



Validation of Thematic Apperception Test (TAT) in the Iranian Population

Sajjad Elahdadi^{1✉}, Asghar Minaei², Mohammad Reza Falsafinejad³, Mohammad-kazem Atef-Vahid⁴

1. PhD Candidate in Assessment and measurement, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran. E-mail: alahdadi_s@atu.ac.ir
2. Associate Professor of Assessment and Measurement, Department, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran. E-mail: asghar.minaei@yahoo.com
3. Associate Professor of Assessment and Measurement, Department, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran. E-mail: falsafinejad@yahoo.co.uk
4. Associate Professor, Department of Clinical Psychology, Mental Health Research Center and School of Behavioral Sciences and Mental Health (Tehran Institute of Psychiatry), Iran University of Medical Sciences, Tehran, Iran. E-mail: kazemv@yahoo.com

ARTICLE INFO

Article type:

Research Article

Article history:

Received 19 September 2025

Received in revised form 09 February 2026

Accepted 19 March 2026

Published Online 23 July 2026

Keywords:

Thematic Apperception Test (TAT),
Social Cognition and Object Relations Scale - Global Rating Method (SCORS-G),
validity,
reliability.

ABSTRACT

Background: Any psychological test should undergo validation before being used in a specific population to ensure its reliability and validity in the target context, and projective tests are no exception. Although the Thematic Apperception Test (TAT) has been used in Iran for many years, its psychometric properties have not yet been validated in the Iranian population.

Aims: The present study aimed to validate the Thematic Apperception Test (TAT) using the Social Cognition and Object Relations Scale - Global Rating Method (SCORS-G) scoring system in the Iranian population.

Methods: This was a descriptive study aimed at validating the instrument. The statistical population included Iranian adults aged 18 to 65 residing in Tehran in the winter of 2024, from which 154 participants were selected through convenience sampling. The research instruments comprised the Thematic Apperception Test (TAT) Murray & Morgan (1943), the Minnesota Multiphasic Personality Inventory-2 (MMPI-2; 71-item form) butcher et al (2001), and the Social Cognition and Object Relations Scale- Global Rating Method (SCORS-G) scoring system. Participants were divided into clinical and non-clinical groups based on MMPI-2 scores. Data were analyzed using Pearson correlation, Hotelling's T^2 , intraclass correlation coefficient (ICC), and Cronbach's alpha via SPSS-24.

Results: Results showed that TAT using SCORS-G scoring system showed significant differences between clinical and non-clinical groups across five scales ($p \leq 0.01$). Furthermore, all scales of MMPI-2 were significantly correlated with at least one SCORS-G scale, with correlation coefficients ranging from 0.16 to 0.43. ICC for the eight SCORS-G scales ranged from 0.77 to 0.92, while Cronbach's alpha values ranged from 0.64 to 0.95.

Conclusion: The Thematic Apperception Test (TAT), scored with the Social Cognition and Object Relations Scale-Global Rating Method (SCORS-G), demonstrates acceptable validity and reliability in the Iranian population, and its use is recommended in clinical and research settings, particularly for assessing cognitive and affective components in adults.

Citation: Elahdadi, S., Minaei, A., Falsafinejad, M.R., & Atef-Vahid, M.K. (2026). Validation of Thematic Apperception Test (TAT) in the Iranian population. *Journal of Psychological Science*, 25(161), 1-19. [10.66224/jps.25.161.4](https://doi.org/10.66224/jps.25.161.4)

Journal of Psychological Science, Vol. 25, No. 161, 2026

© The Author(s). DOI: [10.66224/jps.25.161.4](https://doi.org/10.66224/jps.25.161.4)



✉ **Corresponding Author:** Sajjad Elahdadi, PhD Candidate in Assessment and measurement, Faculty of Psychology and Educational Sciences, Allameh Tabataba'i University, Tehran, Iran.
E-mail: alahdadi_s@atu.ac.ir, Tel: (+98)9364222421

Extended Abstract

Introduction

One of the most widely used projective tests is the Thematic Apperception Test (TAT), recognized as the fourth most frequently employed psychological test among adults (Narvon, 2005). For the proper use of psychological tests, it is essential to examine their validity and reliability within the target population (Lilienfeld., 2006). In this regard, efforts have consistently been made to enhance the test's reliability by introducing more objective methods and approaches. Recently, the Social Cognition and Object Relations Scale (SCORS) has received considerable attention. First developed by Westen (1985), this scale provides a structured framework for scoring story-based tests (Westen, 1990). The advantages of this scoring method include establishing standardized scoring conditions, ease of scoring, and creating a common foundation for further research. The SCORS – Global Scoring Method (SCORS-G) is the latest version of this scale, which, unlike previous versions that used Component Scoring, applies a global scoring approach (Stein & Slavin-Mulford, 2017). Since its introduction, numerous studies have employed this method to assess the reliability and validity of the TAT. These studies have examined construct validity (Hibbard et al., 1995; Bram, 2014; Haggerty et al., 2015), convergent validity (Ackerman et al., 2001; Hibbard et al., 2001), internal consistency reliability (Hibbard et al., 2001; Inslegers et al., 2012), inter-rater reliability (Inslegers et al., 2012; Stein et al., 2012; Siefert et al., 2018), factorial Structure (Bram., 2014), known-groups validity (Bouvet et al., 2020), and test-retest reliability (Ridenour et al., 2022), with most studies reporting satisfactory results. Furthermore, a systematic review indicated that the TAT scored using SCORS-G demonstrates relatively adequate validity and reliability across various studies (Alahdadi et al., 2025).

Given the wide application of the TAT (Aronow et al, 2001) and the necessity of validating psychological instruments within target populations, it is important to examine its psychometric properties in the Iranian population. Therefore, the present study aims to

validate the Thematic Apperception Test (TAT) using the SCORS – Global Scoring Method (SCORS-G) in an Iranian sample.

Method

This study employed a descriptive design of the instrument validation type. To examine the validity of the instrument, the known-groups validity method and convergent validity with the MMPI-2 were used. Reliability was assessed through internal consistency (Cronbach's alpha) and the intraclass correlation coefficient (ICC).

The study population consisted of Iranian adults aged 18 to 65 years residing in Tehran in the winter of 2024–2025. Inclusion criteria were: being an Iranian adult aged 18–65 years, residing in Tehran during the study period, and willingness to participate in the research. The sample size of 150 participants was determined based on references in the literature (Martin & Hollins Martin., 2017) and recruited using convenience sampling. Data were collected using the Thematic Apperception Test (TAT) Murray & Morgan (1943) and the short form (71 items) of the Minnesota Multiphasic Personality Inventory–2 (MMPI-2) Butcher et al (2001).

For data collection, participants first completed the MMPI-2 either in person or online. Responses were reviewed, and those with valid scores on the test's validity scales were subsequently contacted and invited to participate in the second stage, namely administration of the TAT. The TAT was administered under standardized, controlled conditions at the Rooyesh Clinic in Tehran, for which prior authorization had been obtained. Before administration, participants received a full explanation of the study procedures. The test was then administered, and participants' verbal responses were audio-recorded. Finally, the responses were independently scored by two raters. The data collection was conducted over approximately three months.

To ensure scoring accuracy, two months prior to data collection the raters studied the SCORS-G scoring protocol thoroughly and completed the practice exercises provided in the scoring manual to gain proficiency in scoring. In the final step, blind to participants' other information, the raters

independently scored the TAT narratives. Data were analyzed using Hotelling's T^2 test, Pearson's correlation coefficient, Cronbach's alpha, and the intraclass correlation coefficient (ICC) through SPSS-26.

Results

A total of 154 individuals participated in this study. Participants' ages ranged from 18 to 51 years, with a mean of 32.55 (SD = 5.86). Of the total sample, 27.3% (n = 42) were male and 72.7% (n = 112) were female.

To assess construct validity using the known-groups method, the sample was divided into two groups. The

nonclinical group consisted of individuals who did not obtain a score under 70 on any of the MMPI-2 scales, while the clinical group included those who scored 70 or above on at least one of the main MMPI-2 scales. For each participant, the mean scores of the eight SCORS-G scales across 20 TAT stories were calculated. In addition, the mean scores of the two raters were computed and then subjected to Hotelling's T^2 test. The results of this analysis, presented in Table 1, showed that the combined SCORS-G scales significantly differentiated between the clinical and nonclinical groups ($F=4.56$, $p<.001$).

Table 1. Hotelling's T^2 Test Comparing SCORS-G Scales Across Clinical and Nonclinical Groups

exams	value	df	Df error	f	p
Pillai's Trace	0.2	8.000	145.000	4.56	<0.001
Wilks' Lambda	0.8	8.000	145.000	4.56	<0.001
Hotelling's Trace	0.25	8.000	145.000	4.56	<0.001
Roy's Largest Root	0.25	8.000	145.000	4.56	<0.001

The results of the between-group effects (Table 2) indicated that five scales—including Affective Quality of Representations (AFF), Emotional Investment in Relationships (EIR), Experience and Management of Aggressive Impulses (AGG), Self-Esteem (SE), and Identity and Coherence of Self

(ICS)—were significant ($p < .001$). That is, the mean scores of the clinical and nonclinical groups differed significantly on these scales. No significant differences were observed between the two groups on the other scales ($p > .05$).

Table 2. Between-group effects of Hotelling's T^2 test for the SCORS-G scales in the nonclinical and clinical groups

exams	Group	N	Mean	SD	f	p
Complexity of Representation of People (COM)	Nonclinical	122	3.84	0.60	0.56	0.45
	Clinical	32	3.74	0.63		
Affective Quality of Relationships (AFF)	Nonclinical	122	3.77	0.50	22.78	0.000
	Clinical	32	3.31	0.43		
Emotional Investment in Relationships (EIR)	Nonclinical	122	3.99	0.55	12.33	0.001
	Clinical	32	2.61	0.56		
Emotional Investment in values and Moral standards (EIM)	Nonclinical	122	3.88	0.20	2.22	0.14
	Clinical	32	3.81	0.27		
Anderstanding of Social Causality (SC)	Nonclinical	122	3.56	0.68	1.94	0.17
	Clinical	32	3.38	0.54		
Experience and Management of Aggressive Impulses (AGG)	Nonclinical	122	3.80	0.22	8.34	0.004
	Clinical	32	3.67	0.23		
Self-Esteem (SE)	Nonclinical	122	3.77	0.43	11.97	0.001
	Clinical	32	3.47	0.44		
Identity and Coherence of Self (ICS)	Nonclinical	122	3.89	0.46	15.00	0.000
	Clinical	32	3.53	0.48		

To examine convergent validity, a correlation matrix was created using Pearson's correlation coefficients between the MMPI-2 scales and the SCORS-G

scales, as presented in Table 3. The results indicated that all MMPI-2 scales showed at least one significant Correlation with the SCORS-G scales. These correlations ranged from $r = .16$, $p \geq .05$ for the L

scale with ICS, to $r = .49$, $p \leq .01$ for the Ma scale with EIR.

