Journal Website

Article history:
Received 13 July 2026
Revised 12 September 2026
Accepted 04 October 2026
Initial Publication 27 October 2026
Final Publication 01 December 2026

Psychological Research in Individuals with Exceptional Needs

OPEN PEER-REVIEW REPORT

E-ISSN: 3060-6713

Effectiveness of a Peer-Mediated Intervention on Social Skills in Children with Autism Spectrum Disorder

Maedeh. Mirzaei Malati[1*], Razieh. Ghodrati Dizbani[2], Maryam. Tahmasebi[3], Maryam. Samandari Matouf[4]

[1] Department of Psychology, Ra.C., Islamic Azad University, Rasht, Iran
[2] Department of Health Psychology, Ra.C., Islamic Azad University, Rasht, Iran
[3] Department of Clinical Psychology, La.C., Islamic Azad University, Lahijan, Iran
[4] Department of Psychology, La.C., Islamic Azad University, Lahijan, Iran

*** Corresponding author email address**: maedeh.mirzaeimalati@iau.ac.ir

Editor	Reviewers
Shahrooz Nemati Professor, Department of Educational Sciences, Faculty of Educational Science and Psychology, University of Tabriz, Iran Sh.Nemati@Tabrizu.ac.ir	**Reviewer 1:** Hooman Namvar Assisstant Professor, Department of Psychology, Saveh Branch, Islamic Azad University, Saveh, Iran. Email: hnamvar@iau-saveh.ac.ir **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 participant description requires clarification because the manuscript states that the sample consisted of “six-year-old children,” whereas the Findings section reports mean ages of 6.20 ± 0.75 years and 6.35 ± 0.82 years. These standard deviations imply a fairly wide age distribution and are difficult to reconcile with a strict inclusion criterion of exactly 6 years of age. The authors should report age in months or provide the actual age range for each group and clarify whether “6 years” referred to a broader developmental interval, such as 6.00–6.99 years, or whether children outside that range were included. This issue is important because developmental variation in language, play, and social cognition can materially affect response to a peer-mediated intervention.

In the Findings paragraph, the sentence “The two groups were homogeneous in terms of age and gender ($p > .05$)” should be revised and supported with actual statistical information. The manuscript should report the gender distribution in each group, the statistical test used for age comparison, the test used for gender distribution, test statistics, degrees of freedom where

applicable, and exact p values. Moreover, nonsignificant baseline tests should not be interpreted as proof that groups are "homogeneous," particularly with only 15 participants per group, because such analyses have low power to detect potentially meaningful baseline imbalances. Descriptive baseline comparability is preferable, supplemented by standardized differences if possible.

The manuscript should clarify the reported effect-size statistic. Table 2 labels the effect size as "η^2," but SPSS repeated-measures procedures commonly report partial eta squared (ηp^2), not classical eta squared. These quantities are not interchangeable and can differ substantially. The values .45, .35, and .33 are relatively large and therefore materially influence interpretation of intervention effectiveness. The authors should verify the original output, identify whether these values represent eta squared or partial eta squared, use the appropriate notation consistently, and preferably provide confidence intervals for effect sizes. Interpretation should also distinguish statistical significance from the magnitude and clinical importance of the intervention effect.

The "Power" values of .85, .82, and .80 in Table 2 require explanation and should likely be reconsidered. If these are observed/post hoc power estimates generated from the obtained effect sizes, they provide little information beyond the corresponding p values and are generally not recommended as evidence of study adequacy. The important power analysis is the a priori calculation used in study planning. The authors should clarify whether the reported values represent observed power, prospective power, or another metric. If they are post hoc power estimates, removing this column and instead reporting confidence intervals for effect sizes would provide more informative evidence about statistical precision.

The Bonferroni pairwise-comparison results contain a sign inconsistency that should be corrected. Table 1 shows that the experimental-group total score decreased from 82.50 at pretest to 57.70 at posttest. If the contrast is explicitly defined as "Pretest – Posttest," the mean difference should be +24.80, not -24.80. Likewise, "Pretest – Follow-up" should be approximately +23.30 based on the reported means, whereas the table gives -23.30. In contrast, "Posttest – Follow-up" correctly yields approximately -1.50. The authors should verify the ordering used by SPSS and either reverse the labels or correct the signs. This is not merely cosmetic, because the direction of the effect is central to the interpretation of an outcome for which lower scores indicate improvement.

