Supplementary Document 1. Procedures for accounting for allele frequencies of unsampled individuals in calculating effect sizes for extreme samples.

Example: Hettema et al. (2008) sampled individuals from the top 20th and bottom 20th percentile of latent genetic risk for neuroticism. Data obtained from the first author included frequencies by COMT Val158Met genotype and neuroticism status. For instance, data from female participants in the Stage 1 sample (sample "A" in Table 1 of the manuscript) are:

	Met/Met	Val/Met	Val/Val
Top 20th percentile neuroticism	21	43	23
Bottom 20th percentile neuroticism	29	47	10

Total *N* = 173

Based on the total *N* of 173 for 40% of the original sample (top 20th + bottom 20th percentile), we calculated the original sample size of 432.5 (= 173 / 0.4). According to Palmetier et al. (1999), the genotype frequencies for Met/Met, Val/Met and Val/Val in Mixed European populations are 0.297, 0.527 and 0.176, respectively. Based on these genotype frequencies, we calculated the expected number of women by COMT Val158Met genotype in the Hettema et al. (2008) sample. For instance, the expected number of women with the Met/Met genotype was calculated as (432.5 x 0.297) = 128.5; see row 2 of the table below for other genotype *n*'s. In the next step, we calculated the expected number of individuals falling in the middle 60% (100% - top 20th percentile – bottom 20th %percentile) by subtracting the actual genotype *n* from the expected genotype *n*. For example, for Met/Met, *n* for the middle 60% individuals was calculated as (128.5 – 21 – 29 = 78.5); see row 3 for other genotypes. We also calculated the expected genotype *n* for the bottom 80th percentile of the neuroticism by adding the expected *n* for the middle 60% of the neuroticism distribution and the actual *n* for the bottom 20th% percentile (e.g., 29+78.5 = 107.5 for Met/Met); see row 4 below.

	Met/Met	Val/Met	Val/Val
Palmetier et al. (1999) genotype frequency	0.297	0.527	0.176
Expected genotype *n* in full sample of 433	128.45	227.93	76.12
Expected genotype *n* in middle 60% of distribution	78.45	137.93	43.12
Expected genotype *n* in bottom 80% of distribution	107.45	184.93	53.12

From the two tables, we rounded off the frequencies into integers and formed a table of frequencies by COMT Val158Met genotype and neuroticism status (top 20% vs. bottom 80%):

	Met/Met	Val/Met	Val/Val
Top 20th percentile neuroticism	21	43	23
Bottom 80th percentile neuroticism	107	185	53

To obtain the odds of high neuroticism (top 20%) given Met/Met genotype, relative to the odds of high neuroticism given Val/Val genotype, we first calculated the following table of cell probabilities:

	Met/Met	Val/Val
Top 20th percentile neuroticism	21 (p_{MM20}=0.4773)	23 (p_{VV20}=0.5227)
Bottom 80th percentile neuroticism	107 (p_{MM80}=0.6688)	53 (p_{VV80}=0.3313)

$$ES_{OR} = \frac{p_{MM20} * p_{VV80}}{p_{MM80} * p_{VV20}} = 0.4523$$

The log odds ratio was calculated as:

$$ES_{LOR} = ln\,(ES_{OR})$$

$$= ln\,(0.4523)$$

$$= -0.7934$$

To transform ES_{LOR} into an effect size comparable with effect sizes in mean differences metric, the Cox method was applied:

$$ES_{Cox} = ES_{LOR} / 1.65$$

$$= (-0.7934) / 1.65$$

$$= -0.4809$$

The sampling variance of the ES_{Cox} was calculated as follows:

$$SE_{Cox} = 0.367\left[\frac{1}{O_{MM20}} + \frac{1}{O_{MM80}} + \frac{1}{O_{VV20}} + \frac{1}{O_{VV80}}\right]$$

$$= 0.367\left[\frac{1}{21} + \frac{1}{107} + \frac{1}{23} + \frac{1}{53}\right]$$

$$= 0.0438$$

Supplementary Table 1. Effect size (standard error) of the association between COMT Val158Met polymorphism and any anxiety-related traits by genotypic contrast and sex for each study.