Table 3. Correlations between MMPI-2 scales and SCORS-G scores (TAT)

	COM	AFF	EIR	EIM	SC	AGG	SE	ICS
L	-0.14	0.14	0.05	0.06	-0.17*	0.18*	0.11	0.16*
F	-0.05	-0.17*	-0.21**	-0.08	-0.06	-0.25**	-0.17*	-0.09
K	0.23*	0.34**	0.37**	0.07	0.30**	0.11	0.25**	0.33**
Hs	-0.13	-0.21*	-0.13	-0.09	-0.08	-0.12	-0.13	-0.19*
D	-0.25**	0.04	-0.13	0.20*	-0.27*	-0.13	0.03	0.09
Hy	-0.14	0.11	-0.04	0.14	-0.14	0.02	0.13	0.26**
Pd	-0.18*	-0.23**	-0.43**	-0.12	-0.17*	-0.14	-0.31**	-0.15
Pa	-0.20*	-0.08	-0.33**	-0.13	-0.22**	0.09	-0.13	-0.03
Pt	-0.16	-0.06	-0.31**	0.06	-0.20**	-0.03	-0.08	-0.11
Sc	-0.07	-0.06	-0.20*	0.20*	-0.14	-0.03	-0.07	-0.05
Ma	-0.28**	-0.40**	-0.49**	-0.09	-0.30**	-0.26**	-0.41**	-0.29**

Reliability of the test was assessed using the intraclass correlation coefficient (ICC), with the results presented in Table 4. The lowest ICC was observed for the Identity and Coherence of Self (ICS) scale (ICC = .77), whereas the highest ICC was found for the Emotional Investment in Relationships (EIR) scale (ICC = .92).

Additionally, internal consistency reliability was assessed using Cronbach's alpha, with the results presented in Table 4. Cronbach's alpha for the SCORS-G scales ranged from .64 for the Experience and Management of Aggressive Impulses (AGG) scale to .95 for the Social Causality (SC) scale. The overall Cronbach's alpha for all scales was 0.97.

Table 4. Intraclass correlation coefficients (ICC) and Cronbach's alpha for SCORS-G scores (TAT)

exams	n	Cronbach's alpha	ICC
Complexity of Representation of People (COM)	20	0.94	0.84
Affective Quality of Relationships (AFF)	20	0.85	0.84
Emotional Investment in Relationships (EIR)	20	0.85	0.92
Emotional Investment in values and Moral standards (EIM)	20	0.69	0.86
Understanding of Social Causality (SC)	20	0.95	0.87
Experience and Management of Aggressive Impulses (AGG)	20	0.64	0.83
Self-Esteem (SE)	20	0.83	0.79
Identity and Coherence of Self (ICS)	20	0.88	0.77
Total	20	0.97	0.88

Conclusion

The present study aimed to validate the Thematic Apperception Test (TAT) using the SCORS-G scoring system in an Iranian population. The known-groups validity results indicated that all affective component scales of SCORS-G, except for the EIM scale, significantly differentiated between clinical and non-clinical groups, whereas the cognitive component did not. These findings are consistent with the results of Hibbard et al. (1995) and Ackerman et al. (2001), and partially aligned with those of Bouvet et al. (2020) and Bram. (2014). The discrepancies in results may be attributed to differences in the study

samples, the cards used, or the limited sensitivity of some scales to group characteristics.

Convergent validity results showed that out of 88 possible correlations, 38 were significant, which is consistent with the findings of Ackerman et al. (2001) and Haggerty et al. (2015). These results suggest that SCORS-G and MMPI-2 demonstrate notable convergence in certain dimensions, while simultaneously offering unique features that provide valuable complementary information for psychological assessment.

Inter-rater reliability assessed through ICC indicated an adequate level of agreement among raters. This finding is in line with previous studies, all of which

reported satisfactory inter-rater reliability (Inslegers et al., 2012; Siefert et al., 2017; Stein et al., 2012; Ridenour et al., 2022). Moreover, internal consistency reliability, examined via Cronbach's alpha, demonstrated acceptable to excellent levels across scales. These findings are consistent with Alahdadi et al. (2025) and, to some extent, with Inslegers et al. (2012) and Hibbard et al. (2010), though not for all scales. Such differences may be due to the smaller number of cards used in some studies, variations in study samples, and cultural differences. Among the limitations of the present study were the restricted training and practice period for raters (about two months), the use of convenience sampling, and the relatively small sample size. Future research may address these limitations by including more experienced raters, employing random sampling, and utilizing larger samples to further evaluate the instrument and extract the factorial structure of the scales within the Iranian population, thereby providing more comprehensive methodological evidence.

In conclusion, the findings indicated that the TAT, scored using SCORS-G, demonstrated acceptable reliability and validity in the Iranian population. Reliability was supported by both inter-rater agreement and internal consistency. Significant differences between clinical and non-clinical groups and meaningful correlations with some MMPI-2 scales provide considerable evidence for the test's psychometric properties. Some scales and correlations were not significant, which may reflect the actual lack of association between variables in those areas or the current scoring method, highlighting directions for further research. Overall, The Thematic Apperception Test (TAT), scored with the Social Cognition and Object Relations Scale–Global Rating Method (SCORS-G), demonstrates acceptable validity and reliability in the Iranian population, and its use is recommended in clinical and research settings, particularly for assessing cognitive and affective components in adults.

Ethical Considerations

Compliance with ethical guidelines: This research is based on the doctoral dissertation of the first author in the field of Assessment and Measurement at the Faculty of Psychology and

Educational Sciences, Allameh Tabataba'i University that Its proposal was defended on January 15, 2024. To adhere to ethical principles, prior to the administration of the tests, the purpose of the study, the procedures, the method of audio recording, and the participants' rights were fully explained. Written informed consent was obtained from all participants, and it was emphasized that participation was entirely voluntary and that they could withdraw from the study at any stage without any negative consequences. To ensure privacy and confidentiality, the names and personal details of the participants were not mentioned at any stage of reporting or publishing the results. The collected data were used solely for research purposes and were not accessible to unrelated individuals or organizations. Additionally, upon request, participants were provided with the results of the tests and the overall findings of the study.

Funding: This study was conducted as a PhD thesis with no financial support.

Authors' contribution: This article is derived from the first author's doctoral dissertation, conducted under the supervision of the second and third authors and the advisement of the fourth author.

Conflict of interest: The authors declare no conflict of interest for this study.

Acknowledgments: The researchers would like to express their gratitude to the people who participated in this study.



اعتباریابی آزمون اندریافت موضوع (TAT) در جمعیت ایرانی

سجاد اله دادی^۱، اصغر مینائی^۲، محمدرضا فلسفی نژاد^۳، محمد کاظم عاطف وحید^۴

۱. کاندیدای دکتری تخصصی سنجش و اندازه‌گیری، گروه سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه:

alahdadi_s@atu.ac.ir

۲. دانشیار، گروه سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: asghar.minaei@yahoo.com

۳. دانشیار، گروه سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه: falsafinejad@yahoo.co.uk

۴. دانشیار، گروه روانشناسی بالینی، مرکز تحقیقات بهداشت روان و دانشکده علوم رفتاری و سلامت روان (انستیتو روانپزشکی تهران)، دانشگاه علوم پزشکی ایران، تهران، ایران. رایانامه:

kazemv@yahoo.com

چکیده

مشخصات مقاله

زمینه: هر آزمون روان‌شناختی پیش از استفاده در یک جمعیت خاص باید اعتباریابی شود تا روایی و پایایی آن در بافت مورد نظر تأیید گردد و آزمون‌های فراکن هم از این موضوع مستثنا نیستند؛ اما با وجود اینکه آزمون اندریافت موضوع (TAT) سال‌هاست که در ایران استفاده می‌شود، تاکنون اعتباریابی آن در جمعیت ایرانی صورت نگرفته است.

هدف: هدف پژوهش حاضر، اعتباریابی آزمون اندریافت موضوع (TAT) با نمره‌گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی (SCORS-G) در جمعیت ایرانی بود.

روش: روش پژوهش توصیفی و از نوع اعتباریابی ابزار بود و جامعه‌ی آماری آن شامل افراد بزرگسال بین ۱۸ تا ۶۵ سال ایرانی ساکن شهر تهران در زمستان ۱۴۰۳ بود که از بین آن‌ها تعداد ۱۵۴ نفر به روش نمونه‌گیری در دسترس انتخاب شدند. ابزارهای پژوهش شامل آزمون اندریافت موضوع موری و مورگان (۱۹۴۳)، سیاهه شخصیتی چندوجهی مینه‌سوتا ۲ - فرم ۷۱ سؤالی بوچر و همکاران (۲۰۰۱) و روش نمره‌گذاری مقیاس کلی شناخت اجتماعی و روابط موضوعی بود. نمونه با استفاده از سیاهه شخصیتی چندوجهی مینه‌سوتا ۲ به دو گروه بالینی و غیربالینی تقسیم شد. داده‌ها با استفاده از آزمون‌های همبستگی پیرسون، T^2 هاتلینگ، ضریب همبستگی درون طبقه‌ای و ضریب آلفای کرونباخ و از طریق نرم‌افزار SPSS-24 تحلیل شد.

یافته‌ها: نتایج نشان داد آزمون اندریافت موضوع با نمره‌گذاری مقیاس کلی شناخت اجتماعی و روابط موضوعی در ۵ مقیاس بین دو گروه بالینی و غیربالینی تفاوت معنادار ایجاد کرد ($P \leq 0/01$). همچنین همه مقیاس‌های سیاهه شخصیتی چندوجهی مینه‌سوتا ۲ هر کدام حداقل با یکی از مقیاس‌های مقیاس کلی شناخت اجتماعی و روابط موضوعی همبستگی معنادار داشتند. مقدار این همبستگی‌ها از ۰/۱۶ تا ۰/۴۳ متغیر بود. ضریب همبستگی درون طبقه‌ای برای هشت مقیاس از ۰/۷۷ تا ۰/۹۲ متغیر بود. ضریب آلفای کرونباخ هم برای هشت مقیاس از ۰/۶۴ تا ۰/۹۵ بود.

نتیجه‌گیری: آزمون اندریافت موضوع (TAT) با نمره‌گذاری بر اساس مقیاس کلی شناخت اجتماعی و روابط موضوعی (SCORS-G)، از روایی و پایایی مناسبی در جمعیت ایرانی برخوردار است و استفاده از آن به‌ویژه در محیط‌های بالینی و پژوهشی برای سنجش مؤلفه‌های شناختی و عاطفی در بزرگسالان توصیه می‌شود.

نوع مقاله:

پژوهشی

تاریخچه مقاله:

دریافت: ۱۴۰۴/۰۶/۲۸

بازنگری: ۱۴۰۴/۱۱/۲۰

پذیرش: ۱۴۰۴/۱۲/۲۸

انتشار برخط: ۱۴۰۵/۰۵/۰۱

کلیدواژه‌ها:

آزمون اندریافت موضوع، مقیاس کلی شناخت اجتماعی و روابط موضوعی، روایی، پایایی

استناد: اله دادی، سجاد؛ مینائی، اصغر؛ فلسفی نژاد، محمدرضا؛ و عاطف وحید، محمدکاظم (۱۴۰۵). اعتباریابی آزمون اندریافت موضوع (TAT) در جمعیت ایرانی. *مجله علوم روانشناختی*، دوره ۲۵، شماره ۱، ۱۹-۱۶۱.

مجله علوم روانشناختی، دوره ۲۵، شماره ۱، ۱۴۰۵. DOI: 10.66224/jps.25.161.4



© نویسندگان.

✉ **نویسنده مسئول:** سجاد اله دادی، کاندیدای دکتری تخصصی سنجش و اندازه‌گیری، دانشکده روانشناسی و علوم تربیتی، دانشگاه علامه طباطبائی، تهران، ایران. رایانامه:

alahdadi_s@atu.ac.ir

تلفن: ۰۹۳۶۴۲۲۲۴۲۱

مقدمه

امروزه اهمیت استفاده از آزمون‌های فرافکن در روانشناسی، بر کسی پوشیده نیست؛ تا حدی که حتی برخی پژوهش‌ها، یادگیری آزمون‌های فرافکن برای روانشناسان را با یادگیری تشریح برای پزشکان مقایسه کرده و از کاهشی بودن روند آن در برنامه‌های تحصیلات تکمیلی انتقاد می‌کنند و مدعی‌اند که آزمون‌های فرافکن باید بخش جدایی‌ناپذیری از آموزش روانشناسان باشد (جوی، ۲۰۲۰). مزیت اصلی در این روش این است که سوگیری‌ها، عدم صداقت و دفاع‌های فرد، کنار زده می‌شوند و واقعیت اصلی، نمایان می‌شود؛ حتی اگر فرد، سعی در فریب درمانگر داشته باشد. به همین سبب از آزمون‌های فرافکن به عنوان ایکس ری - ذهنی نیز نام برده می‌شود (نسلی و دی فریسه، ۲۰۲۰).