The Bonferroni analysis should be described more precisely because the manuscript states that pairwise comparisons were conducted "in the experimental group," yet the repeated-measures ANOVA evaluates a Group × Time interaction. The authors should explain whether the pairwise tests were estimated marginal mean comparisons stratified by group, whether separate within-group repeated-measures analyses were performed, or whether simple effects of time were examined following the significant interaction. Ideally, post hoc analysis should directly decompose the significant Time × Group interaction, including between-group comparisons at posttest and follow-up or appropriately adjusted simple effects. Reporting only within-experimental-group change does not fully establish that improvement was attributable to the intervention rather than to temporal change.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

The manuscript requires substantially greater detail about autism diagnosis and eligibility determination. The Study Design paragraph refers to "a confirmed diagnosis of autism spectrum disorder based on a specialist clinical evaluation and documented diagnostic records," while the GARS-3 was used for diagnostic screening. The authors should specify which professional discipline established the ASD diagnosis, which diagnostic criteria were applied, whether DSM-5-TR criteria were used prospectively or diagnoses were accepted from existing files, and whether GARS-3 scores were used as an inclusion threshold. The exact GARS-3 Autism Index criterion, if any, should be reported. Without this information, it is unclear whether the sample represents clinically diagnosed ASD, screening-positive children, or both.

The manuscript repeatedly describes the primary outcome as "social skills," although the principal instrument is the SRS-2. For example, Table 1 labels the total score as "Total Social Skills Score." This terminology should be reconsidered. The SRS-

E-ISSN: 3060-6713

2 primarily quantifies autism-related social impairment/social responsiveness and includes Restricted Interests and Repetitive Behavior, which is not itself a social-skill construct. Therefore, labeling the SRS-2 total score as a direct measure of "social skills" risks construct overinterpretation. The authors should either reformulate the primary outcome as social responsiveness/social impairment or provide a strong conceptual justification for operationalizing "social skills" using the SRS-2. The title, objectives, hypotheses, Results, and Discussion should use terminology consistent with the measured construct.

There are internal numerical inconsistencies in Table 1 that should be corrected before publication. In the experimental group at posttest, the reported component means sum to 57.20 (7.40 + 9.80 + 17.20 + 8.90 + 13.90), whereas the reported total mean is 57.70. At follow-up, the component means sum to 59.00, whereas the total mean is reported as 59.20. In the control group at follow-up, the component means sum to 81.55, whereas the reported total is 81.40. The authors should verify whether these differences arise from transcription, rounding, use of independently computed raw totals, or another scoring procedure. If total scores are calculated independently from item-level data, this should be explained; otherwise, the numerical values should be corrected.

The statistical assumptions paragraph contains an important inconsistency between the reported Mauchly's test and the degrees of freedom shown in the repeated-measures ANOVA table. The manuscript states that "The assumption of sphericity was also confirmed using Mauchly's test … ($p = .082$)." With three measurement occasions, an uncorrected within-subject effect would normally have a numerator degree of freedom of 2.00. However, Table 2 reports 1.40 for the time and Time × Group effects, suggesting application of a Greenhouse–Geisser or similar correction with an epsilon of approximately .70. If sphericity was not violated, uncorrected degrees of freedom should generally be reported. If a correction was applied, the authors should report Mauchly's W, the correction used, epsilon, corrected denominator degrees of freedom, and the rationale for using corrected statistics.

Table 2 contains an apparent arithmetic inconsistency in the Time effect. The table reports a sum of squares of 32.74, df = 1.40, and a mean square of 53.09. These values are mathematically incompatible because mean square should equal sum of squares divided by degrees of freedom; 32.74/1.40 is approximately 23.39, not 53.09. By contrast, the Time × Group row is internally consistent because 43.61/1.40 is approximately 31.15. The authors should return to the original SPSS output and reconstruct the entire ANOVA table directly from the software output, including error terms, denominator degrees of freedom, and corrected statistics. This is a major issue because transcription errors in the inferential table undermine confidence in the reported significance tests.

The repeated-measures ANOVA table is incomplete for scientific interpretation because denominator degrees of freedom and the corresponding error terms are not reported. For example, the statement "$F = 23.28$, $p < .001$" is insufficient without specification of F(df1, df2). The authors should report full APA-formatted statistics, such as F(2, 56) or the appropriate corrected values, for time, group, and Time × Group. The analysis table should also identify the relevant within-subject and between-subject error terms. This is particularly important in a small-sample repeated-measures design because the degrees of freedom provide essential information about model specification, corrections, and the validity of the reported tests.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor's decision: Accepted.
Editor in Chief's decision: Accepted.

E-ISSN: 3060-6713