Archives of
Rehabilitation

2026. Vol 27. Issue 2

Research Paper

Development and Psychometric Evaluation of the Persian Receptive Figurative Language Test (P-RFLT) in Healthy Adults

Zahra Amini[1], Mehdi Purmohammad[2], Alireza Mordi[3], *Leila Ghasisin[4]

1. Institute for Cognitive Sciences, Shahid Beheshti University, Tehran, Iran.
2. Department of Psychology, MacEwan University, Edmonton, Alberta, Canada.
3. Department of Educational Psychology, University of Alberta, Edmonton, Alberta, Canada.
4. Department of Speech Therapy, Musculoskeletal Research Center, Faculty of Rehabilitation Sciences, Isfahan University of Medical Sciences, Isfahan, Iran.

Use your device to scan and read the article online

Citation Amini Z, Purmohammad M, Mordi A, Ghasisin L. Development and Psychometric Evaluation of the Persian Receptive Figurative Language Test (P-RFLT) in Healthy Adults. Archives of Rehabilitation. 2026; 27(2):198-217. https://doi.org/10.32598/RJ.27.2.141.6

https://doi.org/10.32598/RJ.27.2.141.6

ABSTRACT

Objective Figurative language is a fundamental component of linguistic–cognitive abilities that plays a critical role in the comprehension of indirect meanings, abstract thinking, and everyday communicative interactions. It pertains to an individual's capacity to interpret implied meanings and is inherently context-dependent. The comprehension of figurative language requires advanced linguistic and cognitive skills. In Persian, no comprehensive tool currently exists that can systematically assess various types of figurative language. Therefore, the present study aimed to develop the Persian receptive figurative language test (P-RFLT) and evaluate its psychometric properties in healthy Persian-speaking adults.

Materials & Methods This is a methodological study with a cross-sectional design conducted during 2023-2024. Test items were developed based on theoretical frameworks of figurative language drawn from cognitive linguistics and pragmatics, encompassing four components: metaphors, idioms, proverbs, and similes. A panel of linguistics experts evaluated face and content validity. Following item development and a pilot study, the final version of the test was administered to 337 healthy Persian-speaking adults aged 25-65 years. Construct validity was assessed by exploratory factor analysis (EFA). Internal consistency was assessed using Cronbach's α coefficient, and test re-test reliability using the intraclass correlation coefficient (ICC).

Results The items for all four components of the P-RFLT demonstrated adequate face and content validity. Construct validity was confirmed by EFA, with the Kaiser–Meyer–Olkin (KMO) measure and factor loadings indicating that each component was independently measured and capable of discriminating among participants. Cronbach's α coefficients ranged from 0.75 to 0.80, and the ICC ranged from 0.71 to 0.81, demonstrating acceptable reliability of the instrument.

Conclusion The P-RFLT has satisfactory validity and reliability and can be used as a reliable tool in clinical assessments, cognitive rehabilitation, and language-and-cognition research.

Keywords Figurative language, Psychometrics, Adults, Persian

Received: 10 Dec 2025
Revised: 25 May 2026
Accepted: 10 Jun 2026

Corresponding Author:
Leila Ghasisin, Associate Professor.
Address: Department of Speech Therapy, Musculoskeletal Research Center, Faculty of Rehabilitation Sciences, Isfahan University of Medical Sciences, Isfahan, Iran.
Tel: +98 (913) 3553753
E-Mail: ghasisin@rehab.mui.ac.ir

Copyright © 2026 The Author(s);
This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-By-NC 4.0: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

English Version

Introduction

Figurative language is a cognitive and discursive tool that operates beyond literal meaning and depends on an individual's ability to interpret implied meanings in a context-dependent manner. Understanding figurative language requires advanced linguistic and cognitive skills [1]. According to the lexical concepts and cognitive models theory (LCCM), figurative language processing results from the simultaneous interaction of linguistic knowledge and cognitive structures, whereby individuals interpret the non-literal meaning of utterances by drawing upon lexical items, syntactic structures, and prior experiential knowledge. In this framework, metaphors, idioms, proverbs, and similes are considered a shared cognitive–linguistic construct [1, 2]. According to Giora, the comprehension of figurative language depends more on the salience degree of meaning, which is shaped by an individual's experience, rather than on the distinction between literal and non-literal meaning. This salience [3]. Similarly, Lakoff and Johnson considered figurative language to be an inseparable part of everyday language and thought, arguing that humans rely on it to express abstract concepts [4].

