Neuroticism and the brain
Servaas, Michelle N.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2015

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):
Chapter 7

FILLING THE GAP: RELATIONSHIP BETWEEN THE SEROTONIN-
TRANSPORTER-LINKED POLYMORPHIC REGION AND AMYGDALA
ACTIVATION

JOJANNEKE A. BASTIAANSEN
MICHELLE N. SERVAAS
JAN-BERNARD C. MARSMAN
JOHAN ORMEL
ILJA M. NOLTE
HARRIETTE RIESE
ANDRÉ ALEMAN

PSYCHOLOGICAL SCIENCE, 2014; 25(11):2058-66
7.1 Abstract

The alleged association between the serotonin transporter polymorphism (5-HTTLPR) and amygdala activation forms a cornerstone of the common view on the 5-HTTLPR as a potential risk factor for affective disorders. A recent meta-analysis showed that this association is statistically significant (Hedge $g = + 0.35$), but at the same time warned against estimate distortions due to publication bias (Murphy et al., 2013). Here, we report a replication study of this relationship in 120 participants which failed to find an association of 5-HTTLPR variation with amygdala activation during a widely used emotional face matching task. Moreover, we show that the pooled meta-analytic effect size is no longer significant ($g = + 0.20$, $p = 0.06$) when unpublished studies are included and the meta-analysis is updated with our study. These findings cast doubt on previously reported substantial effects, suggesting the 5-HTTLPR-amygdala association is either much smaller than previously thought or conditional on other factors.
7.2 Introduction

Advances in neuroimaging techniques have as yet been unsuccessful in providing clinically relevant and reliable biomarkers for psychiatric disorders (Kapur et al., 2012). One major difficulty is that psychiatric disorders are presumed to be the “endpoints of multiple converging pathophysiological pathways” (Linden, 2012: p. 14), which cut across diagnostic borders and are not necessarily similar in individuals of the same diagnostic category (Insel et al., 2010). A potentially fruitful approach to transcend diagnostic boundaries and probe specific biological pathways associated with psychiatric disorders has been offered by the field of imaging genetics. This approach investigates how genetic variants with a known neurochemical effect influence specific components of a biological pathway (Linden, 2012; Meyer-Lindenberg and Weinberger, 2006).

A premier finding in the field of imaging genetics has been an increased amygdala response to negative emotional stimuli in carriers of the short (S) allele of the serotonin transporter polymorphism (5-HTTLPR) (Hariri et al., 2002). S-carriers had previously been found to score higher on questionnaires measuring the personality trait neuroticism (Lesch et al., 1996), an important risk factor for the development of affective disorders (Lahey, 2009; Ormel et al., 2013b). Since these early findings, the idea of the 5-HTTLPR as a potential genetic risk factor for affective disorders has taken a firm hold in the literature (e.g. Bogdan et al., 2013; Canli et al., 2006). The most recent meta-analysis on the increasing number of genetic imaging studies has indicated that there is a statistically significant, but small effect of the 5-HTTLPR on amygdala activation (Hedge’s g = +0.35 for published studies, Murphy et al., 2013). However, the reported excess of significance among published studies with on average low statistical power (Murphy et al., 2013) points to a potential publication bias, which could mean that even their modest estimate represents an overestimation of the true effect size (Borenstein et al., 2009). In addition to the relative absence of (small and hence) low-accuracy studies with null findings, there is a gap with regard to high-accuracy studies (Munafo et al., 2008; Murphy et al., 2013). Therefore, the current state of the literature could undermine the validity of (otherwise methodologically sound) meta-analyses (Borenstein et al., 2009).

In this article, we reevaluate the claim that the short variant of the 5-HTTLPR relates to amplified amygdala activation by 1) performing fMRI scans in a relatively large sample of 120 participants during a task that has consistently been found to engage the amygdala (Bertolino et al., 2005; Hariri et al., 2000, 2002; Tessitore et al., 2002), and 2) running an updated meta-analysis by adding our findings to the published as well as unpublished studies reported by Murphy and colleagues (2013). We increased our study’s sensitivity a priori by targeting a relatively homogeneous sample of women, who are known to have a higher risk of developing affective disorders compared to men (Parker and Brotchie, 2010). The sample size gave our study 99.99%
power\(^1\) to detect the effect reported in the original study (Hariri et al., 2002). Moreover, it is the largest of studies on 5-HTTLPR and amygdala response to negative stimuli thus far; our sample size is only surpassed by a study on 5-HTTLPR genotype and baseline perfusion in the amygdala (Viviani et al., 2010: \(n = 183\)). Thus, it provides an important step in filling the high-accuracy study gap evident in meta-analyses (Munafo et al., 2008; Murphy et al., 2013).