Study	Year	Sex	Measure (construct)	Ethnicity	Met/Met vs. Val/Val	Met/Met vs. Val/*	Met/* vs. Val/Val
Benjamin	2000	MF	TPQ (HA)	Unknown	-.041 (.138)	-.022 (.118)	-.034 (.100)
Henderson	2000	MF	Combined	Caucasian	-.012 (.096)	.009 (.077)	-.024 (.079)
		M			-.172 (.142)	-.125 (.112)	-.092 (.118)
		F			.123 (.130)	.113 (.106)	.053 (.107)
Eley	2003	F	NEO-FFI (N)	Caucasian	.128 (.282)	.318 (.208)	-.103 (.255)
Kim (A)	2004	MF	TCI (HA)	Caucasian	-.187 (.161)	-.246 (.128)	.013 (.130)
		M			-.253 (.229)	-.231 (.186)	-.095 (.183)
		F			-.160 (.227)	-.252 (.176)	.054 (.186)
Kim (B)	2004	F	TCI (HA)	Asian	.054 (.461)	.232 (.429)	-.223 (.264)
Tsai	2004	F	TPQ (HA)	Asian	-.075 (.425)	-.011 (.419)	-.009 (.186)
Harris	2005	MF	HADS (Anx)	Caucasian	.077 (.127)	-.017 (.099)	.108 (.106)
		M			.133 (.194)	.074 (.151)	.109 (.163)
		F			.042 (.168)	-.064 (.131)	.110 (.139)
Olsson	2005, 2007	MF	Combined	Caucasian	-.163 (.111)	-.115 (.089)	-.103 (.093)
		M			-.147 (.173)	-.068 (.140)	-.139 (.147)
		F			-.161 (.147)	-.128 (.117)	-.085 (.122)
Reuter	2005	MF	Combined	Caucasian	.192 (.147)	.144 (.120)	.110 (.121)
		M			.042 (.281)	-.028 (.222)	.084 (.237)
		F			.281 (.173)	.226 (.142)	.138 (.141)
Hoth	2006	MF	Combined	Caucasian	.041 (.144)	-.026 (.111)	.075 (.124)
		M			.103 (.190)	.008 (.148)	.131 (.164)
		F			-.036 (.222)	-.054 (.168)	-.006 (.189)
Ivanova	2006	MF	TCI (HA)	Caucasian	-.060 (.259)	-.013 (.227)	-.076 (.209)
		M			-.275 (.405)	-.221 (.355)	-.187 (.332)
		F			.133 (.338)	.171 (.297)	.004 (.268)
Kim	2006	MF	TCI (HA)	Asian	-.648 (.209)	-.481 (.200)	-.358 (.119)
		M			-.436 (.278)	-.336 (.267)	-.179 (.171)
		F			-.840 (.317)	-.625 (.303)	-.535 (.167)
Reuter	2006	MF	BIS (BI)	Caucasian	.136 (.157)	.095 (.132)	.100 (.130)
		M			.525 (.232)	.391 (.191)	.364 (.189)
		F			-.185 (.218)	-.161 (.183)	-.106 (.179)
Tochigi	2006	MF	NEO-PI-R (N)	Asian	-.073 (.211)	.011 (.201)	-.153 (.128)
Golimbet	2007	MF	Combined	Caucasian	.212 (.181)	.070 (.152)	.224 (.143)
		M			-.061 (.310)	-.092 (.246)	-.007 (.258)
		F			.403 (.227)	.209 (.193)	.367 (.174)
Hashimoto	2007	MF	TCI (HA)	Asian	.611 (.272)	.468 (.260)	.467 (.173)
		M			.658 (.491)	.523 (.476)	.570 (.306)
		F			.561 (.327)	.425 (.311)	.391 (.211)
Ishii	2007	MF	TCI (HA)	Asian	.051 (.174)	.083 (.169)	-.059 (.092)
		M			-.316 (.228)	-.245 (.219)	-.160 (.128)
		F			.714 (.276)	.671 (.269)	.098 (.132)
Urata	2007	F	NEO-FFI (N)	Asian	.152 (.283)	.116 (.272)	.093 (.158)
Bækken	2008	MF	HADS (Anx)	Caucasian	.024 (.066)	.013 (.048)	.019 (.059)
		M			-.103 (.107)	-.060 (.079)	-.076 (.094)
		F			.124 (.085)	.073 (.062)	.091 (.076)
Hettema (A)	2008	MF	EPQ-R-S (N)	Caucasian	-.296 (.148)	-.139 (.115)	-.243 (.125)
		M			-.101 (.213)	-.069 (.158)	-.064 (.183)
		F			-.481 (.209)	-.210 (.167)	-.414 (.173)