یکی از مهم‌ترین آزمون‌های فرافکن، آزمون اندریافت موضوع^۱ (TAT) است (نرون، ۲۰۰۵). این آزمون در سال ۱۹۳۵ توسط هنری موری و کریستین مورگان ساخته شد؛ عقیده هنری موری این بود که نیازها و فشارهای ناخودآگاه در افراد وجود دارد که انگیزه رفتارهای دیگر از جمله رفتارهای ناسازگار می‌شود (مورگان و موری، ۱۹۳۵). کاربرد آزمون TAT به زودی گسترش یافت و تا جایی پیش رفت که چهارمین آزمون روانی پرکاربرد در بین بزرگسالان شد (نرون، ۲۰۰۵).

اما برای استفاده از آزمون‌های روانشناسی لازم است اعتباریابی آن در جمعیت‌های هدف بررسی شود. همان‌گونه که لیلینفیلد (۲۰۰۶) تأکید می‌کند، این ابزارها بدون ارزیابی و اثبات اعتبار و پایایی نمی‌توانند به طور مطمئن برای تشخیص یا پژوهش مورد استفاده قرار گیرند؛ بنابراین، انجام مطالعات اعتبارسنجی برای هر جمعیت و زمینه بالینی، به ویژه برای TAT که از محرک‌های تصویری متنوع استفاده می‌کند، امری حیاتی است و پایه‌ی استفاده علمی و بالینی این آزمون‌ها را تشکیل می‌دهد (لیلینفیلد، ۲۰۰۶).

اما با وجود چالش‌هایی که برای اعتباریابی TAT وجود دارد (از جمله تعدد شیوه‌های نمره‌گذاری، کیفی بودن برخی شیوه‌ها، تفسیر وابسته به قضاوت

متخصص و غیره)، همچنان نیاز به استفاده از روش‌هایی جهت اعتباریابی این ابزارها اجتناب‌ناپذیر است و در همین راستا همواره تلاش شده است با معرفی روش‌ها و رویکردهای عینی‌تر به افزایش اعتبار این آزمون‌ها کمک شود. در همین راستا اخیراً مقیاس شناخت اجتماعی و روابط موضوعی^۲ (SCORS) توجه زیادی را به خود معطوف کرده است. این مقیاس که نخستین بار توسط وستن و همکارانش (۱۹۸۵) توسعه یافت، برای نمره‌گذاری آزمون‌هایی که محتوای داستانی دارند، یک چارچوب اندازه‌گیری را معرفی می‌کند (وستن، ۱۹۹۰). در سال ۱۹۹۵ این مقیاس بازبینی شد و پاسخ‌ها که در آغاز طیف لیکرت ۵ گزینه‌ای بود، به طیف ۷ گزینه‌ای تغییر کرد. همچنین ابعاد آن هم که تا آن زمان ۴ بعد بود، به ۸ بعد افزایش پیدا کرد. ابعاد جدید شامل این موارد بود: پیچیدگی بازنمایی‌های افراد^۳ (COM)، کیفیت عاطفی بازنمایی‌ها^۴ (AFF)، سرمایه‌گذاری هیجانی در روابط^۵ (EIR)، سرمایه‌گذاری هیجانی در ارزش‌ها و استانداردهای اخلاقی^۶ (EIM)، درک علت اجتماعی^۷ (SC)، تجربه و مدیریت تکانه‌های خشم^۸ (AGG)، عزت نفس^۹ (SE)، هویت و انسجام خود^{۱۰} (ICS) (استین و اسلاون-مالفورد، ۲۰۱۷). نمرات بالاتر در این نمره‌گذاری نشان دهنده انطباق و سلامت بیشتر و نمرات پایین‌تر نشان دهنده آسیب بیشتر است. این مقیاس‌ها در دو مؤلفه کلی شناختی (شامل پیچیدگی بازنمایی‌های افراد و درک علت اجتماعی) و عاطفی (شامل بقیه مقیاس‌ها) قرار می‌گیرند. از مزیت‌های این روش نمره‌گذاری ایجاد شرایط استاندارد نمره‌گذاری، سهولت نمره‌گذاری و ایجاد زمینه مشترک برای پژوهش‌های بیشتر است. مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی^{۱۱} (SCORS-G) آخرین نسخه از این مقیاس است که برخلاف نسخه‌های قبلی که از روش نمره‌گذاری جز به جز^{۱۲} استفاده می‌کردند، از روش نمره‌گذاری کلی استفاده می‌کند. این روش بر اساس برداشت جامع و کل‌نگر از روایت صورت می‌پذیرد و موجب می‌شود که ابعاد روانی، شناختی و روابط ابژه‌ای به صورت یکپارچه و در قالب یک امتیاز کلی، بدون تمرکز بر جزئیات منفرد، تحلیل شوند (استین و اسلاون-مالفورد،

8. Experience and Management of Aggressive Impulses

9. Self-Esteem

10. Identity and Coherence of Self

11. The Social Cognition and Object Relations Scale-Global Rating Method

12. Component Scoring

1. Thematic Apperception Test

2. Social Cognition and Object Relations Scale

3. Complexity Of Representations

4. Affective quality of representations

5. Emotional Investment in Relationships

6. Emotional Investment in values and Moral standards

7. Understanding of Social Causality

۲۰۱۷). از زمان معرفی روش مقیاس شناخت اجتماعی و روابط موضوعی و نسخه نهایی آن یعنی مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی پژوهش‌های مختلفی به اعتباریابی آزمون اندریافت موضوع با این روش نمره گذاری پرداخته‌اند. از جمله آن، پژوهش هیبرد و همکارانش (۱۹۹۵) بود که روایی سازه آزمون اندریافت با مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی را بررسی کردند؛ آنان نتیجه گرفتند که بازنمایی‌های عاطفی از ابژه با شدت آسیب روانی ارتباط دارد و نه بازنمایی‌های شناختی (هیبرد و همکاران، ۱۹۹۵). در پژوهشی دیگر آکرمن و همکارانش (۲۰۰۱) به بررسی روایی همگرایی آزمون اندریافت موضوع و رورشاخ در حیطه روابط موضوعی پرداختند. یافته‌های این پژوهش هم نشان از همبستگی‌های متوسط معنادار (از ۰/۲۵ تا ۰/۴۴) بین مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی با تعدادی از مقیاس‌های نمره گذاری رورشاخ^۱ (MOA) بود که عمدتاً مربوط به مقیاس‌های عاطفی مقیاس شناخت اجتماعی و روابط موضوعی بودند (آکرمن و همکاران، ۲۰۰۱).

هیبرد و همکارانش (۲۰۰۱) در پژوهشی دیگر به بررسی همسانی درونی نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی از کارت‌های آزمون اندریافت موضوع پرداختند. یافته‌ها نشان داد مقیاس‌های شناختی، همسانی درونی بیشتری نسبت به مقیاس‌های عاطفی کسب کردند. همچنین یافته استفاده از تعداد ۱۰ الی ۱۲ کارت را برای دستیابی به همسانی درونی بالاتر از ۰/۷ پیشنهاد کردند (هیبرد و همکاران، ۲۰۰۱). یافته‌های پژوهش اینسلجرز و همکارانش (۲۰۱۲) با هدف تعیین روایی و پایایی نمرات مقیاس شناخت اجتماعی و روابط موضوعی به دست آمده از داستان‌های آزمون اندریافت موضوع نشان داد توافق بین ارزیاب‌ها در سطح بالایی قرار داشت؛ اما همسانی درونی مقیاس‌ها بخصوص برای مقیاس‌های عاطفی همسو با پژوهش هیبرد و همکاران (۲۰۱۰) پایین بود (اینسلجرز و همکاران، ۲۰۱۲). پژوهش استین و همکارانش (۲۰۱۲) بر روی نمونه بالینی نشان دهنده پایایی بالای توافق بین ارزیاب‌ها بود. برای این منظور از ضریب همبستگی درون طبقه‌ای^۲ (ICC) استفاده شد که همه ضرایب بالاتر از ۰/۷۹ بودند (استین و همکاران، ۲۰۱۲). همچنین نتایج تحلیل مؤلفه‌های اصلی نشان از وجود سه

مؤلفه (به جای دو مؤلفه پژوهش اصلی) بود. مقدار ضریب ICC در پژوهش سیفرت و همکاران نیز برای همه مقیاس‌ها بالای ۰/۷۹ گزارش شد که مناسب بود و نتایج تحلیل عاملی اکتشافی هر دو مدل دو عاملی و سه عاملی را تأیید کرد (سیفرت و همکاران، ۲۰۱۸).

برم (۲۰۱۴) روایی ساختار عاملی را بررسی کردند و نتایج آن تأییدکننده دو عامل شناختی و عاطفی بود که با یافته‌های پژوهش اصلی سازگار بود و حاکی از روایی سازه ابزار بود. همچنین همبستگی‌های متوسط معنادار بین مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی با سلامت روانی، جسمی و کارکرد بین فردی وجود داشت. علاوه بر آن نمونه غیربالینی در هر هشت مقیاس، سالم‌تر از نمونه‌های بالینی پژوهش‌های قبلی بود. برم نتایج مربوط به روایی را رضایت‌بخش اعلام کرد (برم، ۲۰۱۴).

از دیگر پژوهش‌ها، هگرتی و همکارانش (۲۰۱۵) یک نمونه بالینی را بررسی کردند. در این پژوهش هر ارزیاب یک منبع اطلاعات بیمار را ارزیابی می‌کرد. در نهایت روایی بین ارزیاب‌ها در ابعاد مختلف از ۰/۱۷ تا ۰/۶۶ متغیر بود؛ اما روایی سازه و افزایشی در ارتباط با متغیرهایی مثل میزان مشارکت در درمان، کارکرد شخصی، کیفیت روابط با همسالان، کارکرد در مدرسه، ارزیابی کارکرد کلی^۳ (GAF)، تاریخچه رفتار اختلال خوردن و تاریخچه آسیب به خود (به جز خودکشی) بررسی شد که حاکی از روایی مناسب این ابزار برای نوجوانان بستری در بخش‌های روان‌پزشکی بود (هگرتی و همکاران، ۲۰۱۵).

