

Effectiveness of a Neuro-Play Therapy-Based Intervention on Working Memory, Self-Control, and Cognitive Flexibility in Children with Learning Disabilities

Soraya. Hamidi¹, Mojtaba. Tamadoni^{2*}, Afsaneh. Boostan³

¹ Department of Counseling, Ki.C., Islamic Azad University, Kish, Iran

² Department of Counseling, ST.C., Islamic Azad University, Tehran, Iran

³ Department of Counseling, Ro.C., Islamic Azad University, Roudehen, Iran

* Corresponding author email address: tamadonim336@gmail.com

E d i t o r	R e v i e w e r s
Seyyed Mohsen Asgharinekah Associate Professor, Department of Counseling and Educational Psychology, Faculty of Education and Psychology, Ferdowsi University of Mashhad, Mashhad, Iran asghari-n@um.ac.ir	Reviewer 1: Roodabeh Hooshmandi Department of Psychology and Counseling, KMAN Research Institute, Richmond Hill, Ontario, Canada. Email: roodhooshmandi@kmanresce.ca Reviewer 2: Thseen Nazir Professor of Psychology and Counseling Department, Ibn Haldun University, Istanbul, Turkey. Email: thseen.nazir@ihu.edu.tr

1. Round 1

1.1. Reviewer 1

Reviewer:

The exclusion criterion referring to “severe psychiatric disorders such as autism spectrum disorder, schizophrenia, or bipolar disorder” requires conceptual reconsideration and clearer reporting. Autism spectrum disorder is classified as a neurodevelopmental disorder rather than a psychiatric disorder in the same category as schizophrenia or bipolar disorder. The authors should revise this wording and explain how comorbid conditions were assessed. More importantly, common comorbidities among children with specific learning disabilities—particularly ADHD, anxiety disorders, language disorders, and developmental coordination disorder—should be addressed because these conditions are strongly related to executive functioning and could confound outcomes involving self-control, working memory, and cognitive flexibility.

The description of the “Daneman and Carpenter Working Memory Capacity Test” should be revised for developmental appropriateness. The original Daneman and Carpenter paradigm was developed within reading-span research, whereas the present sample consists of children aged 7–11 years with learning disabilities. The authors should explain whether the

instrument was adapted for children, whether the sentences were developmentally and linguistically appropriate for this age range, and whether a validated Persian child version was used. Furthermore, because children with reading or language-related learning disabilities may perform poorly on sentence-based working-memory tasks partly because of linguistic processing deficits, the authors should discuss whether the measure provides a sufficiently independent assessment of working memory rather than conflating working-memory capacity with language competence.

The Self-Control Scale description raises a potentially major measurement-validity concern. The manuscript states that the “Tangney Self-Control Scale” was administered to children aged 7–11 years, yet this instrument is conventionally a self-report scale developed primarily for older respondents. The authors must clarify who completed the scale—the child, parent, therapist, or another informant—and provide evidence that the selected version has been validated for children within this age range. If children themselves completed the instrument, the authors should explain how item comprehension was ensured, particularly among participants with learning disabilities. Without age-appropriate validity evidence, conclusions regarding intervention-related changes in self-control may be questionable.

The manuscript reports “Box’s $M = 120.67$, $F(90, 8900.78) = 1.03$, $p = .40$ ” as evidence for equality of covariance matrices. The authors should verify this result carefully against the SPSS output and clarify why Box’s M was included in what is described as a simple mixed repeated-measures ANOVA. In addition, because Box’s M is highly sensitive to sample size and multivariate normality, the manuscript should not treat a nonsignificant result as proof that matrices are identical. The assumptions paragraph would be stronger if it also addressed univariate and multivariate outliers, normality of model residuals, independence of observations, and whether any missing values were present.