7.3 Methods

7.3.1 Participants

This study was part of a larger project on the neural correlates of neuroticism (e.g. Servaas et al., 2014a, 2014b). Participants were 120 female students (mean age 20.8 ± 2.0 years) from different faculties of the University of Groningen, who were selected from a larger sample of 240 students on the basis of their scores on the neuroticism scale (12 items) of the NEO Five-Factor Inventory (Hoekstra et al., 1996). Half of the scanned participants were drawn from the highest 25% of neuroticism scores and the other half was randomly selected from the lower 75% of scores. The selection procedure resulted in a normal distribution of neuroticism scores (as reassessed by the 48-item neuroticism scale of the NEO Personality Inventory Revised, NEO-PI-R, Hoekstra et al., 1996) around a mean of 135.5 (± 18.9), which is in a similar range as the female student norm group in the NEO-manual (\(n = 690, 143.6 ± 21.0\)). Thus, despite our overselection of relatively high neuroticism scores, the mean level of neuroticism in our participants was not higher than in the norm group. Furthermore, we measured two other personality traits from the NEO, extraversion and conscientiousness; the scores were in the same range as the norm group: 165.9 ± 18.6 for extraversion (159.6 ± 18.9 for the norm group) and 169.3 ± 19.8 for conscientiousness (160.2 ± 20.4 for the norm group). All participants were scanned in the first ten days of their menstrual cycle or in their pill-free week to control for menstrual cycle-related effects on (neural correlates of) mood, stress sensitivity and neurocognitive function (Andreano and Cahill, 2010; Goldstein et al., 2010; Symonds et al., 2004). None of the participants had any present or past neurological or psychiatric disorders or used medication that could influence task results. Participants were right-handed native Dutch speakers of Caucasian descent, between 18-25 years of age, had normal hearing and (corrected-to-normal) vision and were eligible for MRI research. All gave written informed consent to participate in the study, which was approved by the Institutional Review Board of the University Medical Center Groningen, Groningen, the Netherlands.

7.3.2 Genotyping

DNA extraction and genotyping were performed at the department of Laboratory Medicine of the University Medical Center Groningen, Groningen, the Netherlands. Saliva was collected in

\(^1\) Statistical power was calculated by G‘Power (Faul et al., 2007) at an \(\alpha\) level of 0.05 for the effect size of the original study (Hariri et al., 2002) as reported in the meta-analysis by Murphy and colleagues (2013), Hedge’s \(g = + 0.97\).
Oragene saliva collection and preservation kits (DNAGenotek, Ontario, Canada), and DNA was extracted according to the protocol of the manufacturer. For SLC6A4, the 5-HTTLPR S/La/Lg variants were determined using PCR with Forward primer FAM-5’TGAATGCCAGCACCTAACC-3’ and Reverse primer 5’TCTCGGTCCACCTAGACG-3’ (35 cycles of 30 seconds at 95°C, 30 seconds at 61°C and 1 minute at 72°C), and subsequent ingestion of the PCR product with the restriction enzyme Msp-I for at least 3 hours at 37 °C. The resulting restriction fragments were separated using capillary electrophoresis (ABI 3130 analyzer; Applied Biosystems, Nieuwerkerk a/d IJssel, the Netherlands) and fragment lengths were estimated using the ABI Prism® GeneMapper™ software, version 3.0 (Applied Biosystems). The La, Lg and S variants were determined by the detection of fragments 325 base pairs (bp), 152 bp or 284 bp, respectively. The method is based on and validated using the method as described by Doornbos et al. (2009). Analyses were performed in two runs; in each run, two control samples for 5-HTTLPR were genotyped for between plate reproducibility. All duplicates showed 100% concordance.

The 5-HTTLPR Lg variant behaves as the low-expressing S allele (Hu et al., 2006; Wendland et al., 2006). In the remainder of this paper, the S and Lg variants will be referred to as the short (S) allele and the La variant as the L allele. In this paper, we will compare S-carriers with L/L homozygotes (Hariri et al., 2002), because the serotonin reuptake of cells with the homozygous L variant is twice as high compared to cells that carry either one or two copies of the S variant (Lesch et al., 1996).

