

Development of a Skills Training Educational Package for Children with Autism Spectrum Disorder Based on Theory of Mind and Iranian Cultural Characteristics

Mina. Nirookar¹, Eghbal. Zarei^{2*}, Kobra. Hajializadeh³

¹ Ph.D. Student, Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran

² Professor, Department of Psychology and Counseling, Faculty of Humanities, University of Hormozgan, Bandar Abbas, Iran

³ Department of Psychology, BA.C., Islamic Azad University, Bandar Abbas, Iran

* Corresponding author email address: e.zarei@hormozgan.ac.ir

E d i t o r	R e v i e w e r s
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 [✉] Assistant 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:

In the Methods paragraph beginning with “This study was conducted with the ultimate aim of developing the conceptual framework and preliminary content of the educational package,” the methodological identity of the study is insufficiently resolved. The manuscript simultaneously describes the study as qualitative-exploratory, grounded theory, qualitative content analysis, and a mixed qualitative–quantitative validation process. These are not interchangeable methodological labels. If grounded theory was the principal approach, the authors should identify the specific grounded-theory tradition used, explain how theoretical sampling was implemented, describe constant comparison, memo writing, category saturation, and the process by which a core category was selected. If qualitative content analysis was instead the primary method, then terminology such as “paradigmatic model,” “causal conditions,” and “central phenomenon” requires methodological justification. The manuscript should therefore state one coherent analytic framework and align the sampling, coding, reporting, and quality criteria accordingly.

The claim that sampling continued “until theoretical saturation was reached” requires considerably more methodological evidence. The manuscript does not explain how saturation was operationalized, at what interview it was first observed, how many additional interviews were conducted to verify it, or whether saturation was assessed at the code, category, or theoretical level. For a grounded-theory study, simply stating that saturation occurred is insufficient. The authors should report, for example, that no new categories emerged after interview X and that interviews X+1 through X+n were conducted to confirm saturation. A saturation table or brief narrative audit trail would substantially improve methodological transparency and justify the final sample size.

The participant description is inadequate for evaluating the credibility and transferability of the qualitative findings. The Methods identify participants only broadly as clinical psychologists, child psychiatrists, occupational therapists, speech therapists, and educational instructors with at least five years of experience. The manuscript should report the exact number of participants in each professional group, sex distribution, age range or mean age, academic qualification, years of autism-related experience, practice setting, and geographic location. Because this study specifically claims cultural adaptation to the Iranian context, it is also important to indicate whether experts represented multiple regions and service environments or were concentrated in a single city or institutional network. Without these characteristics, readers cannot adequately judge the breadth or representativeness of the expert panel.

The interview procedure requires substantially greater detail. The Methods state that semi-structured interviews were used and that the process involved planning, explanation of objectives, asking core questions, and analysis, but the actual interview guide is not presented. The authors should specify the number of questions, provide examples of core questions and probes, indicate the average interview duration, describe whether interviews were conducted face-to-face or remotely, identify who conducted them and that researcher’s professional relationship to participants, and explain whether the interview guide was pilot-tested. These details are essential for reproducibility and for assessing whether the resulting categories emerged from sufficiently open inquiry rather than from highly leading questions based on the researchers’ pre-existing Theory of Mind framework.

The manuscript should provide a rigorous qualitative trustworthiness framework rather than relying primarily on face validity, content validity, and Cronbach’s alpha. In the Methods, the authors state that scientific rigor was ensured through “face and content validity” and internal consistency reliability, but these criteria are principally associated with measurement instruments rather than with qualitative analysis. For the interview component, the authors should address credibility, dependability, confirmability, and transferability through procedures such as member checking, peer debriefing, independent coding, audit trails, reflexive documentation, negative-case analysis, or triangulation. If multiple coders were involved, inter-coder agreement or consensus procedures should be reported. If only one researcher coded all transcripts, this should be explicitly acknowledged and mitigated through independent review of a sample of transcripts.

The sentence in the Methods stating that “Cronbach’s alpha was used to assess the internal consistency reliability of the questions” is incomplete because no actual reliability coefficient is subsequently reported in the Findings. If a questionnaire or checklist was quantitatively validated, the manuscript must identify that instrument, specify the number of items, response scale, scoring method, sample used for reliability analysis, and report the resulting alpha coefficient or coefficients. If Cronbach’s alpha was not actually calculated, this statement should be removed. At present, mentioning the method without reporting its numerical result creates the impression of an incomplete quantitative validation phase.

The manuscript similarly states that face and content validity of the educational package and questionnaires were evaluated by experts, but no quantitative or qualitative validity indices are provided. If formal content validation was undertaken, the authors should report the number of raters and the method used, such as Content Validity Ratio, Content Validity Index, item-level CVI, scale-level CVI, or another established procedure. If validity was assessed qualitatively, the authors should describe the criteria used by experts, the revisions made in response to their comments, and how disagreement was resolved. Given that development and validation of the package is the principal contribution of the article, the validation phase should be one of the most thoroughly reported aspects of the manuscript rather than a brief methodological statement.

The manuscript should distinguish much more carefully between “expected outcomes,” “clinical indicators,” and “observed outcomes.” In the grounded-theory findings, outcomes include increased eye contact, initiation with peers, improved turn-taking, increased empathy, improved Theory of Mind, and reductions in stereotyped behavior. These are expert-proposed indicators of successful intervention, not outcomes observed in treated children. Yet later narrative passages sometimes

describe them as results of the package. Every table and paragraph should explicitly identify these as anticipated or theoretically expected outcomes unless empirical intervention data are added. This distinction is essential to prevent readers from interpreting the paper as an effectiveness study.

