

Suzanne Vrshek-Schallhorn, Ph.D.
Gail M. Corneau, M.A.
Lisa R. Starr, Ph.D.

Department of Psychology, University of North Carolina at Greensboro, Greensboro (Vrshek-Schallhorn, Corneau); Department of Psychology, University of Rochester, Rochester, N.Y. (Starr).

Send correspondence to Dr. Vrshek-Schallhorn (suzanne.schallhorn@uncg.edu) and Dr. Starr (lisa.starr@rochester.edu).

The authors report no financial relationships with commercial interests.

Accepted May 6, 2019.

Am J Psychiatry 2019; 176:667–668; doi: 10.1176/appi.ajp.2019.19040374

Measurement Error Cannot Account for Failed Replications of Historic Candidate Gene-by-Environment Hypotheses: Response to Vrshek-Schallhorn et al.

TO THE EDITOR: Vrshek-Schallhorn et al. dispute our conclusion that historic candidate gene-by-environment (G×E) hypotheses were incorrect, criticizing the measures of environmental stressors we employed. We appreciate the opportunity to respond to their points.

We agree with their point that measures of environment are important, as we emphasized in our article (1). Our results demonstrate that both interpersonal and noninterpersonal environmental measures influence depression liability (see Table S6 in the online supplement to our article). The additive effects of these environmental measures replicate previous findings that a variety of stressors affect depression liability. However, Vrshek-Schallhorn et al. suggest that the measures we used are completely obscuring real candidate G×E effects. Below, we argue that this cannot be the case and that several of their comments are inaccurate.

They refer to the use of artificially dichotomized questionnaires, implying that we chose arbitrary cutoff points for continuous measures. In actuality, the binary stress measures we examined were inherently dichotomous, indicating whether or not participants endorsed one of a handful of events. For example, exposure to trauma in childhood was coded affirmatively if participants stated that they had been subject to sexual or physical abuse in childhood. On the other hand, we did not dichotomize the Townsend deprivation index, a continuous measure of socioeconomic hardship. In addition, Vrshek-Schallhorn et al. state that our worst-case-scenario measurement error simulations “vastly underestimate the amount of random error introduced by inadequate stress measures,” suggesting that a noisy stress questionnaire measure might account for as little as 16% of the variance of its adequately measured analogue. In response, we introduced this degree of error variance in both the depression and stress measures in our simulations. Even with this extreme degree of error, we still observed greater than 90% power to detect even modest interaction effects by candidate gene standards (odds ratio ≥ 1.26).

Systematic error can influence results, as Vrshek-Schallhorn et al. assert. However, the several severe systematic measurement error regimes we examined (see section S4.3.3 in the online supplement to our article) demonstrate that misclassification of environment measures cannot account for the lack of candidate gene or candidate G×E replication. In truth, we are unable to construct a plausible measurement error model that reconciles the validity of previously reported candidate gene findings with our observations that every stressor we examined evidenced substantial, highly significant effects on every depression measure (see section S6 in the online supplement to our article) but that no candidate gene polymorphism or stressor-by-polymorphism interaction had detectable effects, despite ~100% power across a broad array of measurement error scenarios. This is not to say that measurement error is unimportant. However, with respect to the large effects reported in the candidate gene literature in small samples, measurement error cannot account for the lack of support for historical candidate gene hypotheses in our study (1) or in other large, collaborative studies (2) that have investigated the genetic underpinnings of depression, even in carefully phenotyped studies specifically testing genome-wide G×E hypotheses (3). Instead, the most plausible explanation for these failures to replicate is that the original candidate gene findings were false positives.

Vrshek-Schallhorn et al. also state that we failed to distinguish between weakly correlated types of stress that may interact in different ways with genetic variants. The great majority of scenarios wherein either or both of these types of stress interact with a candidate genetic variant should still produce detectable genetic variant main effects—as noted, none were detected (see section S7 in the online supplement to our article). In the unlikely case of a complete crossover interaction where the variant has no main effect, any interactions detectable in candidate G×E samples as small as those cited by our critics would in turn induce detectable differences in variance across genotypes in large samples, even allowing for extreme measurement error. We found no evidence for such heteroscedasticity (results available upon request).