از جمله روش‌های دیگری که در بررسی روایی به کار برده شد، بررسی توانایی ابزار در ایجاد تمایز بین گروه‌های بالینی و غیربالینی است که تحت عنوان روایی گروه‌های شناخته شده^۴ معروف است. بووت و همکارانش در سال ۲۰۲۰ پژوهشی را با هدف تعیین این نوع روایی برای داستان‌های آزمون اندریافت موضوع با نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی در نمونه فرانسوی انجام دادند. نتایج نشان داد که ابزار قادر است بین گروه‌های غیربالینی و بالینی تمایز قائل شود. به جز در دو مورد (کیفیت عاطفی بازنمایی‌ها و سرمایه گذاری هیجانی در روابط) بقیه ابعاد در بین گروه‌های غیربالینی و بالینی تفاوت معنادار نشان

³. Global Assessment of Functioning

⁴. Known-Groups Validity

¹. The Rorschach Mutuality of Autonomy Scale

². Intraclass correlation coefficients

دادند. پژوهشگران نتایج اعتباریابی را رضایت بخش اعلام کردند و بر لزوم پژوهش های بیشتر در این مورد تأکید کردند (بووت و همکاران، ۲۰۲۰). از معدود پژوهش های طولی هم رایدنور و همکارانش در سال ۲۰۲۲، به بررسی پایایی آزمون اندریافت موضوع با نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی در یک بازه زمانی ۱۸ ماهه پرداختند. یافته ها حاکی از همبستگی های بین ۰/۳ تا ۰/۵۲ در ۸ مقیاس بود. پژوهشگران گزارش کردند که گرچه این مقادیر از حد رضایت بخش کمتر است؛ اما با توجه به اینکه بیماران تحت درمان بوده اند و نیز به دلیل طولانی بودن فاصله بین دو آزمون، نتایج حاصل را قابل قبول ارزیابی کردند. همچنین نتایج توافق بین ارزیابان برای هر هشت مقیاس، خوب تا عالی گزارش شد (رایدنور و همکاران، ۲۰۲۲).

در نهایت اله دادی و همکاران (۲۰۲۵) در پژوهش مروری خود با عنوان مروری بر اعتباریابی آزمون اندریافت موضوع به روش نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی تعداد ۵۲ پژوهش را بررسی کردند که به صورت مستقیم یا غیرمستقیم به اعتباریابی آزمون اندریافت موضوع به روش مقیاس شناخت اجتماعی و روابط موضوعی یا مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی مرتبط بود (اله دادی و همکاران، ۲۰۲۵). نتایج پژوهش های مربوط به روایی ملاکی نشان داد در ۲۱ مورد از ۲۳ مورد (۹۱ درصد)، نمره های آزمون اندریافت موضوع، بین گروه های مورد مطالعه تمایز ایجاد کرد؛ در ۵ مورد از ۶ مورد (۸۳ درصد) تغییرات درمانی را پیش بینی کرد؛ در ۱۶ مورد از ۱۸ مورد (۸۹ درصد) با متغیرهای ملاک همبستگی معنادار داشت و در همه ۴ مورد (۱۰۰ درصد) مربوط به بررسی کشش کارت ها هم موفقیت داشت. همچنین بررسی روایی سازه همگرا نشان داد از ۷ پژوهش صورت گرفته ۵ مورد (۷۱ درصد) روایی همگرا را تأیید کردند. بررسی روایی به روش ساختار عاملی هم نشان داد در همه ۶ پژوهش (۱۰۰ درصد)، مدل های دو یا سه عاملی با داده ها برازش داشتند. نتایج پژوهش های مربوط به پایایی هم نشان داد از ۴۷ پژوهش صورت گرفته به روش توافق بین ارزیاب ها ۴۶ مورد (۹۸ درصد)، با تأیید پایایی همراه بود؛ از ۹ پژوهش صورت گرفته به روش همسانی درونی، ۷ مورد (۷۸ درصد) پایایی را مناسب گزارش کردند و در نهایت تنها پژوهش

مربوط به آزمون - باز آزمون هم همبستگی های قابل قبولی در جهت تأیید پایایی گزارش کردند (اله دادی و همکاران، ۲۰۲۵). بنابراین با توجه به گسترش روش نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی و پتانسیل بالقوه آزمون اندریافت موضوع جهت کاربرد در مراکز روان سنجی، کلینیک های مشاوره و روان درمانی، بیمارستان های روان پزشکی، مراکز اصلاح و تربیت، مراکز قضایی، ندامتگاه ها، استخدام نیروی انسانی در مشاغل حساس و غیره و همچنین کاربرد آن در پژوهش های مختلف (آرونوف و همکاران، ۲۰۰۱)، ضرورت دارد اعتباریابی آن در فرهنگ ایرانی بررسی شود؛ اما با وجود اینکه آزمون اندریافت موضوع سال ها است که در ایران استفاده می شود؛ تاکنون اعتباریابی آن در جمعیت ایرانی بررسی نشده است. لذا هدف پژوهش حاضر اعتباریابی آزمون اندریافت موضوع (TAT) با نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی (SCORS-G) در جمعیت ایرانی است.

روش

الف) طرح پژوهش و شرکت کنندگان: این پژوهش به روش توصیفی و از نوع اعتباریابی ابزار انجام شد. بدین منظور از روش های روایی گروه های شناخته شده و روایی همگرا با سیاهه شخصیتی چندوجهی مینه سوتا-۲ جهت بررسی روایی ابزار استفاده شد. برای بررسی پایایی ابزار نیز از روش های همسانی درونی (آلفای کرونباخ) و ضریب همبستگی درون طبقه ای^۱ (ICC) استفاده شد. همسانی درونی ابزار، از طریق ضریب آلفای کرونباخ حاصل از نمرات هشت مقیاس، مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی بررسی شد و برای محاسبه ضریب همبستگی درون طبقه ای، پس از اجرای آزمون بر روی نمونه، نتایج آن توسط دو نفر ارزیاب به طور جداگانه، نمره گذاری و تفسیر شد. پس از جمع آوری داده ها، در نهایت ضریب همبستگی درون طبقه ای محاسبه شد. جامعه پژوهش شامل افراد بزرگسال بین ۱۸ تا ۶۵ سال ایرانی ساکن شهر تهران در زمستان ۱۴۰۳ بود. معیارهای ورود به پژوهش در این بخش شامل این موارد بود: بزرگسال ایرانی ۱۸ تا ۶۵ سال بودن، ساکن تهران بودن در زمستان سال ۱۴۰۳ و داشتن تمایل به مشارکت در پژوهش. معیارهای

^۱. Intra-class Correlation Coefficient

خروج هم شامل: کامل نکردن آزمون‌ها، عدم تمایل به ادامه همکاری و انصراف از پژوهش در هر مرحله از کار بود. با توجه به وسعت جامعه پژوهش و زمان‌بر بودن اجرای آزمون اندریافت موضوع (TAT) که نیازمند همکاری داوطلبانه شرکت‌کنندگان بود، نمونه‌گیری به شیوه در دسترس انجام شد. این روش ضمن رعایت محدودیت‌های عملی، امکان جمع‌آوری داده‌های معتبر و کافی برای تحلیل‌های مربوط به اعتباریابی ابزار را فراهم کرد. برای تعیین حجم نمونه، با توجه به توصیه‌های موجود در پژوهش‌های اعتباریابی مبنی بر کفایت حداقل ۱۰۰ نفر برای برآورد شاخص‌های پایایی و روایی (مارتین و هالینس مارتین، ۲۰۱۶)، حجم نمونه ۱۵۰ نفر در نظر گرفته شد تا توان آماری پژوهش افزایش یافته و نتایج با دقت بیشتری برآورد شود.

(ب) ابزار: آزمون اندریافت موضوع^۱ (TAT): یک آزمون فرافکن است که در دهه ۱۹۳۰ توسط هنری موری و کریستینا مورگان در دانشگاه هاروارد طراحی شد و برای بررسی افکار، انگیزه‌ها، نگرش‌ها و نیازهای پنهان فرد کاربرد دارد. آزمون شامل ۳۰ کارت تصویری سیاه و سفید مبهم است که برخی از آن‌ها مخصوص مردان، برخی مخصوص زنان و برخی برای هر دو جنس طراحی شده‌اند و برای هر نفر با توجه به سن و جنسیت حداکثر ۲۰ کارت قابل اجرا است. تصاویر کارت‌ها شامل زمینه‌های اجتماعی، خانوادگی، بین فردی و دیگر موقعیت‌های مبهم است تا فرد بتواند داستانی درباره شخصیت‌ها و موقعیت‌ها بسازد. فرض نظری پشت آزمون اندریافت موضوع این است که فرد، عناصر ناخودآگاه ذهن خود را از طریق داستان‌پردازی آشکار می‌کند و پژوهشگر می‌تواند با تحلیل مضامین داستان‌ها، ویژگی‌های شخصیتی و الگوهای روانی پنهان را شناسایی کند. تاکنون روش‌های نمره‌گذاری مختلفی برای آزمون اندریافت موضوع معرفی شده است که به‌طور کلی می‌توان آن‌ها را در دو گروه روش‌های کمی و روش‌های کیفی دسته‌بندی کرد. در روش‌های کمی مثل روش مک کلند و یا روش استفانسن، روان‌شناسان به درجه و اهمیت مقولات مطرح شده در داستان‌ها نمره می‌دهند و تفسیر خود را بر بنیاد ارزش‌های کمی این نمرات استوار می‌کنند؛ در حالی که در روش‌های کیفی مثل روش موری (۱۹۴۳-۱۹۳۵)، هنری (۱۹۵۶)، امبردان (۱۹۶۹)، پیتروسکی (۱۹۵۰) و بلک (۱۹۶۵)، تحلیل بر پایه محتوای داستان‌ها،

^۱. Thematic Apperception Test

مضامین تکرارشونده، سبک روایت و بازنمایی روابط بین‌فردی انجام می‌گیرد و تأکید اصلی بر تفسیر بالینی و معنایی پاسخ‌ها است (بهرامی، ۱۳۹۳). روایی همگرا در برخی پژوهش‌ها که از نمره‌گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی استفاده کرده‌اند، ۰/۲۵ تا ۰/۴۴ گزارش شده است (هیرد و همکاران، ۲۰۰۱). پژوهش‌های پایایی هم نشان داده‌اند که همبستگی بین ارزیابی‌های مختلف از پاسخ‌های یکسان، متغیر است و می‌تواند بین ۰/۳۷ تا ۰/۹۰ باشد. این تفاوت‌ها نشان‌دهنده تأثیر تفسیرهای فردی و نبود استانداردهای واحد در ارزیابی‌هاست (لاندى، ۱۹۸۵). این مقدار برای برخی پژوهش‌ها که از روش مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی استفاده کرده‌اند، برای همه مقیاس‌ها بالای ۰/۷۹ گزارش شد (استین و همکاران، ۲۰۱۲). ضمناً در جستجوی منابع داخلی، پژوهشی که روایی و پایایی TAT را بررسی کرده باشد، یافت نشد. میزان همبستگی بین ارزیاب‌ها در پژوهش حاضر ۰/۸۸ بدست آمد.