### 7.3.3 Behavioral task

During the 60-minute fMRI session, participants completed a resting state scan and four tasks, which were programmed in E-prime (Psychology Software Tools, Pittsburgh, PA). All participants started with a modified version of the widely used emotional face matching task developed by Hariri and colleagues (2000, 2002). The blocked paradigm consisted of an emotion task interleaved with a sensorimotor control task. In the control blocks, participants had to indicate on a button box which of two shapes presented at the bottom of the screen matched a target geometrical shape (circle, horizontal or vertical ellipse) presented at the top of the screen. In the emotion blocks, participants had to match the affect of a target face to one of two simultaneously presented faces. There were three different images for each combination of gender type (male/female) and type of affect (angry/fearful). Each trio of images was presented for maximally 5 seconds. The task consisted of two periods with each four control and four emotion blocks. Each block consisted of 10 trials. After every second block, there was a rest period of 10 seconds. We used a self-paced task, in which participants were instructed to respond as quickly as possible. Response accuracy and reaction times were recorded throughout the task.

### 7.3.4 fMRI data collection

Scans were obtained using a 3T Philips Intera Quaser (Best, the Netherlands) equipped with
a synergy SENSE 32-channel head coil. Functional images were acquired using a T2*-weighted echo-planar sequence with 39 descending axial slices without slice gap to cover the entire cortex (voxel size = 3.5 x 3.5 x 3.5 mm, TR = 2000 ms, TE = 30 ms, FOV = 224 x 224 mm). In addition, high-resolution T1-weighted structural images were acquired containing 170 slices (voxel size = 1 x 1 x 1 mm, TR = 9 ms, TE = 8 ms, FOV = 256 x 231 mm). To reduce artifacts from the nasal cavities, images were tilted 10° from the AC-PC transverse plane.

7.3.5 Preprocessing fMRI data

All image processing was performed using the Statistical Parametric Mapping 8 (SPM8) package (Wellcome Department of Cognitive Neurology, London, UK; http://www.fil.ion.ucl.ac.uk) in Matlab R2009a (The MathWorks Inc., Natick, MA). Data preprocessing comprised realignment to correct for shifts in head position, coregistration, spatial normalization into a standard space using a Montreal Neurological Institute (MNI) T1 template and smoothing with an 8 mm full-width half-maximum (FWHM) isotropic Gaussian kernel to minimize noise and accommodate residual variations in neuroanatomy between participants. Data of three participants were excluded due to excessive head movement (n=1) or structural abnormalities (n=2).

7.3.6 Statistical analysis

Descriptive statistics were calculated and independent sample t-tests were used to identify genotype effects on neuroticism scores and response speed. Allele frequencies of the three genotypes of 5-HTTLPR (L/L, L/S, S/S) were calculated and analyzed for deviation from Hardy-Weinberg equilibrium (HWE) using a chi-square test with 1 degree of freedom.

7.3.7 Task activation effects

Emotion and control trials were analyzed at the subject level by boxcar functions convolved with the hemodynamic response function after 128 s high-pass filtering to remove low-frequency noise and slow drifts in the signal. Head movement was accommodated by six motion regressors and their first temporal derivatives. For each participant, a voxel-by-voxel t-map was computed for the emotion versus control contrast. A one-sample t-test was performed on these images to determine the task effect across the sample (p<0.05, family wise error (FWE)-corrected). The images were then used in a second-level random-effects (RFX) model with between-subject factor genotype and tested for the contrast S-carriers > L/L homozygotes (FWE-p<0.05). Neuroticism scores were added as a covariate in a second step to investigate whether sample selection influenced the results. Because of our a priori hypothesis on the amygdala, we used a Small Volume Correction (SVC) for this structure, which was anatomically defined using the AAL template of the WFU Pick Atlas (Version 3.0) (Maldjian et al., 2003; Tzourio-Mazoyer et al., 2002).
7.3.8 Meta-analysis update

For each participant, we extracted for the emotion versus control contrast mean activation in left and right amygdala regions of interest based on the AAL template of the WFU Pick Atlas (Version 3.0) (Maldjian et al., 2003; Tzourio-Mazoyer et al., 2002). Then, we calculated the effect size (Hedge's g) of the association between the 5-HTTLPR and amygdala activation based on the group means and standard deviations for the S-carriers and L/L homozygotes. This data from our individual study was added to comparable information from the 29 published and 5 unpublished studies reported in the meta-analysis of Murphy and colleagues (2013). Finally, we reran their RFX analysis to update the estimate of the pooled effect size using the Comprehensive Meta-Analysis (v2) statistical software package (Biostat, Englewood, NJ, USA). The effect sizes of the individual studies were plotted against study accuracy/precision. Study precision is calculated as 1/standard error, which is closely related to the study's sample size. Furthermore, the effect sizes of the individual studies were plotted against year of publication. Between-study heterogeneity was assessed with the I² statistic.

