

Doubting the Efficacy/Effectiveness of Electroencephalographic Neurofeedback in Treating Children With Attention-Deficit/Hyperactivity Disorder Is As Yet Unjustified

To the Editor: Van Dongen-Boomsma et al¹ assessed the efficacy of electroencephalographic (EEG) neurofeedback in children with attention-deficit/hyperactivity disorder by comparing an experimental group (n = 22) receiving EEG neurofeedback to a placebo group (inactive neurofeedback) (n = 19). Results did not show significant differences between the 2 groups. The authors report that, post hoc, the “sample had 80% power to detect a treatment effect of 0.90” and that “it is unlikely that [the] negative results were due to limited statistical power,”^{1(p826)} concluding that their study sheds serious doubt on the effectiveness of EEG neurofeedback as a way of treating children with attention-deficit/hyperactivity disorder.

First, we want to acknowledge the robust design and methodology of this study. However, we are not sure what the authors meant by stating in their discussion that “the sample had 80% power to detect a treatment effect of 0.90,” and we cannot find an explanation of the rationale behind such a statement. How would statistical power be an issue if the effect size was as high as 0.90 in a sample of 41 participants? We ran a post hoc power analysis, using G*Power Version 3.1.5,^{2,3} for a sample size of 41 participants, α of 5% and an effect size of 0.90 resulting from repeated-measures analysis of variance within-group and between-groups interaction tests for 2 groups and 2 measuring times, and found that this sample would have achieved 100% power to detect a treatment effect size of 0.90, not 80% as reported by the authors. Therefore, there is a 100% chance of finding statistically significant differences ($P \leq .05$) between the EEG neurofeedback group and the placebo group,⁴ although none were found. Impossible, unless the effect size equaled zero. We suspect that the authors made a calculation error.

To verify this, we first calculated η^2 effect size for all the comparisons in the group \times time effect (interaction) using the following formula⁵: $(F \times df_1) / [(F \times df_1) + df_2]$. We found that $0.13\% \leq \eta^2 \leq 5.48\%$ (Table 1). Then, with these effect sizes, we ran a post hoc power analysis for each comparison, following the same procedure, using G*Power Version 3.1.5.^{2,3} Our calculations show that the achieved power varies between 5% and 11% (Table 1). On the basis of these findings, the risk of a type II error in this study varies between 89% and 95%. Hence, the authors' conclusion that it is unlikely that the negative results were due to a lack of statistical power is conceptually wrong.

We read van Dongen-Boomsma and colleagues¹ study with great interest, and, given the η^2 effect sizes we calculated, we agree that the differences between the effect of EEG neurofeedback and the placebo neurofeedback are small, and we believe that they probably would not be much higher in a larger sample if inclusion criteria were based on the same demographic and clinical characteristics. Nonetheless, given the lack of statistical power we showed above, it is not methodologically sound to describe the study as one with negative results. Clarification from the authors of the rationale and what they meant in saying that “the sample had 80% power to detect a treatment effect of 0.90” would be useful. Further research with sound statistical power and validity is needed to corroborate van Dongen-Boomsma and colleagues' conclusion doubting the efficacy/effectiveness of EEG neurofeedback, which has so far been demonstrated in other methodologically sound studies.^{6,7}

REFERENCES

1. van Dongen-Boomsma M, Vollebregt MA, Slaats-Willems D, et al. A randomized placebo-controlled trial of electroencephalographic (EEG) neurofeedback in children with attention-deficit/hyperactivity disorder. *J Clin Psychiatry*. 2013;74(8):821–827.
2. Faul F, Erdfelder E, Lang A-G, et al. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175–191.
3. Faul F, Erdfelder E, Buchner A, et al. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods*. 2009;41(4):1149–1160.
4. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Hillsdale, NJ: Lawrence Erlbaum; 1988.
5. Guilford JP. *Fundamental Statistics in Psychology and Education*. New York, NY: McGraw-Hill; 1965.
6. Arns M, de Ridder S, Strehl U, et al. Efficacy of neurofeedback treatment in ADHD: the effects on inattention, impulsivity and hyperactivity: a meta-analysis. *Clin EEG Neurosci*. 2009;40(3):180–189.
7. Gevensleben H, Holl B, Albrecht B, et al. Is neurofeedback an efficacious treatment for ADHD? a randomised controlled clinical trial. *J Child Psychol Psychiatry*. 2009;50(7):780–789.

Emmanuelle Dagenais, BSc
Ariane Leroux-Boudreault, BSc
Ghassan El-Baalbaki, PhD
 el-baalbaki.ghassan@uqam.ca
Jean Bégin, PhD

Author affiliations: Department of Psychology, University of Quebec (all authors), and Faculty of Medicine, McGill University (Dr El-Baalbaki), Montreal, Quebec, Canada.

Potential conflicts of interest: None reported.

Funding/support: None reported.

J Clin Psychiatry 2014;75(00):000–000 (doi:10.4088/JCP.14lr09043).

© Copyright 2014 Physicians Postgraduate Press, Inc.

Table 1. Effect Sizes and Post Hoc Power Calculation

Measure	Group \times Time Effect			Effect Size η^2 (%)	Achieved Power (%)
	F	df	P Value		
ADHD-RS-IV investigator-rated score					
Total symptoms	0.36	1,39	.554	0.91	5
Inattention symptoms	0.17	1,39	.682	0.43	5
Hyperactivity/impulsivity symptoms	2.26	1,39	.141	5.48	11
ADHD-RS-IV teacher-rated score					
Total symptoms	0.45	1,37	.509	1.20	5
Inattention symptoms	0.25	1,37	.624	0.67	5
Hyperactivity/impulsivity symptoms	0.53	1,37	.473	1.41	5
CGI-I score			.092		
CGAS score	1.96	1,38	.169	4.90	9
SDQ score	0.05	1,37	.818	0.13	5
PSERS score	0.10	1,39	.754	0.26	5

Abbreviations: ADHD-RS-IV = ADHD Rating Scale IV, CGAS = Children's Global Assessment Scale, CGI-I = Clinical Global Impressions-Improvement scale, PSERS = Pittsburgh Side Effects Rating Scale, SDQ = Sleep Disorders Questionnaire.