سیاهه شخصیتی چندوجهی مینه‌سوتا^۲ (MMPI-2) فرم کوتاه: فرم اصلی این پرسشنامه در سال ۱۹۴۳ توسط هاتاوی و مک کینلی در دانشگاه مینه‌سوتا تهیه و تدوین شد و فرم کوتاه ۷۱ سؤال آن نخستین بار در کتاب راهنمای سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ (بوچر و همکاران، ۲۰۰۱) معرفی گردید. هدف از تدوین این فرم کوتاه، فراهم‌سازی ابزاری سریع‌تر و کارآمدتر برای موقعیت‌هایی بود که اجرای فرم کامل ۵۶۷ سؤال امکان‌پذیر نیست، به‌ویژه در پژوهش‌ها یا غربالگری‌های بالینی که با محدودیت زمان همراه است. این نسخه آزمون دارای سه مقیاس روایی و هشت مقیاس بالینی است. مقیاس‌های روایی شامل مقیاس‌های دروغ‌پردازی (L)، نمره روایی (F) و نمره اصلاح (K) نام دارد که میزان صداقت، تمارض و مقاومت آزمودنی‌ها را در پاسخ به پرسش‌ها می‌سنجد. مقیاس‌های بالینی عبارت‌اند از خودبیمارانگاری (Hs)، افسردگی (D)، هیستری (H)، انحراف روانی-اجتماعی (Pd)، پارانویا (Pa)، ضعف روانی (Pt)، اسکیزوفرنی (Sc) و مانیا (Ma). پاسخ‌های پرسشنامه به صورت دو گزینه‌ای (بله/خیر) است و نمره‌گذاری آن بر اساس تبدیل پاسخ‌های خام به نمرات استاندارد انجام می‌شود. در این فرایند، ابتدا تعداد پاسخ‌های کلید شده برای هر مقیاس به‌عنوان نمره خام محاسبه شده و سپس با استفاده از جداول هنجاری به

^۲. Minnesota Multiphasic Personality Inventory-2

نمرات T با میانگین ۵۰ و انحراف معیار ۱۰ تبدیل می‌گردد. تفسیر نتایج بر اساس نمرات T صورت می‌گیرد. نتایج پژوهش‌ها نشان داده است که برای فرم کوتاه سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ روایی همگرا در محدوده‌ی ۰,۲۹ تا ۰,۶۳، روایی افتراقی در محدوده‌ی ۰,۰۹ تا ۰,۵۷ و همچنین ضریب آلفای کرونباخ برای مقیاس‌های بالینی این آزمون در محدوده‌ی ۰/۸۵ تا ۰/۹۰ برخوردار بود (بوچر و همکاران، ۲۰۰۱). این آزمون در ایران توسط دهقانی و همکاران هنجاریابی شده و نتایج همسانی درونی به روش کودررپیچادرسون ۲۰ از ۰/۴۳ تا ۰/۸۹ متغیر بود؛ ضرایب همبستگی برای روایی همگرا در محدوده‌ی ۰,۲۹ تا ۰,۶۳ و برای روایی افتراقی در محدوده‌ی ۰,۰۹ تا ۰,۵۷ گزارش شده است (دهقانی و همکاران، ۱۳۸۹). همچنین آلفای کرونباخ برای کل مقیاس‌های روایی و بالینی در پژوهش حاضر ۰/۸۵ به دست آمد.

ج) روش اجرا: ابتدا شرکت‌کنندگان به صورت حضوری یا آنلاین به سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ (MMPI-2) پاسخ دادند. پس از بررسی پاسخنامه‌ها، در صورتی که مقیاس‌های روایی آزمون (مقیاس‌های L، F و K) در محدوده قابل قبول و معتبر قرار داشتند و نشانه‌ای از پاسخ‌دهی نامعتبر یا تحریف‌شده مشاهده نمی‌شد، با شرکت‌کننده تماس گرفته می‌شد تا برای اجرای مرحله دوم یعنی اجرای آزمون اندریافت موضوع (TAT) دعوت شود. آزمون اندریافت موضوع در شرایط حضوری و کنترل‌شده در کلینیک رویش واقع در شهر تهران اجرا شد. مجوزهای لازم برای اجرای آزمون در این کلینیک از قبل اخذ شده بود. پیش از آغاز آزمون، توضیحات لازم درباره هدف پژوهش، محرمانگی اطلاعات و ضبط صدا و نیز نحوه پاسخ‌دهی، به شرکت‌کننده ارائه شد. سپس آزمون اجرا و پاسخ‌های کلامی شرکت‌کنندگان ضبط گردید. فرآیند جمع‌آوری داده‌ها حدود سه ماه به طول انجامید. پس از اجرای آزمون، پاسخ‌ها در اختیار دو ارزیاب مستقل قرار گرفت تا به صورت جداگانه نمره‌گذاری شوند. جهت اطمینان از همسانی و دقت ارزیابی‌ها، از دو ماه پیش از شروع نمره‌گذاری‌ها، پروتکل مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی توسط ارزیاب‌ها به دقت مطالعه شد و آنان تمرین‌های لازم را برای آشنایی عملی با این نظام نمره‌گذاری انجام دادند. این تمرین‌ها در فایل پروتکل موجود بوده و بدین ترتیب است که نمونه

^۱. blind rating

داستان‌های تعریف شده در مورد کارت‌های آزمون اندریافت موضوع آمده است. ارزیاب باید داستان‌ها را به دقت مطالعه کرده و بر اساس دستورالعمل راهنما، نمره‌گذاری کند؛ سپس به نمره‌گذاری‌های صورت گرفته در انتهای راهنما مراجعه کرده و نمره‌های خود را با آن مقایسه کرده تا بتواند کم‌کم نمره‌گذاری خود را به نمره‌گذاری دقیق و درست نزدیک نماید. در نهایت نمره‌گذاری محتوای روایت‌های آزمون اندریافت موضوع بر اساس ابعاد مختلف مدل مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی توسط دو ارزیاب به صورت مستقل انجام شد. ارزیاب‌ها بدون آگاهی از سایر اطلاعات شرکت‌کنندگان (از جمله مشخصات دموگرافیک و نتایج سیاهه شخصیتی چندوجهی مینه‌سوتا-۲) به تحلیل روایت‌های آزمون پرداختند تا از هرگونه سوگیری شناختی جلوگیری شود. این روش به عنوان نمره‌گذاری کور^۱ شناخته می‌شود.

تمامی شرکت‌کنندگان پیش از شروع پژوهش، رضایت آگاهانه کتبی خود را ارائه کردند و پژوهش کاملاً مطابق با اصول اخلاقی، با رعایت محرمانگی و حفاظت از حقوق شرکت‌کنندگان انجام شد. داده‌ها با استفاده از آزمون‌های T² هاتلینگ، ضریب همبستگی پیرسون، ضریب آلفای کرونباخ و ضریب همبستگی درون طبقه‌ای و از طریق SPSS-26 تحلیل شدند.

یافته‌ها

در این پژوهش در مجموع ۱۵۴ نفر مشارکت داشتند. سن شرکت‌کنندگان بین ۱۸ تا ۵۱ سال با میانگین ۳۲/۵۵ و انحراف استاندارد ۵/۸۶ بود. ۲۷/۳ درصد (۴۲ نفر) مشارکت‌کنندگان مرد و ۷۲/۷ درصد (۱۱۲ نفر) زن بودند. از نظر وضعیت تأهل هم ۸۴ نفر (۵۴/۵ درصد) مجرد، ۴۹ نفر (۳۱/۸ درصد) متأهل و ۲۱ نفر (۱۳/۶ درصد) هم مطلقه بودند.

برای تعیین روایی سازه آزمون اندریافت موضوع (TAT) در تمایزگذاری بین گروه‌های بالینی و غیربالینی افراد نمونه به دو گروه تقسیم شدند. گروه اول شامل افرادی بود که در هیچ کدام از مقیاس‌های سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ نمره بالای ۷۰ کسب نکرده بودند و تحت عنوان گروه غیربالینی شناسایی شدند؛ اما گروه دوم شامل افرادی بود که حداقل در یکی از مقیاس‌های بالینی سیاهه شخصیت چندوجهی مینه‌سوتا-۲ نمره

بالتر از ۷۰ کسب کرده بودند و تحت عنوان گروه بالینی شناسایی شدند. میانگین هشت مقیاس، مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی برای روایت های ۲۰ کارت محاسبه شد. همچنین میانگین نمرات دو ارزیاب محاسبه شد. سپس آزمون T^2 هاتلینگ برای هشت

مقیاس، مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی گرفته شد تا مشخص شود که آیا میانگین نمرات مقیاس ها بین دو گروه غیر بالینی و بالینی تفاوت معناداری با هم دارد یا خیر. نتایج در جداول ۱ و ۲ آمده است.

جدول ۱. نتایج آزمون T^2 هاتلینگ برای مقایسه مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی بین گروه بالینی و غیر بالینی

آزمون	مقدار	F	درجه آزادی فرضیه	درجه آزادی خطا	p-value
اثر پیلاي ۱	۰/۲۰	۴/۵۶	۸/۰۰۰	۱۴۵/۰۰۰	۰/۰۰۰
لامبدای ویلکس ۲	۰/۸۰	۴/۵۶	۸/۰۰۰	۱۴۵/۰۰۰	۰/۰۰۰
اثر هاتلینگ ۳	۰/۲۵	۴/۵۶	۸/۰۰۰	۱۴۵/۰۰۰	۰/۰۰۰
بزرگترین ریشه روی ۴	۰/۲۵	۴/۵۶	۸/۰۰۰	۱۴۵/۰۰۰	۰/۰۰۰

نتایج حاصل از آزمون T^2 هاتلینگ در جدول ۱ نشان می دهد ترکیب مقیاس های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی بین دو گروه بالینی و غیر بالینی تفاوت معنادار دارد ($F=۴/۵۶$, $p<۰/۰۰۱$). لازم به ذکر است پیش فرض همسانی کوواریانس که توسط آزمون ام-باکس^۵ صورت گرفت حاکی از نقض این پیش فرض

بود ($F=۴/۳۱$, $p<۰/۰۰۱$). اگرچه همه آماره های آزمون شامل اثر پیلايی، لامبدای ویلکس، اثر هاتلینگ و بزرگترین ریشه روی تفاوت معناداری را نشان دادند ($p<۰/۰۰۱$)؛ اما با توجه به مقاومت آماره اثر پیلايی در برابر نامتوازن بودن گروه ها، تفسیر نتایج بر اساس این شاخص انجام گرفت.

جدول ۲. اثرات بین گروهی آزمون T^2 هاتلینگ برای مقیاس های SCORS-G برای روایت های TAT در دو گروه غیر بالینی و بالینی

مقیاس	گروه ها	تعداد	میانگین	انحراف معیار	F	p-value
پیچیدگی بازنمایی های افراد (COM)	غیر بالینی	۱۲۲	۳/۸۴	۰/۶۰	۰/۵۶	۰/۴۵
	بالینی	۳۲	۳/۷۴	۰/۶۳		
کیفیت عاطفی بازنمایی ها (AFF)	غیر بالینی	۱۲۲	۳/۷۷	۰/۵۰	۲۲/۷۸	۰/۰۰۰
	بالینی	۳۲	۳/۳۱	۰/۴۳		
سرمایه گذاری هیجانی در روابط (EIR)	غیر بالینی	۱۲۲	۳/۹۹	۰/۵۵	۱۲/۳۳	۰/۰۰۱
	بالینی	۳۲	۲/۶۱	۰/۵۶		
سرمایه گذاری هیجانی در ارزش ها و استانداردهای اخلاقی (EIM)	غیر بالینی	۱۲۲	۳/۸۸	۰/۲۰	۲/۲۲	۰/۱۴
	بالینی	۳۲	۳/۸۱	۰/۲۷		
درک علیت اجتماعی (SC)	غیر بالینی	۱۲۲	۳/۵۶	۰/۶۸	۱/۹۴	۰/۱۷

4. Roy's Largest Root

5. M-Box

1. Pillai's Trace

2. Wilks' Lambda

3. Hotelling's Trace

مقیاس	گروه‌ها	تعداد	میانگین	انحراف معیار	F	p-value
تجربه و مدیریت تکانه‌های خشم (AGG)	بالینی	۳۲	۳/۳۸	۰/۵۴	۸/۳۴	۰/۰۰۴
	غیربالینی	۱۲۲	۳/۸۰	۰/۲۲		
	بالینی	۳۲	۳/۶۷	۰/۲۳		
	غیربالینی	۱۲۲	۳/۷۷	۰/۴۳		
عزت نفس (SE)	بالینی	۳۲	۳/۴۷	۰/۴۴	۱۱/۹۷	۰/۰۰۱
	غیربالینی	۱۲۲	۳/۸۹	۰/۴۶		
	بالینی	۳۲	۳/۵۳	۰/۴۸		
	غیربالینی	۱۲۲	۳/۸۹	۰/۴۶		
هویت و انسجام خود (ICS)	بالینی	۳۲	۳/۵۳	۰/۴۸	۱۵/۰۰	۰/۰۰۰
	غیربالینی	۱۲۲	۳/۸۹	۰/۴۶		
	بالینی	۳۲	۳/۵۳	۰/۴۸		
	غیربالینی	۱۲۲	۳/۸۹	۰/۴۶		

جدول اثرات بین گروهی نشان داد تعداد پنج مقیاس شامل کیفیت عاطفی بازنمایی‌ها (AFF)، سرمایه‌گذاری هیجانی در روابط (EIR)، تجربه و مدیریت تکانه‌های خشم (AGG)، عزت نفس (SE) و هویت و انسجام خود (ICS) معنادار بودند ($P < ۰/۰۰۱$)؛ یعنی میانگین نمرات دو گروه افراد بالینی و غیربالینی در این مقیاس‌ها با هم تفاوت معنادار داشت. در سایر مقیاس‌ها تفاوت معناداری بین دو گروه مشاهده نشد ($P > ۰/۰۵$).