Supplementary Table 1 (continued)

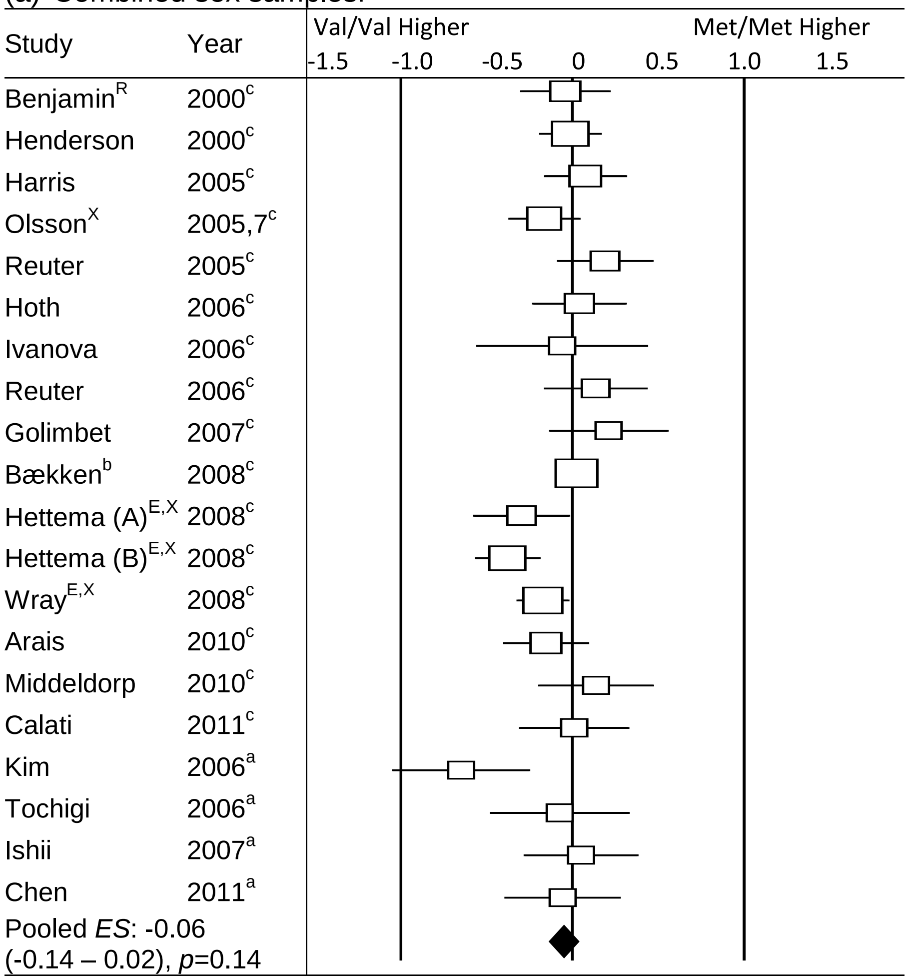
Study	**Year**	**Sex**	**Measure (construct)**	**Ethnicity**	**Met/Met vs. Val/Val**	**Met/Met vs. Val/***	**Met/* vs. Val/Val**
Hettema (B)	2008	MF	EPQ (N)	Caucasian	-.378 (.101)	-.128 (.079)	-.357 (.085)
		M			-.265 (.127)	-.085 (.098)	-.254 (.108)
		F			-.576 (.168)	-.206 (.134)	-.540 (.139)
Montag	2008	F	BIS (BI)	Caucasian	.564 (.255)	.535 (.220)	.303 (.218)
Wray	2008	MF	EPQ-R-S (N)	Caucasian	-.172 (.082)	.050 (.064)	-260 (.071)
		M			-.035 (.129)	.109 (.098)	-.136 (.114)
		F			-.265 (.107)	.008 (.086)	-.344 (.091)
Arias	2010	MF	STAI-T (Anx)	Caucasian	-.153 (.131)	-.205 (.113)	.008 (.102)
		M			-.188 (.195)	-.203 (.168)	-.037 (.153)
		F			-.128 (.176)	-.194 (.151)	.028 (.137)
Hatzimanolis	2010	F	EPQ (N)	Caucasian	.049 (.147)	.101 (.126)	-.043 (.113)
Middeldorp	2010	MF	YASR (Anx)	Caucasian	.138 (.175)	.182 (.154)	-.010 (.130)
		M			-.182 (.261)	-.075 (.239)	-.222 (.184)
		F			.357 (.244)	.341 (.203)	.076 (.188)
Calati	2011	MF	TCI (HA)	Caucasian	.011 (.167)	-.162 (.140)	.187 (.132)
		M			.336 (.255)	-.063 (.202)	.442 (.214)
		F			-.093 (.227)	-.180 (.197)	.061 (.169)
Chen	2011	MF	Combined	Asia	-.056 (.176)	-.062 (.172)	.005 (.085)
		M			-.628 (.282)	-.583 (.276)	-.205 (.129)
		F			.320 (.228)	.257 (.222)	.159 (.115)
Desmeules	2011	F	STAI-T (Anx)	Caucasian	.239 (.327)	.457 (.285)	-.119 (.251)