7.4 Results

7.4.1 Sample characteristics

The prevalence of the 5-HTTLPR variants in our sample (L/L: n = 31, 25.8%; S/L: n = 57, 47.5%; S/S: n = 32, 26.7%) resembled findings from the European HapMap31 (Hu et al., 2006) and did not deviate from HWE (p = 0.58). S-carriers did not differ significantly from L/L homozygotes regarding their mean neuroticism score (136.2 ± 19.2 versus 133.3 ± 18.2, respectively; t\textsubscript{(118)} = 0.76, p = 0.45).

7.4.2 Behavioral data

Task accuracy was very high during the control blocks (99.2% ± 1.1) as well as the emotion blocks (96.0% ± 4.0). Due to these ceiling effects, we did not execute the test for differences in accuracy between S-carriers and L/L homozygotes. There was no significant effect of genotype on response speed for the emotion and control conditions (both p's > 0.48).

7.4.3 Task activation effects

To validate our task, we first tested for a significant task effect (emotion – control) across all participants. Task-related effects were significant in large lateral prefrontal clusters and bilateral amygdala/hippocampus in addition to the cerebellum, parietal regions and primary and extrastriate visual cortical areas (Table 1). There were no significant clusters for the contrast between S-carriers and L/L homozygotes. Small-volume correction for the amygdala did not reveal any significant
voxels either. No significant clusters were found when neuroticism was added to the model as a covariate either.
Table 1 Activations during perceptual processing of negative facial expressions

<table>
<thead>
<tr>
<th>Location</th>
<th>H</th>
<th>Cluster description</th>
<th>Cluster size</th>
<th>Peak T</th>
<th>MNI coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x</td>
</tr>
<tr>
<td>Emotional faces &gt; geometrical shapes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Precuneus</td>
<td>B</td>
<td>Precuneus, Superior Parietal Lobule (SPL), Lingual Gyrus, Calcarine Gyrus, Thalamus</td>
<td>4716</td>
<td>11.44</td>
<td>10 -70 42</td>
</tr>
<tr>
<td>Inferior Occipital Gyrus (IOG)</td>
<td>L</td>
<td>IOG, Fusiform Gyrus (FG), Cerebellum (L/R)</td>
<td>3689</td>
<td>22.83</td>
<td>-20 -96 -6</td>
</tr>
<tr>
<td>Inferior Frontal Gyrus (IFG)</td>
<td>L</td>
<td>IFG (BA44/BA45), Precentral Gyrus, Insula</td>
<td>2942</td>
<td>15.40</td>
<td>-44 16 26</td>
</tr>
<tr>
<td>IFG</td>
<td>R</td>
<td>IFG (BA45/BA44), Insula</td>
<td>2740</td>
<td>14.52</td>
<td>48 22 24</td>
</tr>
<tr>
<td>IOG</td>
<td>R</td>
<td>IOG, FG, Cerebellum</td>
<td>2525</td>
<td>23.37</td>
<td>26 -94 -2</td>
</tr>
<tr>
<td>Supplementary Motor Area (SMA)</td>
<td>B</td>
<td>SMA</td>
<td>848</td>
<td>9.97</td>
<td>-4 14 50</td>
</tr>
<tr>
<td>Superior Temporal Gyrus (STG)</td>
<td>R</td>
<td>STG, Middle Temporal Gyrus (MTG), Inferior Parietal Lobule (IPL)</td>
<td>441</td>
<td>7.34</td>
<td>54 -42 12</td>
</tr>
<tr>
<td>IPL</td>
<td>L</td>
<td>IPL</td>
<td>438</td>
<td>10.21</td>
<td>-36 -58 44</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>R</td>
<td>Amygdala (LB/SF/CM), Hippocampus (CA)</td>
<td>406</td>
<td>8.75</td>
<td>22 -10 -12</td>
</tr>
<tr>
<td>Hippocampus</td>
<td>L</td>
<td>Amygdala (LB/SF/CM), Hippocampus (CA)</td>
<td>294</td>
<td>9.02</td>
<td>-20 -10 -12</td>
</tr>
<tr>
<td>Thalamus</td>
<td>L</td>
<td>Thalamus (visual, temporal, parietal), Hippocampus (CA)</td>
<td>141</td>
<td>6.24</td>
<td>-22 -30 -2</td>
</tr>
<tr>
<td>Angular Gyrus</td>
<td>R</td>
<td>IPL/Angular Gyrus</td>
<td>107</td>
<td>6.46</td>
<td>38 -58 44</td>
</tr>
<tr>
<td>Cerebellum</td>
<td>R</td>
<td>Cerebellum</td>
<td>55</td>
<td>6.89</td>
<td>34 -72 -50</td>
</tr>
<tr>
<td>MTG</td>
<td>L</td>
<td>MTG</td>
<td>41</td>
<td>5.54</td>
<td>-54 -46 8</td>
</tr>
<tr>
<td>White Matter (WM)</td>
<td>R</td>
<td>WM</td>
<td>25</td>
<td>5.47</td>
<td>8 0 26</td>
</tr>
<tr>
<td>Middle Frontal Gyrus (MFG)</td>
<td>R</td>
<td>MFG</td>
<td>16</td>
<td>5.21</td>
<td>40 52 18</td>
</tr>
</tbody>
</table>
Table 1 Continued