برای بررسی روایی همگرا از روش ضریب همبستگی پیرسون بین مقیاس‌های MMPI-2 و مقیاس‌های شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی ماتریس همبستگی تشکیل شد که در جدول ۳ نشان داده شده است. لازم به ذکر است ابتدا از نمرات مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی به دست آمده توسط دو ارزیاب میانگین گرفته شد و سپس همبستگی آن با سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ محاسبه شد.

جدول ۳. همبستگی بین مقیاس‌های MMPI-2 با میانگین نمرات مقیاس‌های SCORS-G برای روایت‌های TAT

ICS	SE	AGG	SC	EIM	EIR	AFF	COM	
۰/۱۶*	۰/۱۱	۰/۱۸*	-۰/۱۷*	۰/۰۶	۰/۰۵	۰/۱۴	-۰/۱۴	L
-۰/۰۹	-۰/۱۷*	-۰/۲۵**	-۰/۰۶	-۰/۰۸	-۰/۲۱**	-۰/۱۷*	-۰/۰۵	F
۰/۳۳**	۰/۲۵**	۰/۱۱	۰/۳۰**	۰/۰۷	۰/۳۷**	۰/۳۴**	۰/۲۳*	K
-۰/۱۹*	-۰/۱۳	-۰/۱۲	-۰/۰۸	-۰/۰۹	-۰/۱۳	-۰/۲۱*	-۰/۱۳	Hs
۰/۰۹	۰/۰۳	-۰/۱۳	-۰/۲۷*	۰/۲۰*	-۰/۱۳	۰/۰۴	-۰/۲۵**	D
۰/۲۶**	۰/۱۳	۰/۰۲	-۰/۱۴	۰/۱۴	-۰/۰۴	۰/۱۱	-۰/۱۴	Hy
-۰/۱۵	-۰/۳۱**	-۰/۱۴	-۰/۱۷*	-۰/۱۲	-۰/۴۳**	-۰/۲۳**	-۰/۱۸*	Pd
-۰/۰۳	-۰/۱۳	۰/۰۹	-۰/۲۲**	-۰/۱۳	-۰/۳۳**	-۰/۰۸	-۰/۲۰*	Pa
-۰/۱۱	-۰/۰۸	-۰/۰۳	-۰/۲۰**	۰/۰۶	-۰/۳۱**	-۰/۰۶	-۰/۱۶	Pt
-۰/۰۵	-۰/۰۷	-۰/۰۳	-۰/۱۴	۰/۲۰*	-۰/۲۰*	-۰/۰۶	-۰/۰۷	Sc
-۰/۲۹**	-۰/۴۱**	-۰/۲۶**	-۰/۳۰**	-۰/۰۹	-۰/۴۹**	-۰/۴۰**	-۰/۲۸**	Ma

***: $p \leq 0.01$; **: $p \leq 0.05$; *: $p \leq 0.05$ ؛ پیش‌بینی بازنمایی‌های افراد (COM)، کیفیت عاطفی بازنمایی‌ها (AFF)، سرمایه‌گذاری هیجانی در روابط (EIR)، سرمایه‌گذاری هیجانی در ارزش‌ها و استانداردهای اخلاقی (EIM)، درک علیت اجتماعی (SC)، تجربه و مدیریت تکانه‌های خشم (AGG)، عزت نفس (SE)، هویت و انسجام خود (ICS)، مقیاس دروغ‌پردازی (L)، مقیاس تظاهر به بدی (F)، مقیاس انکار (K)، مقاومت خودبیمارانگاری (Hs)، افسردگی (D)، هیستری (H)، انحراف روانی-اجتماعی (Pd)، پارانو یا (Pa)، ضعف روانی (Pt)، اسکیزوفرنی (Sc)، مانیا (Ma) و درون‌گرایی اجتماعی (Si).

نتایج نشان می‌دهد همه مقیاس‌های آزمون سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ حداقل در یک مورد با مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی همبستگی معنادار دارند. این همبستگی‌ها از مقدار $(r=0/16; \leq 0/05)$ برای همبستگی مقیاس‌های دروغ پردازی (L) با هویت و انسجام خود (ICS) تا مقدار $(r=0/49; P \leq 0/01)$ برای مقیاس‌های مانیا (Ma) با سرمایه‌گذاری هیجانی در روابط (EIR) متغیر بود. از منظر دیگر مقیاس مانیا (Ma) با ۷ مقیاس مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی همبستگی معنادار داشت که از این نظر بیشتر از سایر مقیاس‌های سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ بود. پس از آن مقیاس انکار (K) با ۶ مقیاس مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی همبستگی معنادار و مقیاس انحراف روانی اجتماعی (Pd) با پنج مقیاس همبستگی معنادار رتبه‌های بعدی را کسب کردند.

برای بررسی پایایی آزمون از روش ضریب همبستگی درون طبقه‌ای (ICC) استفاده شد. برای این منظور ابتدا میانگین نمرات هر مقیاس برای داستان‌های ۲۰ کارت به ازای هر نفر محاسبه شد؛ بنابراین برای هر شرکت‌کننده ۸ نمره به دست آمد که میانگین نمرات ۲۰ داستان بازگو شده بودند که یک بار توسط ارزیاب اول و بار دوم توسط ارزیاب دوم نمره‌گذاری شده بودند. سپس ضریب همبستگی درون طبقه‌ای برای دو جفت نمره به دست آمده محاسبه شد که نتایج آن در جدول ۴ آمده است. کمترین ضریب همبستگی درون طبقه‌ای مربوط به مقیاس هویت و انسجام خود (ICS) بود که مقدار $0/77$ به دست آمده و بیشترین مقدار برای مقیاس سرمایه‌گذاری هیجانی در روابط (EIR) بود که مقدار $0/92$ حاصل شد. بر اساس منابع موجود ضرایب کمتر از $0/5$ ضعیف، ضرایب بین $0/5$ تا $0/75$ متوسط، ضرایب بین $0/75$ تا $0/90$ در محدوده خوب و ضرایب بالاتر از $0/90$ در محدوده عالی ارزیابی می‌شود (کو و لی، ۲۰۱۶).

جدول ۴. نتایج ضریب همبستگی درون طبقه‌ای (ICC) برای مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی (SCORS-G)

نتیجه	ICC (average measures)	مقیاس
پایایی بین ارزیاب‌ها بالاست	$0/84$	COM
پایایی بین ارزیاب‌ها بالاست	$0/84$	AFF
پایایی بین ارزیاب‌ها بالاست	$0/92$	EIR
پایایی بین ارزیاب‌ها بالاست	$0/86$	EIM
پایایی بین ارزیاب‌ها بالاست	$0/87$	SC
پایایی بین ارزیاب‌ها بالاست	$0/83$	AGG
پایایی بین ارزیاب‌ها بالاست	$0/79$	SE
پایایی بین ارزیاب‌ها بالاست	$0/77$	ICS
پایایی بین ارزیاب‌ها بالاست	$0/88$	کل

پیچیدگی بازنمایی‌های افراد (COM)، کیفیت عاطفی بازنمایی‌ها (AFF)، سرمایه‌گذاری هیجانی در روابط (EIR)، سرمایه‌گذاری هیجانی در ارزش‌ها و استانداردهای اخلاقی (EIM)، درک علت اجتماعی (SC)، تجربه و مدیریت تکانه‌های خشم (AGG)، عزت‌نفس (SE)، هویت و انسجام خود (ICS)

مدیریت تکانه‌های خشم (AGG) تا $0/95$ (مقیاس شناخت اجتماعی -SC) متغیر بود و آلفای کرونباخ برای کل مقیاس‌ها $0/97$ بود. بر این اساس همسانی درونی برای دو مورد از مقیاس‌ها (سرمایه‌گذاری هیجانی در ارزش‌ها و استانداردهای اخلاقی (EIM) و تجربه و مدیریت تکانه‌های خشم

همچنین برای بررسی پایایی آزمون از ضریب همسانی درونی (آلفای کرونباخ) استفاده شد که نتایج آن در جدول ۵ آمده است. همان‌گونه که ملاحظه می‌شود آلفای کرونباخ برای مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی از $0/64$ (مقیاس تجربه و

(AGG) متوسط، برای چهار مقیاس (کیفیت عاطفی بازنمایی ها (AFF)، سرمایه گذاری هیجانی در روابط (EIR)، عزت نفس (SE)، هویت و انسجام خود (ICS) خوب و در نهایت برای دو مقیاس (پیچیدگی بازنمایی های

افراد (COM) و درک علیت اجتماعی (SC)، عالی ارزیابی شد (گلایم و گلایم، ۲۰۰۳).

جدول ۵. نتایج ضریب آلفای کرونباخ برای مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی (SCORS-G)

مقیاس	تعداد آیتم	آلفای کرونباخ	تفسیر
پیچیدگی بازنمایی های افراد (COM)	۲۰	۰/۹۴	عالی
کیفیت عاطفی بازنمایی ها (AFF)	۲۰	۰/۸۵	خوب
سرمایه گذاری هیجانی در روابط (EIR)	۲۰	۰/۸۵	خوب
سرمایه گذاری هیجانی در ارزش ها و استانداردهای اخلاقی (EIM)	۲۰	۰/۶۹	ضعیف
درک علیت اجتماعی (SC)	۲۰	۰/۹۵	عالی
تجربه و مدیریت تکانه های خشم (AGG)	۲۰	۰/۶۴	ضعیف
عزت نفس (SE)	۲۰	۰/۸۳	خوب
هویت و انسجام خود (ICS)	۲۰	۰/۸۸	خوب
کل	۱۶۰	۰/۹۷	عالی

بحث و نتیجه گیری

پژوهش حاضر با هدف اعتباریابی آزمون اندریافت موضوع (TAT) به روش نمره گذاری مقیاس کلی شناخت اجتماعی و روابط موضوعی (SCORS-G) در جمعیت ایرانی صورت گرفت.