Metaphor is one of the most prominent topics in cognitive semantics and is defined as the understanding of one conceptual domain in terms of another. For instance, in the metaphor "Love is a journey", the abstract concept of love is understood through the concrete concept of a journey [5]. Metaphors are systematic mappings between a source domain and a target domain and, beyond the linguistic level, play a pivotal role in organizing the mind's conceptual structures. They vary in terms of the degree of familiarity and novelty [6]. Conventional metaphors, such as "Time is gold", are more frequent and processed more rapidly, while novel or unconventional metaphors, such as "Zeal is lava", require more extensive cognitive processing [7, 8]. Moreover, although many conceptual metaphors are shared across cultures, language and culture play a significant role in their formation and interpretation [9, 10].

Idioms are another important manifestation of figurative language [11]. Phrases such as "Āb-ghureh nagir" (in Persian) carry a meaning that extends beyond the sum of their constituent words [12]. Research has demonstrated that an individual's familiarity with idioms plays a fundamental role in the speed and accuracy of comprehension, such that familiar idioms are understood more quickly and with greater ease [13].

Proverbs are another form of figurative language that, compared to idioms, have a more fixed syntactic structure and are employed as completely fixed expressions. In contrast, idioms can be used in terms of person and tense [14]. Proverbs reflect metaphorical structures as well as cultural and collective knowledge and are considered a type of "compressed knowledge" that conveys abstract and empirical concepts in a metaphorical form [15, 16]. Understanding familiar proverbs usually requires less cognitive processing, whereas unfamiliar proverbs require greater inferential effort [17]. Proverb-based tests can serve as effective tools for assessing cognitive and linguistic functioning [18].

Similes are another form of figurative expression characterized by an explicit comparison. In Persian, these expressions are considered a subcategory of proverbs with a culturally established, fixed structure. Unlike many proverbs, similes often create an explicit comparison between two phenomena by using words such as "like", "as", or "than". For instance, "A teacher is like a lantern in a dark night" is a simile [19]. Despite the greater structural clarity of similes, their comprehension remains dependent on inferential processes, particularly when the similarity between the two concepts is not obvious [20].

Studies in cognitive neuroscience have demonstrated that the brain hemispheres play different roles in processing figurative language. The left hemisphere is predominantly engaged in processing familiar metaphors, whereas the right hemisphere plays a more prominent role in processing novel and complex metaphors [21, 22]. This distinction is more closely related to factors such as novelty, difficulty, and the need to integrate information, rather than to the metaphorical nature of language per se [23]. The investigation of these processes is also clinically important and can help distinguish healthy individuals from those with cognitive disorders [24].

Given the critical role of figurative language in abstract thinking and human communication, its assessment in both research and clinical settings is essential. In this regard, the figurative language interpretation test (FLIT) was designed to measure the comprehension of metaphors, similes, proverbs, and personification [25]. Also, the assessment of pragmatic abilities and cognitive substrates (APACS) was developed in 2016 to evaluate pragmatic skills and figurative language, including metaphor and irony, in Italian adults [26]. In Iran, several instruments have also been designed to assess certain aspects of figurative language. Nejati and Ramesh developed a proverb test to assess executive functions and language abilities in school-aged children

[27]. Ghavami et al. prepared the Persian version of the Delis–Kaplan executive function system (D-KEFS) test battery for ages 16-40 [28]. Rahimifar et al. also designed the High-Level Language Skills Test for neurological patients and adults, which comprises seven subtests for repeating long sentences, sentence construction, inference, comprehension of complex grammatical sentences, comprehension of ambiguous sentences, word definitions, and proverbs [29].

Despite the availability of the aforementioned instruments, there is still a lack of comprehensive, standardized tests for the simultaneous assessment of metaphors, idioms, proverbs, and similes in the Persian language. This gap has limited the precise evaluation of figurative language abilities, particularly in patients with schizophrenia, traumatic brain injury, and neurological diseases such as frontotemporal dementia and amyotrophic lateral sclerosis [30, 31]. Accordingly, the present study aimed to develop and evaluate the psychometric properties of the Persian receptive figurative language test (P-RFLT) for healthy Persian-speaking adults aged 25-65 years. This age range was selected due to the influence of developmental changes prior to age 25 and the effects of the aging process after age 65 on figurative language processing [32]. By providing a comprehensive instrument for assessing figurative language, this study has broad applicability in clinical evaluations and future research.