<table>
<thead>
<tr>
<th>Location 1</th>
<th>H</th>
<th>Cluster description 2</th>
<th>Cluster size</th>
<th>Peak T</th>
<th>MNI coordinates</th>
</tr>
</thead>
<tbody>
<tr>
<td>FG</td>
<td>R</td>
<td>FG, Hippocampus (CA), Amygdala (LB)</td>
<td>5</td>
<td>5.34</td>
<td>34 -8 -32</td>
</tr>
</tbody>
</table>

Significant clusters for the task effect (emotion – control) at FWE p < 0.05 (T=4.68). Peak T-values and MNI coordinates correspond to the first maximum of the activated cluster. H = hemisphere: L = left, R = right, B = bilateral. 1 This column reports the functional name or location that belongs to the peak voxel. 2 This column lists brain areas that are part of the cluster according to the Anatomy Toolbox.

7.4.4 Meta-analysis update

The effect size estimate for the association between 5-HTTLPR and amygdala activation in our study was $g = -0.14$, 95% CI -0.55, +0.27, $p = 0.51$. When we added our study to the 34 samples reported by Murphy and colleagues (2013), the pooled effect size estimate was no longer significant, $g = +0.20$, 95% CI -0.01, +0.41, $p = 0.06$, with evidence of substantial between-study heterogeneity ($I^2 = 70\%$). The effect sizes of the different studies plotted against study precision are shown in Figure 1. Asymmetry in the published studies (in the direction of excess low-accuracy studies with statistically significant effects) is suggestive of publication bias. Furthermore, a plot of year of publication against effect size is shown in Figure 2. The negative relationship reveals how the strength of evidence has declined over time.

Figure 1 Funnel plot displaying the effect size ($g$) of our study and the published and unpublished studies reported in the meta-analysis by Murphy and colleagues (2013) plotted against study precision. Asymmetry in the published studies (in the direction of excess low-accuracy studies with statistically significant effects) is suggestive of publication bias. The standard error (SE) and hence study precision (1/SE) is related to the sample size.
7.5 Discussion

Recent meta-analyses have suggested that effect size estimations of the association between amygdala activation and the 5-HTTLPR could be distorted by publication bias (Munafo et al., 2008; Murphy et al., 2013). Indeed, our funnel plot (see Figure 1) based on Murphy et al’s (2013) meta-analysis illustrates that high-accuracy studies are needed to improve the effect size estimate and that studies with null findings should also be taken into account. When we added our relatively high-accuracy study in 120 participants to the 29 published and 5 unpublished studies reported by Murphy and colleagues (2013), the pooled effect size was no longer significant (Hedge $g = +0.20$, 95% CI -0.01 - 0.41, $p = 0.06$). Moreover, the previously established pooled effect size estimate ($g = +0.35$; Murphy et al., 2013) did not fall within the confidence interval observed in our sample (95% CI -0.55 - 0.27). Together, these findings suggest that the association between the 5-HTTLPR and amygdala activation has not been replicated robustly and may thus be either much smaller than previously thought, non-existent or conditional on other factors.