نتایج مربوط به آزمون روایی گروه های شناخته شده نشان داد که پنج مقیاس شامل کیفیت عاطفی بازنمایی ها (AFF)، سرمایه گذاری هیجانی در روابط (EIR)، تجربه و مدیریت تکانه های خشم (AGG)، عزت نفس (SE) و هویت و انسجام خود (ICS) به طور معناداری تفاوت میان گروه های افراد بالینی و غیر بالینی را نشان دادند؛ در حالی که مقیاس های پیچیدگی بازنمایی های افراد (COM)، سرمایه گذاری هیجانی در ارزش ها و استانداردهای اخلاقی (EIM) و درک علیت اجتماعی (SC) تفاوت معناداری بین دو گروه بالینی و غیر بالینی نشان ندادند؛ به عبارت دیگر همه مقیاس های مؤلفه عاطفی به جز مقیاس سرمایه گذاری هیجانی در ارزش ها و استانداردهای اخلاقی (EIM)، توانستند بین دو گروه تفاوت معنادار ایجاد کنند و مؤلفه شناختی شامل دو مقیاس پیچیدگی بازنمایی های افراد (COM) و درک علیت اجتماعی (SC) تفاوت معناداری بین دو گروه ایجاد نکرد. در مقایسه با پژوهش های قبلی در چند مقیاس تفاوت وجود داشت. در پژوهش بووت و همکاران (۲۰۲۰) همه مقیاس ها به جز دو مقیاس کیفیت عاطفی بازنمایی ها (AFF) و سرمایه گذاری هیجانی در روابط (EIR) بین دو گروه بالینی و غیر بالینی تفاوت معنادار ایجاد کردند و در پژوهش برم

(۲۰۱۴) همه مقیاس ها برای نمونه بالینی، میانگین پایین تر نسبت به نمونه های غیر بالینی سایر پژوهش ها نشان دادند. یک دلیل احتمالی برای این تفاوت ها می تواند ناشی از تفاوت نمونه مورد بررسی باشد. به نحوی که در آن پژوهش ها نمونه های بالینی شامل بیماران بستری در بیمارستان بودند، در حالی که در پژوهش حاضر، گروه بالینی به افرادی اطلاق شد که نمره سیاهه شخصیتی چندوجهی مینه سوتا-۲ فرم ۷۱ پرستی آن ها بالاتر از ۷۰ بود. علت دیگر می تواند ناشی از کارت های آزمون اندریافت موضوع (TAT) مورد استفاده باشد؛ به نحوی که در پژوهش های مذکور از دسته ۵ تایی و ۹ تایی از کارت های پر کاربرد استفاده شد؛ در حالی که در این پژوهش همه ۲۰ کارت استفاده شد. دلایل احتمالی دیگر می تواند ناشی از عدم حساسیت برخی مقیاس ها به ویژگی های گروه یا محدودیت های نمونه باشد. این یافته ها با نتایج سایر مطالعات تعیین روایی نیز همسو هستند. به طور خاص، مشابه پژوهش هیبرد و همکاران (۱۹۹۵) در زمینه روایی سازه و پژوهش آکرمن و همکاران (۲۰۰۱) در زمینه روایی همگرا، نتایج نشان داد که مؤلفه های عاطفی ابزار نسبت به مؤلفه های شناختی روایی بالاتری دارند. این همسویی با مطالعات پیشین نشان می دهد که ابزار مورد استفاده توانایی مناسبی در ارزیابی ابعاد عاطفی ویژگی های روان شناختی دارد و اعتبار آن در اندازه گیری روان شناختی، به ویژه در ابعاد عاطفی، مورد تأیید قرار می گیرد.

بررسی روایی همگرا نشان داد که برخی مقیاس‌های سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ با مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی همبستگی معنادار داشتند. از مجموع ۸۸ روابط ممکن، ۳۸ مورد از آن‌ها همبستگی معنادار نشان دادند. تحلیل همبستگی‌ها نشان داد که برخی مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی مانند سرمایه‌گذاری هیجانی در روابط (EIR) و درک علت اجتماعی (SC) با تعداد بیشتری (هفت مورد) از مقیاس‌های سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ همبستگی معنادار داشتند، در حالی که برخی مقیاس‌های دیگر مانند سرمایه‌گذاری هیجانی در ارزش‌ها و استانداردهای اخلاقی (EIM) و تجربه و مدیریت تکانه‌های خشم (AGG) تنها با دو یا سه مقیاس همبستگی نشان دادند. در مقایسه با پژوهش‌های پیشین، یافته‌های پژوهش آکرمین و همکاران (۲۰۰۱) حاکی از همبستگی‌های متوسط معنادار مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی با مقیاس‌های نمره‌گذاری آزمون رورشاخ (MOA) بود و یافته‌های هگرتی و همکاران (۲۰۱۵) هم حاکی از همبستگی‌های متوسط مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی با ارزیابی کارکرد کلی (GAF) بود.

در پژوهش حاضر نیز هم‌راستا با پژوهش‌های پیشین نتایج حاکی از همبستگی‌های متوسط معنادار در برخی از مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی با ابزارهای مشابه است که در اینجا سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ بود. این الگو بیانگر آن است که برخی مقیاس‌های مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی دامنه روان‌شناختی گسترده‌تری را پوشش می‌دهند و توانایی بیشتری در بازتاب ویژگی‌های مختلف روان‌شناختی افراد دارند، در حالی که برخی مقیاس‌ها ممکن است به ابعاد خاص و محدودتری متمرکز باشند. یافته‌ها همچنین تأکید می‌کنند که ابزارهای روان‌شناختی متفاوت، هرچند همسو در برخی ابعاد، می‌توانند حساسیت متفاوتی نسبت به ویژگی‌های خاص نشان دهند و این نکته اهمیت توجه به ترکیب مناسب مقیاس‌ها را در تحلیل و تفسیر نتایج برجسته می‌کند. به طور کلی، این نتایج نشان می‌دهند که مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی و سیاهه شخصیتی چندوجهی مینه‌سوتا-۲ در برخی ابعاد همگرایی قابل توجهی دارند و هم‌زمان

ویژگی‌های منحصر به فرد هر ابزار می‌تواند اطلاعات تکمیلی ارزشمندی برای ارزیابی روان‌شناختی فراهم کند. بررسی پایایی بین ارزیاب‌ها با استفاده از ضریب همبستگی درون طبقه‌ای (ICC) نشان داد که توافق نمره‌گذاران در سطح مناسبی قرار دارد و ابزار توانایی بالایی در تولید نتایج پایا و قابل اعتماد دارد. کمترین و بیشترین میزان توافق بین مقیاس‌ها متفاوت بود که این تفاوت‌ها می‌تواند ناشی از ویژگی‌های خاص هر مقیاس یا پیچیدگی محتوای آن باشد. این یافته‌ها با مطالعات پیشین همسو هستند که همگی پایایی بین ارزیاب‌ها را مناسب ارزیابی کردند (اینسلیجرز و همکاران (۲۰۱۲)، سیفرت و همکاران (۲۰۱۷)، استین و همکاران (۲۰۱۲)، رایدنور و همکاران (۲۰۲۲)، اله دادی و همکاران (۲۰۲۵)). به طور کلی، توافق مناسب بین ارزیاب‌ها نشان می‌دهد که روش نمره‌گذاری، معیارهای روشن و استاندارد دارد و می‌تواند در کاربردهای بالینی و پژوهشی به عنوان معیاری معتبر مورد استفاده قرار گیرد. علاوه بر این، تفاوت‌های نسبی بین مقیاس‌ها می‌تواند راهنمایی برای اصلاح و بهبود ابزار، آموزش ارزیابان و تمرکز پژوهش‌های بعدی باشد تا حساسیت و دقت ارزیابی در تمام ابعاد افزایش یابد. این نتایج همچنین اهمیت توجه به ویژگی‌های فردی و فرهنگی نمونه‌ها را در اجرای آزمون و تفسیر نتایج تأکید می‌کنند و نشان می‌دهند که پایایی ابزار علاوه بر اعتبار روان‌شناختی، نقش کلیدی در کاربردهای عملی و پژوهشی دارد.

برای بررسی پایایی همسانی درونی آزمون از ضریب آلفای کرونباخ استفاده شد که نشان داد انسجام داخلی مقیاس‌ها در سطح قابل قبول تا عالی قرار دارد. برخی مقیاس‌ها همسانی درونی متوسطی داشتند، در حالی که چند مقیاس دیگر همسانی خوبی و دو مقیاس نیز انسجام درونی عالی نشان دادند. همچنین همسانی کلی مقیاس بسیار عالی گزارش شد. این یافته‌ها با نتایج پژوهش‌های اینسلیجرز و همکاران (۲۰۱۲) و هیبرد و همکاران (۲۰۱۰) همسو نبود؛ زیرا آنان همسانی درونی را برای مقیاس‌های مؤلفه عاطفی پایین و برای مقیاس‌های مؤلفه شناختی در سطح مطلوب گزارش کرده بودند. به نظر می‌رسد بخشی از این تفاوت‌ها ناشی از تفاوت در روش اجرا باشد؛ به‌ویژه آنکه در پژوهش‌های یادشده از تعداد کارت‌های کمتری استفاده شده و این امر می‌تواند بر میزان پوشش و تنوع پاسخ‌ها تأثیر گذاشته باشد. همچنین، ویژگی‌های نمونه‌های پژوهش از جمله ترکیب بالینی و غیربالینی و نیز تفاوت‌های فرهنگی میان جوامع مورد مطالعه، از عواملی

نمونه‌گیری تصادفی و حجم نمونه بزرگ‌تر را ندارد. به همین دلیل، در پژوهش‌های آتی می‌توان با نمونه‌های وسیع‌تر و نمونه‌گیری تصادفی، این ابزار را بررسی کرد و علاوه بر پایایی و روایی، ساختار عاملی مقیاس‌ها را نیز استخراج نمود تا شواهد روش‌شناختی جامع‌تری فراهم شود. در نهایت، آزمون اندریافت موضوع (TAT) با نمره‌گذاری بر اساس مقیاس کلی شناخت اجتماعی و روابط موضوعی (SCORS-G)، از روایی و پایایی مناسبی در جمعیت ایرانی برخوردار است و استفاده از آن به‌ویژه در محیط‌های بالینی و پژوهشی برای سنجش مؤلفه‌های شناختی و عاطفی در بزرگسالان توصیه می‌شود.

ملاحظات اخلاقی

پیروزی از اصول اخلاقی پژوهش: این پژوهش برگرفته از رساله دکتری نویسنده اول در رشته سنجش و اندازه‌گیری در دانشکده روانشناسی و علوم تربیتی دانشگاه علامه طباطبائی است که پروپوزال آن در تاریخ ۱۴۰۲/۱۰/۲۵ دفاع شده است. به جهت رعایت اصول اخلاقی پیش از آغاز آزمون‌ها، هدف پژوهش، شیوه اجرا، نحوه ضبط صدا و حقوق مشارکت‌کنندگان به‌طور کامل برای آن‌ها توضیح داده شد. رضایت آگاهانه کتبی از تمامی مشارکت‌کنندگان دریافت شد و تأکید شد که مشارکت کاملاً داوطلبانه است و در هر مرحله می‌توانند بدون هیچ پیامد منفی از ادامه همکاری انصراف دهند. برای حفظ حریم خصوصی و محرمانگی اطلاعات، نام و مشخصات فردی مشارکت‌کنندگان در هیچ مرحله‌ای از گزارش‌دهی و انتشار نتایج ذکر نشد. داده‌های جمع‌آوری شده صرفاً برای اهداف پژوهشی استفاده شده و در دسترس افراد یا نهادهای غیرمرتبط قرار نخواهد گرفت. همچنین، در صورت تمایل مشارکت‌کنندگان، نتایج آزمون‌ها و یافته‌های پژوهش در اختیار آن‌ها قرار خواهد گرفت.

حامی مالی: این پژوهش در قالب رساله دکتری و بدون حمایت مالی می‌باشد.

نقش هر یک از نویسندگان: این مقاله از رساله دکتری نویسنده اول و به راهنمایی نویسندگان دوم و سوم و مشاوره نویسنده چهارم استخراج شده است.

تضاد منافع: نویسندگان همچنین اعلام می‌دارند که در نتایج این پژوهش هیچ گونه تضاد منافی وجود ندارد.