Table 2 combines time effects, time $\times$ group interactions, and between-subject group effects for three outcomes, which is efficient, but the interpretation should emphasize that the time $\times$ group interaction is the principal evidence of differential intervention effects. At present, the discussion repeatedly interprets the “main effect of time” as evidence of treatment efficacy. A significant main effect of time only demonstrates that scores changed across measurement occasions when groups are considered collectively; it does not establish that the experimental treatment caused the change. The Results and Discussion should therefore prioritize the interaction effects and subsequent simple-effects or pairwise comparisons rather than using significant time effects as primary evidence for efficacy.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

A similar concern applies to the “Dennis and Vander Wal Cognitive Flexibility Inventory,” described as a 20-item self-report measure administered to children aged 7–11 years. The Cognitive Flexibility Inventory was originally designed to assess perceived cognitive flexibility and alternative thinking rather than performance-based executive set shifting, and its developmental appropriateness for young children must be demonstrated. The authors should provide evidence of validation in comparable child populations or explain any adaptation procedure. They should also distinguish clearly between self-perceived cognitive flexibility measured by questionnaire and executive cognitive flexibility measured by neuropsychological tasks such as the Wisconsin Card Sorting Test, Trail Making Test, or dimensional change paradigms. The current manuscript sometimes interprets questionnaire scores as direct evidence of neurocognitive change, which exceeds what this instrument alone can establish.

The intervention description contains an internal inconsistency regarding the number of sessions. One section reports that “the intervention was administered...in 11 active therapeutic-training sessions,” whereas the supplied protocol includes a twelfth session described as “summary and termination.” This discrepancy must be resolved throughout the title, abstract, Methods, intervention description, and any protocol reporting. If 12 sessions were actually delivered, the manuscript should state 12 sessions consistently. If the twelfth session did not constitute an intervention session, the authors should clearly explain its status, duration, content, and whether outcome assessments occurred before or after it. Protocol fidelity cannot be evaluated reliably while session dosage remains ambiguous.

The intervention protocol requires substantially more information regarding standardization and treatment fidelity. Although individual activities such as “animal puzzles,” “direct and reverse sit-down/stand-up,” “colored cards,” and “object-memory games” are described, it is unclear whether every child received the same tasks, duration, difficulty progression, therapist prompts, reinforcement schedule, and performance criteria. The manuscript should state who delivered the intervention, the therapist's professional qualifications and training, group size per session, whether a treatment manual was followed, and whether adherence was independently monitored. A fidelity checklist, supervision procedure, or percentage adherence to protocol would considerably strengthen confidence that observed outcomes resulted from a reproducible intervention rather than therapist-specific variation.

The control condition is described as receiving “no neuro-play therapy, cognitive rehabilitation, or comparable psychological intervention during the corresponding period.” This passive control design limits causal interpretation because the experimental group additionally received therapist attention, peer interaction, structured activities, novelty, and approximately two hours of engagement twice weekly. The authors should acknowledge that nonspecific treatment effects cannot be separated from the specific effects of neuro-play therapy. Ideally, an active control group receiving an equivalent amount of structured non-neurocognitive activity would provide a stronger test. At minimum, this issue should be explicitly included in the limitations and the discussion should avoid attributing all observed improvements specifically to neuroplastic or neurocognitive mechanisms.

The statistical analysis contains a major issue in the reported degrees of freedom. For a design with two groups and three repeated measurement occasions, the manuscript reports the time $\times$ group interaction as “F(4, 114),” while the main time effect is reported as F(2, 114). For a conventional 2×3 mixed ANOVA, the numerator degrees of freedom for a two-group time $\times$ group interaction would ordinarily correspond to the time factor multiplied by group df and would therefore require verification. Similarly, between-subject effects are reported as “F(2, 57)” despite there being only two groups, which ordinarily implies one numerator degree of freedom for group. The authors should return to the original SPSS output and verify all reported degrees of freedom, F statistics, and effect sizes. These apparent inconsistencies are sufficiently important that all inferential tables should be reconstructed directly from the statistical output before publication.

The statement that “Mauchly's test of sphericity and Levene's test of equality of error variances supported the relevant assumptions” is too general and does not provide the statistics needed to verify those assumptions. Mauchly's W, chi-square, degrees of freedom, and p value should be reported for each dependent variable. If sphericity was violated for any outcome, Greenhouse–Geisser or Huynh–Feldt corrected degrees of freedom and p values should be used. Similarly, Levene's test should ideally be reported for each outcome at each measurement occasion. Merely stating that assumptions were “confirmed” is insufficient, particularly because the reported repeated-measures degrees of freedom appear unusual and may reflect corrections or model-specification issues that are presently unexplained.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor's decision: Accepted.

Editor in Chief's decision: Accepted.