The trajectory of increasing complexity in candidate gene research follows a pattern of reactions to repeated replication failures. In the 1990s, it was hypothesized that specific, common polymorphisms within serotonergic and other neurotransmitter genes would explain substantial variation in depression liability. In the 2000s, it was hypothesized that moderation of genetic effects by environmental stressors would explain inconsistent main effect findings. Now, Vrshek-Schallhorn et al. suggest that our null findings can be explained by different types of stressors that produce interaction effects in opposite directions, or by catastrophic measurement error, and that their hypotheses cannot be adequately tested in existing well-powered samples. We instead suggest that these lines of inquiry are fundamentally flawed; neither the notion that common variants have large

effects on complex traits nor the relevance of these specific candidate polymorphisms to depression is supported by well-powered studies (1–5).

The genetic architectures of traits such as depression have revealed themselves to be hyperpolygenic—thousands of common variants influence liability, each exerting an effect far too small to be detected in conventional samples, even with precise phenotypic measurement. The era of using small samples to investigate specific common genetic variants based on historical biological hypotheses that are unsupported by genome-wide data is drawing to a close.

REFERENCES

1. Border R, Johnson EC, Evans LM, et al: No support for historical candidate gene or candidate gene-by-interaction hypotheses for major depression across multiple large samples. *Am J Psychiatry* 2019; 176:376–387
2. Culverhouse RC, Saccone NL, Horton AC, et al: Collaborative meta-analysis finds no evidence of a strong interaction between stress and 5-HTTLPR genotype contributing to the development of depression. *Mol Psychiatry* 2018; 23:133–142
3. Peterson RE, Cai N, Dahl AW, et al: Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression. *Am J Psychiatry* 2018; 175:545–554
4. CONVERGE consortium: Sparse whole-genome sequencing identifies two loci for major depressive disorder. *Nature* 2015; 523:588–591
5. Wray NR, Ripke S, Mattheisen M, et al: Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nat Genet* 2018; 50:668–681

Richard Border, M.A., M.S.
Emma C. Johnson, Ph.D.
Luke M. Evans, Ph.D.
Matthew C. Keller, Ph.D.

Institute for Behavioral Genetics (Border, Evans, Keller), Department of Psychology and Neuroscience (Border, Keller), and Department of Ecology and Evolutionary Biology (Evans), University of Colorado Boulder, Boulder; Department of Psychiatry, Washington University School of Medicine, St. Louis (Johnson).

Send correspondence to Mr. Border (richard.border@colorado.edu) and Dr. Keller (matthew.c.keller@gmail.com).

The authors' disclosures accompany the original article.

Accepted May 6, 2019.

Am J Psychiatry 2019; 176:668–669; doi: 10.1176/appi.ajp.2019.19040374r

Correction to Popova et al.

When the article “Efficacy and Safety of Flexibly Dosed Esketamine Nasal Spray Combined With a Newly Initiated Oral Antidepressant in Treatment-Resistant Depression: A Randomized Double-Blind Active-Controlled Study,” by Vanina Popova et al. (doi: 10.1176/appi.ajp.2019.19020172) was published online on May 21, 2019, and in the June 2019 issue, one of the percentages in Table 1 was misreported. The percentage of participants in the antidepressant plus placebo group who had been treated with three or more antidepressant medications should have been listed as 33.9%. This correction was made and the article was reposted on July 11, 2019.

Corrections to Narendran et al.

After the article “Nociceptin Receptors Upregulated in Cocaine Use Disorder: A Positron Emission Tomography Imaging Study Using [¹¹C]NOP-1A,” by Rajesh Narendran, M.D., et al. (doi: 10.1176/appi.ajp.2019.18081007) was published online on May 6, 2019, and appeared in the June 2019 issue, the authors discovered an error in matching healthy control to cocaine use disorder subjects that led to data from one of the healthy control subjects being entered twice. Calculations were corrected to contrast 24 cocaine use disorder and 25 (i.e., not 26) healthy control subjects. The primary result and conclusions hold with the revised analyses. The article was reposted on July 17, 2019 with revisions to text, tables, figures, and statistical analyses to reflect the corrected healthy control group N of 25. There are some changes of note for secondary results: in Table 1, the between-group difference in education is not significant at $p < 0.05$; in the linear mixed-model analysis of regional [¹¹C]NOP-1A VT, the effect of diagnosis falls short of significance when tobacco use status is included as an additional factor; unpaired *t* tests comparing [¹¹C]NOP-1A distribution volume were significant in four, not six, of the 11 regions of interest and the midbrain comparison did not survive the false discovery rate correction; the negative relationships between age and VT in the medial prefrontal cortex and dorso-lateral prefrontal cortex in the healthy control group do not reach significance.