تشکر و قدردانی: پژوهشگران مایل‌اند صمیمانه از تمامی افرادی که در این پژوهش شرکت کرده‌اند، قدردانی نمایند.

هستند که می‌توانند در الگوی ضرایب پایایی اثرگذار باشند. از طرفی این یافته‌ها با یافته‌های اله‌دادی و همکاران (۲۰۲۵) همسو بود. نتایج آن پژوهش، حاکی از تأیید پایایی به روش همسانی درونی ابزار در پژوهش‌های مختلف بود. در خصوص تفاوت نسبی در سطح همسانی درونی بین مقیاس‌ها شایان ذکر است که این تفاوت‌ها می‌تواند راهنمایی برای بازنگری و بهبود مقیاس‌ها در پژوهش‌های آینده باشد تا انسجام و دقت اندازه‌گیری در تمامی ابعاد افزایش یابد و ابزار در کاربردهای بالینی و پژوهشی به شکلی قابل اعتماد مورد استفاده قرار گیرد.

یکی از ملاحظات مهم در تفسیر نتایج، مربوط به شیوه نمره‌گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی است. در این نظام، در شرایطی که محتوای مرتبط با یک مقیاس خاص، در روایت وجود نداشته باشد، از نمره‌گذاری پیش‌فرض^۱ استفاده می‌شود؛ بدین معنا که برای آن مقیاس، نمره‌ای ثابت در نظر گرفته می‌شود. اگرچه این رویکرد از منظر افزایش پایایی بین ارزیاب‌ها و کاهش اختلافات فردی در نمره‌گذاری حائز اهمیت است، اما می‌تواند محدودیت‌هایی نیز به همراه داشته باشد. به‌طور خاص، استفاده از نمره‌های ثابت ممکن است به کاهش واریانس واقعی داده‌ها منجر شود و در نتیجه بر نتایجی همچون ضرایب همبستگی یا شاخص‌های همسانی درونی اثر بگذارد. از این منظر، بخشی از الگوی همبستگی‌ها یا پایایی مشاهده‌شده ممکن است متأثر از این قاعده‌ی نمره‌گذاری باشد و نه صرفاً بازتابی از ویژگی‌های روان‌شناختی آزمودنی‌ها؛ بنابراین، لازم است در تفسیر یافته‌ها این ملاحظه به عنوان یک محدودیت روش‌شناختی مد نظر قرار گیرد.

از جمله محدودیت‌های پژوهش حاضر، نیاز به دانش و مهارت کافی در نمره‌گذاری داستان‌ها به روش مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره‌گذاری کلی بود. هرچه ارزیاب‌ها دفعات بیشتری آزمون را اجرا کنند، مهارتشان در نمره‌گذاری افزایش می‌یابد و دقت نتایج بالاتر می‌رود. با توجه به اینکه در این پژوهش تنها حدود دو ماه فرصت برای یادگیری و تمرین ارزیاب‌ها فراهم بود، ممکن است محدودیت مهارت نسبی بر برخی نتایج تأثیر گذاشته باشد. علاوه بر محدودیت‌های یاد شده، محدودیت‌های عملی و منابع پژوهش مانند زمان مورد نیاز برای اجرا و نمره‌گذاری آزمون اندریافت موضوع و امکانات موجود، اجازه استفاده از

^۱. default rating

References

- Alahdadi, S., Minaei, A., Falsafinejad, MR., & Atef-Vahid, MK. (2025). A systematic review on the validation of the thematic apperception test (TAT) with the scoring method of the Global Scale of Social Cognition and Object Relations (SCORS-G). [in Persian] *Contemporary Psychology*, (in press).
- Aronow, E., Altman Weiss, K., & Reznikoff, M. (2001). *A Practical Guide to the Thematic Apperception Test: The TAT in Clinical Practice*. Routledge. <https://doi.org/10.4324/9780203766330>
- Ackerman, S. J., Hilsenroth, M. J., Clemence, A. J., Weatherill, R., & Fowler, J. C. (2001). Convergent validity of Rorschach and TAT scales of object relations. *Journal of Personality Assessment*, 77(2), 295-306. doi.org/10.1207/S15327752JPA7702_11
- Bahrami, H. (2014). *Psychological tests: Theoretical foundations and applied techniques* [In Persian]. Tehran: *Allameh Tabataba'i University Press*. https://book.atu.ac.ir/book_172.html
- Bouvet, C., Prime, C., Camart, N., Fouques, D., & Zebdi, R. (2020). Discriminant Validity (Clinical Vs Non Clinical) of the French Version of the Social Cognition and Object Relation Scale, French Version, for TAT Narratives. *Quebec Review of Psychology*, 41(1), 45. doi.org/10.7202/1070662ar
- Bram, A. D. (2014). Object relations, interpersonal functioning, and health in a nonclinical sample: Construct validation and norms for the TAT SCORS-G. *Psychoanalytic Psychology*, 31(3), 314. <https://doi.org/10.1037/a0036286>
- Butcher, J. N., Graham, J. R., Ben-Porath, Y. S., Tellegen, A., Dahlstrom, W. G., & Kaemmer, B. (2001). *MMPI-2: Manual for administration, scoring, and interpretation* (Rev. ed.). *University of Minnesota Press*. www.upress.umn.edu/test-division/mmpi-2
- Dehghani, M., Ashouri, A., Habibi Asgarabad, M., & Khatibi Mirali, A. (2010). Norming and psychometric properties of the Minnesota Multiphasic Personality Inventory-2 (MMPI-2). *Journal of Psychology*, 14(3), 299-318. <https://www.sid.ir/paper/471423/fa>
- Gliem, J. A., & Gliem, R. R. (2003). Calculating, interpreting, and reporting Cronbach's Alpha reliability coefficient for Likert-type scales. *Paper presented at the Midwest Research-to-Practice Conference in Adult, Continuing, and Community Education*, Columbus, OH. <https://scholarworks.indianapolis.iu.edu/server/api/core/bitstreams/976cec6a-914f-4e49-84b2-f658d5b26ff9/content>

منابع

- اله دادی، سجاد؛ مینائی، اصغر؛ فلسفی نژاد، محمدرضا و عاطف وحید، محمد کاظم (۱۴۰۴). مروری نظام‌مند بر اعتباریابی آزمون اندریافت موضوع (TAT) با نمره گذاری مقیاس شناخت اجتماعی و روابط موضوعی - روش نمره گذاری کلی (SCORS-G). *مجله روانشناسی معاصر* (در دست چاپ)
- بهرامی، هادی. (۱۳۹۳). آزمون‌های روانی: مبانی نظری و فنون کاربردی. دانشگاه علامه طباطبائی. https://book.atu.ac.ir/book_172.html
- دهقانی، محسن؛ عاشوری، احمد؛ حبیبی، مجتبی؛ و خطیبی طباطبائی، میرعلی. (۱۳۸۹). هنجاریابی و بررسی ویژگی‌های روانسنجی سیاهه شخصیتی چندوجهی مینه سوتا. *مجله روانشناسی*، ۱۴ (پیاپی ۵۵)، ۲۹۹-۳۱۸. <https://www.sid.ir/paper/471423/fa>

- Haggerty, G., Blanchard, M., Baity, M. R., Defife, J. A., Stein, M. B., Siefert, C. J., Zodan, J. (2015). Clinical validity of a dimensional assessment of self-and interpersonal functioning in adolescent inpatients. *Journal of Personality Assessment*, 97(1), 3-12. doi:10.1080/00223891.2014.930744
- Hibbard, S., Hilsenroth, M. J., Hibbard, J. K., & Nash, M. R. (1995). A validity study of two projective object representations measures. *Psychological Assessment*, 7(4), 432. <https://doi.org/10.1037/1040-3590.7.4.432>
- Hibbard, S., Mitchell, D., & Porcerelli, J. (2001). Internal consistency of the object relations and social cognition scales for the Thematic Apperception Test. *Journal of Personality Assessment*, 77(3), 408-419. DOI:10.1207/S15327752JPA7703_03
- Inslegers, R., Vanheule, S., Meganck, R., Debaere, V., Trenson, E., & Desmet, M. (2012). Interpersonal problems and cognitive characteristics of interpersonal representations in alexithymia: a study using a self-report and interview-based measure of alexithymia. *The Journal of nervous and mental disease*, 200(7), 607-613. DOI: 10.1097/NMD.0b013e31825bfad9
- Joy, S. P. (2020). Why Teach Future Psychotherapists to Administer and Interpret Projective Tests? Well, why teach future physicians and surgeons to perform dissections? *SIS Journal of Projective Psychology & Mental Health*, 27 (1). <https://www.proquest.com/openview/17c84162801741ee599fb2034a079498/1?pq-origsite=gscholar&cbl=29082>
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of chiropractic medicine*, 15(2), 155-163. doi.org/10.1016/j.jcm.2016.02.012
- Lilienfeld, S. O. (2006). The scientific status of projective techniques. *Psychological Science in the Public Interest*, 7(2), 1-49. <https://doi.org/10.1111/j.1529-1006.2006.00029.x>
- Lundy A. (1985). The reliability of the Thematic Apperception Test. *J Pers Assess*, 49(2):141-5. doi: 10.1207/s15327752jpa4902_6
- Martin, C. R., & Hollins Martin, C. J. (2017). Minimum sample size requirements for a validation study of the Birth Satisfaction Scale-Revised (BSS-R). *Journal of Nursing and Practice*, 1(2), 25-30. <https://doi.org/10.36959/545/358>
- Morgan, C. D., & Murray, H. A. (1935). A method for investigating fantasies: The Thematic Apperception Test. *Archives of Neurology & Psychiatry*, 34(2), 289-306. DOI:10.1001/archneurpsyc.1935.02250200049005
- Narron, M. (2005). Updating the TAT: a photographic revision of the Thematic Apperception Test: Alliant International University, San Francisco Bay. https://books.google.com/books/about/Updating_the_TAT.html?id=K_4QtwAACAAJ
- Nissley, G. E., & DeFreese, E. (2020). Thematic Apperception Test. *The Wiley Encyclopedia of Personality and Individual Differences: Measurement and Assessment*, 381-385. DOI:10.1002/9781118970843.ch133
- Ridenour, J. M., Lewis, K. C., Siefert, C. J., & Stein, M. B. (2022). Longitudinal stability of social cognition and object relations scale—global rating method dimensional ratings, score ranges and narrative ‘blandness’ in a clinical sample. *Clinical psychology & psychotherapy*, 29(4), 1447-1456. DOI:10.1002/cpp.2729
- Siefert, C. J., Stein, M., Slavin-Mulford, J., Haggerty, G., Sinclair, S. J., Funke, D., & Blais, M. A. (2018). Exploring the factor structure of the Social Cognition and Object Relations—Global Rating Method: Support for two-and three-factor models. *Journal of Personality Assessment*, 100(2), 122-134. DOI:10.1080/00223891.2017.1336716
- Stein, M. B., Slavin-Mulford, J., Sinclair, S. J., Siefert, C. J., & Blais, M. A. (2012). Exploring the construct validity of the social cognition and object relations scale in a clinical sample. *Journal of Personality Assessment*, 94(5), 533-540. DOI:10.1080/00223891.2012.668594
- Stein, M., & Slavin-Mulford, J. (2017). The Social Cognition and Object Relations Scale-Global Rating Method (SCORS-G): A comprehensive guide for clinicians and researchers (1st ed.). *Routledge*. <https://doi.org/10.4324/9781315207629>
- Westen, D., Lohr, N., Silk, K. R., Gold, L., & Kerber, K. (1990). Object relations and social cognition in borderlines, major depressives, and normals: A Thematic Apperception Test analysis. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 2(4), 355. DOI: 10.1207/s15327752jpa5601_6