Notably, the considerable between-study heterogeneity in effect sizes ($I^2 = 70\%$) suggests that the true effect size for the 5-HTTLPR-amygdala association may differ across study populations or designs (Murphy et al., 2013). For instance, tasks that require a cognitive appraisal of negative emotions or measure amygdala activation after exposure to an acute stressor may be more sensitive to genotype effects (Cousijn et al., 2010; Firk et al., 2013; Gillihan et al., 2010; Volman et al., 2013). Stratified analyses by Murphy and colleagues (2013), however, did not provide evidence for moderating roles of study design or population characteristics with potential exceptions for (poor) genotyping quality and (older) age. More and larger studies may be needed, however, to allow for
adequate analysis of relevant moderator variables.

In our study, allele frequencies did not deviate from HWE equilibrium. Moreover, participants were genotyped taking into account the relatively new insight that the Lg allele on the long form of the 5-HTTLPR-rs25531 combination has the same transcriptional activity as the S-allele (Wendland et al., 2006). This procedure most probably increased the difference in level of serotonin expression between our L/L and S-carrier groups, which made our study more sensitive for detecting differences in amygdala activation compared to earlier studies that did not type the Lg allele as a low expression variant (e.g. Canli et al., 2005; Hariri et al., 2002, 2005).

Concerning age, we cannot exclude the possibility that there is a 5-HTTLPR-amygdala association particularly in samples with older participants. The 5-HTTLPR may have a differential impact on the ageing brain (Pacheco et al., 2012), which is typically characterized by less amygdala activity and higher prefrontal activity in response to emotional stimuli (Roalf et al., 2011). However, a specific 5-HTTLPR-amygdala association in older age seems to be at odds with the idea of 5-HTTLPR as a risk factor for the development of affective disorders, which have the highest incidence rates earlier in life (Andrade et al., 2003; Bijl et al., 2002). At this point, the role of age in the expression of the 5-HTTLPR remains uncertain. Stratified analyses will become increasingly powerful in charting study characteristics that influence the magnitude of the observed effect size as the number of studies per stratum increases. At present, however, inadequate sample sizes are the main source of variation in study outcomes (Murphy et al., 2013). Indeed, even our sample size probably falls short to detect the effect size indicated in the previous meta-analysis (Murphy et al., 2013), even though it is large by the standards of the field. Nevertheless, it should be noted that the funnel plot indicates our effect size to be at the centre of the distribution, which raises confidence that this is a precise estimate and the putative 5-HTTLPR-association is less robust than previously thought. Whereas most studies of 5-HTTLPR variation and neural correlates of emotion processing have been focused on the amygdala, other brain regions such as the anterior cingulate cortex (and connectivity with emotion-related regions) could be of interest (Waring et al., 2013), an issue that awaits further investigation.

We can not preclude that the pendulum for the effect size estimate may swing back to the other side of p = 0.05 when future studies with proper sample sizes are added to the meta-analysis. The current collective state of the literature, however, casts doubt on previously reported substantial effects and calls into question simple gene–brain relationships. Regarding the latter, epistatic effects between 5-HTTLPR and other polymorphisms should be taken into account (e.g. Pezawas et al., 2008). Moreover, environmental influences could potentially modify or even mask assumed 5-HTTLPR-determined neural activity (Bogdan et al., 2013, but see Walsh et al., 2012 for a recent null finding). However, the challenges of (multivariate) gene x environment research (e.g. Duncan and Keller, 2011) will be even more pronounced when incorporated in the field of neuroscience in which "small, low-powered studies are endemic" (Button et al., 2013, p. 374).
There is growing concern about the reliability of scientific findings, in particular regarding 'hot topics' such as the 5-HTTLPR-amygdala activation association (Ioannidis, 2005; Park et al., 2014). Not only may there be a publication bias in favor of small studies with positive findings (Munafo et al., 2008; Murphy et al., 2013), but published studies may also make stronger claims than is warranted by the level of support in the reported data (Park et al., 2014). While the standard analytical procedures of meta-analyses reduce the influence that researcher's flexible analytical options can have on the presentation of results (Simmons et al., 2011), it will be impossible to make a reliable estimation of any association when published studies are only the tip of the iceberg. Meta-analyses should therefore be reiteratively updated with studies with proper sample sizes (irrespective of their effect sizes) to shed light on key findings in psychology.

7.6 Acknowledgements

The authors would like to thank A. Sibeijn-Kuiper and J. Streurman-Werdekker for their support in fMRI data acquisition, A.M. Brugman for DNA extraction and genotyping and L. Geerligs and H. Meffert for helpful discussions on the fMRI analyses.