Materials & Methods

This is a methodological study with a cross-sectional design. To develop the P-RFLT, figurative language items were first generated across four domains—metaphor, proverb, idiom, and simile—and then their psychometric properties were evaluated. The study was carried out in three phases: item development, content and face validity assessment, and final test administration.

Participants

The study population comprised 418 adults aged 25-65 years, of whom 41 participated in the item familiarity rating phase, 40 in the pilot study, and 337 in the final phase. The sample size was determined based on Morgan's table [33]. Convenience sampling was conducted over six months during 2023–2024 by attending various locations in Tehran Province, including parks, workplaces, shopping centers, and companies. The inclusion criteria were as follows: age 25-65 years, being a monolingual Persian speaker, at least a primary school education to ensure the ability to respond to the test items, and no uncorrectable visual or auditory impairments—all of which were ascertained through self-report. Additionally, participants' cognitive health was confirmed using the Persian version of the clinical dementia rating scale (CDR-P) [34].

Procedure

Phase 1: Test item development

The P-RFLT items were developed based on Cieślicka et al.'s test, which encompasses four main components of figurative language: metaphor, idiom, proverb, and simile [35]. In the initial phase, expressions pertaining to these components were extensively extracted from various sources and texts using a deductive approach. To control for confounding linguistic variables, the structural features of the items (word count, sentence length, and grammatical pattern) were kept as similar as possible, such that all items were formulated in the form of short and simple sentences with 4-7 words. Given that the number of items should ideally be 2-3 times the final required number [36], a total of 86 metaphors, 112 idioms, 48 proverbs, and 48 similes were initially selected. Metaphors and idioms constitute an extensive and fundamental part of the conceptual and lexical system of language; hence, a larger number of items were selected for these two categories [4].

Phase 2: Content and face validity assessment

In this phase, the items were submitted to 10 doctoral students in cognitive linguistics, of whom 5 were university lecturers, 3 had experience at various research institutes, and 2 had clinical experience in speech-language pathology. The panelists evaluated the items for relevance, clarity, representativeness, and alignment with the test objectives. Subsequently, the content validity ratio (CVR) was calculated using Lawshe's method, and items with a CVR below 0.62 were eliminated. Additionally, the face validity of the items was assessed with respect to clarity, comprehensibility, simplicity of wording, readability, layout, and overall organization. Based on the feedback received, modifications were made, including disambiguation, sentence simplification, and structural rearrangement of certain items. After content and face validity assessment, 28 metaphors, 27 proverbs, 31 similes, and 59 idioms were retained as test items. Subsequently, to prevent ceiling and floor effects, item difficulty was controlled for familiarity level [37]. To assess familiarity, the initial draft of the test was administered to 41 eligible adults, who were asked to rate their degree of familiarity with each item on a five-point Likert scale from "very low" to "very high". Items with

a mean score >3 were classified as familiar, and those with a mean score <3 were classified as unfamiliar [35]. Finally, based on the results of the familiarity analysis and the research team's judgment, the test was compiled, comprising 18 metaphors, 32 idioms, 15 proverbs, and 15 similes, organized sequentially from easy to difficult levels.

To assess the cognitive processing of figurative language, test items were designed in two formats:

Multiple-choice format (three options): In this format, each item was designed as a unit consisting of a stimulus sentence (containing a figurative expression) and three response options: Relevant (reflecting the figurative or expected meaning), a literal-meaning distractor, and irrelevant. Participants were required to select the correct option. Scoring was conducted on a dichotomous basis (0–1). This format was employed to assess the comprehension of metaphors, idioms, and a subset of proverbs and similes. For example, for the Persian item "Amoye man az donya raft" [My uncle left the world], the correct response was "Amoye man mord" [My uncle passed away].

Sentence completion format: A subset of items related to proverbs and similes was designed in a sentence-completion format. In this format, an incomplete sentence is given, and participants are asked to complete it with one or two words. This format was employed to assess deeper semantic processing, access to meanings stored in long-term memory, and the ability to infer semantic relationships [38]. Scoring in this section was also dichotomous (0–1), and orthographic errors were disregarded in the evaluation. For example, for the Persian proverb "Payane shabe siah…" [The end of the dark night…], the expected response was "sepid ast" [Is white] (the proverb means every cloud has a silver lining).