Notes:

1. When a study included multiple anxiety measures, the average effect size across measures was used for analysis.

Supplementary Figure 1. Forest plot of meta-analysis results for (a) combined sex samples; (b) male samples; (c) female samples on all anxiety-related measures; Met/Met vs. Val/Val contrast.

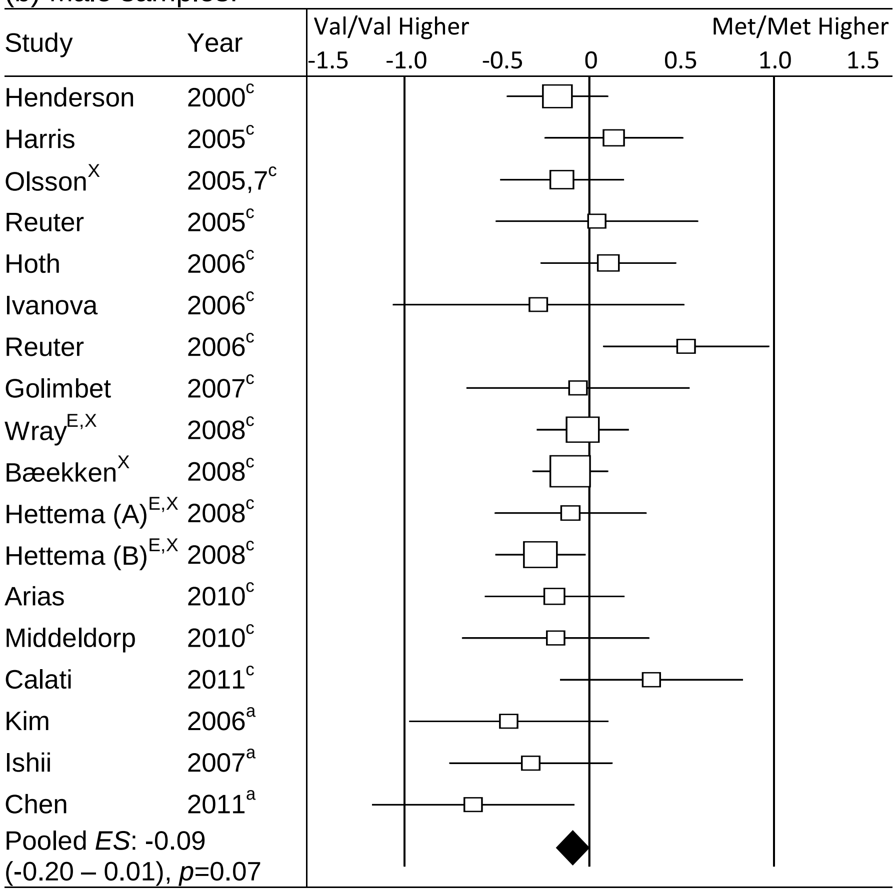
(a) Combined sex samples:



Notes for Supplementary Figure 1a-c:

1. Study subsequently removed in sensitivity analysis due to the following characteristics denoted by superscripts: [E] = extreme sampling, [X] = Cox method used to convert logged odds ratio effect sizes, and [R] = samples with related individuals.
2. Sample ethnicity is denoted by superscripts: [c] = Caucasian, [a] = Asian, [u] = unknown.
3. For Olsson et al. (2005, 2007), data on CIS-R (but not NEO) was removed from the set of sensitivity analysis that accounted for use of logged odds ratio effect sizes.

(b) Male samples:



(c) Female samples: