5.0064	0.0053
Douglas et al., 2011	261 M, AA	Negative	-3.3016	<0.0001
Kretschmer et al., 2014	1118 M/F	Positive	2.3455	0.0071
Li et al., 2010	1294 F	Positive	2.2170	0.0066
Li et al., 2010	1194 M	Positive	0.4867	0.0004
Reif et al., 2007	172 M/F	Positive	2.5115	0.0023
Sadeh et al., 2010	178 M/F	Negative	-2.2734	0.0001
Sadeh et al., 2012	237 M/F	Negative	-1.0178	0.0002
Total effective sample size	**6724**	**Overall two-tailed *p*-value of weighted Liptak-Stouffer**		**0.0006**

Note: F: female; M: male; EA: European Americans; AA: African Americans

*Positive indicates greater sensitivity among S variant *5-HTTLPR* subjects to environmental adversity, as presented in the original study report. Negative indicates a greater sensitivity among L variant *5-HTTLPR* subjects to environmental adversity. For further discussion of the extraction and calculation of the *Harmonized z-statistic*, see methods section and **Supplementary Materials *III)*** *Transformation of p-values into a common z-statistic*.