The developed sample items were administered in a pilot study to 40 eligible participants. Prior to administration, participants were asked to provide feedback on font size, font type, and line spacing to optimize the word readability. All participants approved the prepared draft.

Following the pilot study, the frequency and percentage of selection for each option were calculated, and the attractiveness of the distractors was evaluated. Furthermore, the correlation between the selection of each option and the total test score was examined to determine the discriminative validity of the options. Options that exhibited very low selection frequency, inadequate discriminative power, or were indistinguishable from the distractors were identified and revised, and the final draft of the test was prepared.

Phase 3: Test administration

The final draft of the test was administered to 337 participants who met the inclusion criteria to examine internal consistency, test re-test reliability, and construct validity. Test administration was conducted in a quiet environment. Initially, the purpose of the study was explained verbally to the participants, and upon agreement, they completed an informed consent form, followed by a demographic form and the CDR-P test. The prepared test was then provided to participants in a paper-and-pencil, self-administered form. To familiarize participants with how to respond, two sample items (one with a multiple-choice format and one with a sentence completion format) were placed at the beginning of each section of the test. To mitigate the participants' fatigue, the items were divided into two equal sections, with a brief rest interval between them. Upon completion of test administration, item scoring was conducted by an independent rater with no involvement in the sampling process. To examine test re-test reliability, the test was re-administered to 67 participants at a two-week interval.

Data analysis

Data entry into SPSS software, version 20 was performed by a person unaware of the sampling and scoring processes to enhance data accuracy. For data analysis, both descriptive and inferential statistics were employed. Descriptive statistics, including the Mean±SD, and median, were used to present participants' performance across the subtests. Internal consistency was assessed using Cronbach's α. Test re-test reliability was assessed by calculating the intraclass correlation coefficient (ICC), and construct validity was examined by exploratory factor analysis (EFA).

Results

Participants were 418 healthy adults aged 25-65, including 293 females (70%) and 125 males (30%). In terms of age, 18% were in the 25-34 age group, 51.6% in the 35-44 age group, 23% in the 45-54 age group, and 7.4% in the 55-65 age group. Regarding education level, 1% had lower than high school education, 16.6% had a high school diploma, 7.4% had a postgraduate degree, 47.4% had a bachelor's degree, 22.6% had a master's degree, and 5% had a PhD degree.

Table 1. The number of test items based on CVR and familiarity

Components	No.					
	Items with CVR≥0.62	Familiar Items (Mean ≥3)	Unfamiliar Items (Mean <3)	Familiar Items Selected	Unfamiliar Items Selected	Final Items Selected
Idioms	59	39	20	17	15	32
Metaphors	28	12	16	9	9	18
Proverbs	27	17	10	10	5	15
Similes	31	22	9	10	5	15

Archives of Rehabilitation

After determining the test's content validity based on expert opinions and the calculation of the CVR, familiarity of items was assessed, and the final items were selected (Table 1).

The internal consistency of the subtests, measured using Cronbach's α coefficient, is presented in Table 2, indicating that each subtest of the P-RFLT had adequate internal consistency (α>0.7).

To evaluate test re-test reliability, the ICC at a 95% confidence interval was calculated using a two-way mixed model with absolute agreement (Table 3). The ICC values for the four subtests ranged from 0.71 to 0.81, reflecting adequate test re-test reliability.

Construct validity, assessed using the EFA, demonstrated that the data for each subtest independently supported the intended structure. The Kaiser–Meyer–Olkin (KMO) values for each subtest were 0.739, 0.778, 0.740, and 0.804, respectively. Bartlett's test of sphericity yielded a significance level for all components (P<0.05). Given that the factor loadings for each item within its respective domain were above or close to 0.4, the items for each domain were adequate. Furthermore, to establish normative reference ranges for the healthy sample, the minimum, maximum, mean, and standard deviation were calculated, as shown in Table 4. Based on the obtained mean and median values, the normative reference scores were as follows: 26 for the 32-item idioms subtest, 15 for the 18-item metaphors subtest, 9 for the 10-item proverbs subtest, 9 for the 10-item similes subtest, 4 for the 5-item proverb sentence completion subtest, and 3 for the 5-item simile sentence completion subtest.

Discussion

In the present study, we developed and validated the P-RFLT, comprising four components—metaphors, idioms, proverbs, and similes—for Persian-speaking adults. The findings demonstrated that the developed test had adequate content validity, face validity, and construct validity, with satisfactory internal consistency and test-retest reliability.