The item “reduction in stereotyped behavior, for example from 10 to 3 occurrences” is particularly problematic because it appears numerically specific despite not being based on intervention data. If this example was directly provided by an interview participant, it should be clearly marked as an illustrative expert quotation or hypothetical clinical benchmark rather than presented as a study result. Otherwise, the numeric example should be removed. Quantitative-looking values embedded within qualitative findings can easily be misread as measured treatment effects, especially when the manuscript elsewhere uses the phrase “significant reduction.”

The Discussion paragraph asserting that increased understanding of social situations “reduces uncertainty and therefore decreases repetitive behavior” should be substantially moderated. The cited literature may support associations among emotion regulation, social communication, and repetitive behaviors, but the present study did not test mediation by environmental predictability. The authors should separate three levels of inference: what experts reported, what previous research has demonstrated, and what the authors propose theoretically. A suitable formulation would be that experts perceived greater predictability as a potential mechanism through which ToM-oriented intervention might influence repetitive behavior, and that this proposition is compatible with prior correlational findings, but requires direct empirical testing.

Authors revised the manuscript and uploaded the document.

1.2. Reviewer 2

Reviewer:

Table 1, “Manifestations of Theory of Mind Deficits in Everyday Situations and Social Interactions,” is clinically informative, but the boundary between raw interview evidence and researcher interpretation needs to be clearer. For example, “aggression in response to a friend’s playfulness” is categorized as “failure to understand others’ intentions,” and “standing unnecessarily close” is categorized as a Theory of Mind-related social-interaction deficit. These interpretations may be plausible, but they are not self-evidently equivalent to Theory of Mind deficits and could alternatively reflect pragmatic-language problems, sensory differences, social-learning history, executive dysfunction, or broader autism-related social-communication characteristics. The authors should explain the analytic rules used to assign each open code to a Theory of Mind category and should avoid implying a specific causal cognitive mechanism unless the interview material directly supported that interpretation.

The grounded-theory table describing “causal conditions” should be revised because the term “causal” is used more strongly than the qualitative data permit. Categories such as “deficit in understanding nonliteral language,” “deficit in empathy,” “inability to deceive and understand deception,” and “pragmatic-language deficits” are treated as factors causing weak social skills and stereotyped behaviors. However, the current data consist of expert perceptions rather than longitudinal or experimental evidence establishing causal direction. In grounded-theory terminology, “causal conditions” may be permissible within a Straussian paradigm model, but the manuscript should explicitly clarify that these are analytically derived conditions within the experts’ explanatory model, not empirically demonstrated causal determinants. This distinction is particularly necessary because the Discussion currently interprets some of these associations as mechanistic relationships.

The classification of stereotyped behaviors within the “central phenomenon” requires stronger theoretical and empirical justification. The manuscript identifies both “weak social skills” and “emergence of stereotyped behaviors” as central phenomena and later suggests that improved Theory of Mind could reduce stereotyped behaviors by increasing environmental predictability. However, this pathway was not tested in the present study. Moreover, stereotyped and repetitive behaviors in ASD are heterogeneous and may be influenced by sensory regulation, restricted interests, anxiety, arousal, reinforcement contingencies, and neurological factors in addition to social unpredictability. The authors should therefore present this relationship as an expert-generated hypothesis or conceptual proposition rather than a demonstrated effect, and the conceptual model should distinguish between directly targeted outcomes and theoretically secondary outcomes.

The manuscript should reconsider the terminology “high-functioning autism” and “level 1 autism,” which are used inconsistently across sections. The Abstract concludes that the package can support rehabilitation of “high-functioning children with autism,” while later sections define the package as intended for “high-functioning autism spectrum children (level 1).” These terms are not synonymous, and “high-functioning autism” is increasingly regarded as imprecise because it may obscure substantial adaptive, language, sensory, or support needs. The authors should specify the intended clinical population using operational criteria, preferably DSM-5 support level, language ability, intellectual functioning, and age range. If the target population is children aged 6–12 years with ASD level 1 and sufficient receptive language to participate in ToM tasks, this should be stated consistently in the title, Abstract, Methods, package description, and Discussion.

The developmental sequencing of the 40-session package requires reconciliation across the manuscript. One section presents sessions 1–4 as basic emotion recognition, 5–8 as joint attention, 9–12 as desires, 13–16 as first-order false belief, 17–20 as behavior prediction, 21–24 as deception, 25–28 as second-order false belief, 29–32 as empathy, 33–36 as humor, and 37–40 as irony and idioms. However, the later “general structure of the 40-session package” allocates 8 sessions to emotional foundations, 6 to joint attention, 6 to desires, 8 to first-order false belief, 4 to behavior prediction, 4 to second-order false belief, and 4 to empathy and humor. These are materially different intervention structures. The authors must decide which version represents the final validated package and remove or clearly label earlier developmental drafts. Otherwise, clinicians cannot reproduce the intervention reliably.

Authors revised the manuscript and uploaded the document.

2. Revised

Editor’s decision: Accepted.

Editor in Chief’s decision: Accepted.