The analyses indicated that the test items coherently covered various types of figurative language and that the four-factor structure obtained after EFA was consistent with cognitive and pragmatic theoretical frameworks, although differed at the linguistic level. The types of figurative language depend on shared semantic processing networks at the cognitive level. The congruence of the

Table 2. Cronbach's alpha of the P-RFLT subtests

Subtest	Cronbach's α
Idioms	0.80
Metaphors	0.78
Proverbs	0.75
Similes	0.78

Archives of Rehabilitation

Table 3. The ICC values for test-retest reliability of the P-RFLT subtests

Variables		Mean±SD	ICC (95% CI)
Overall test	Baseline	66.1±5.7	0.68 (0.60, 0.75)
	Two weeks later	65.9±4.8	
Idioms	Baseline	26.94±2.38	0.71 (0.67, 0.83)
	Two weeks later	26.23±2.69	
Metaphors	Baseline	14.63±1.96	0.79 (0.69, 0.88)
	Two weeks later	15±1.59	
Proverbs	Baseline	9.16±0.93	0.77 (0.71, 0.85)
	Two weeks later	9±1.26	
Similes	Baseline	10.94±0.87	0.81 (0.65, 0.93)
	Two weeks later	11.22±0.98	

Archives of Rehabilitation

factor structure with the theoretical framework indicated that the empirical data support the conceptual model of figurative language within cognitive linguistics and pragmatics [4, 5].

A comparison of the present test with the FLIT [25] revealed that both instruments directly assess figurative language types, including metaphor, simile, and proverb, and are aligned in terms of design. Both tests have acceptable face and content validity, confirmed by experts, and comparable psychometric properties, including internal consistency. However, the FLIT primarily focuses on the global assessment of the ability to process metaphorical and indirect items, with results presented as a total score or on a limited set of subscales, while in the P-RFLT, figurative language was delineated as a multi-component construct, with each component subjected to independent psychometric analysis.

The findings of the present study are consistent with the results reported by Arcara and Bambini [26] for the APACS regarding reliability. Moreover, from a structural standpoint, the factor analysis results are consistent with those of the APACS, in which figurative language types were similarly reported as measurable components. This suggests that the cognitive organization of figurative language in both instruments is amenable to structural modeling. Nevertheless, the APACS study employed more advanced psychometric models as part of instrument standardization in English. In contrast, we used EFA to yield a four-component structure in Persian. In comparison with the study by Ghavami et al. on the Persian D-KEFS test battery [28], both studies empha-

Table 4. Score ranges for the P-RFLT subtests

Subtest	Minimum	Maximum	Median	Mean±SD
Idioms	16	32	26.23	26.09±2.7
Metaphors	8	18	15	14.92±2.1
Proverbs	4	10	8.95	8.9±1.01
Similes	3	10	8.66	8.5±1.4
Total	39	76	66.24	66.1±5.7

Archives of Rehabilitation

sized the importance of figurative language and the role of cultural context in its comprehension. However, their study focused on a specific dimension of figurative language and pragmatic analyses, whereas we examined figurative language as a multi-component construct within a structured psychometric instrument.

The variation in psychometric tests, particularly reliability tests, across components can be explained in terms of semantic processing and cognitive load. Different types of figurative language vary in semantic transparency and the degree of cognitive inference required. Specifically, similes, due to the presence of explicit comparative markers, have greater semantic transparency and are associated with lower executive demand and more direct processing; conversely, metaphors—particularly when unfamiliar—require inferential processing and more complex conceptual mappings. Furthermore, the degree of cultural familiarity with items plays a significant role in participant performance, such that familiar items (especially common idioms and proverbs) are processed more automatically and with a lower cognitive load. In contrast, unfamiliar items impose a greater executive demand on the cognitive system [7, 20]. From a neurological perspective, these differences may also be associated with the degree of engagement of classical language networks in the left hemisphere, abstract semantic processing networks in the right hemisphere, and executive networks in the frontal–parietal lobe. Accordingly, variations in familiarity, semantic transparency, and the type of linguistic structure can lead to differences in item difficulty and, consequently, changes in reliability indices and response patterns, which partially account for the discrepancies observed across studies [21].

In the present study, the primary focus was on the test's psychometric properties rather than its diagnostic function. Therefore, rather than establishing a cutoff score, the mean score of healthy individuals was used as a normative reference. This approach is consistent with the initial stages of instrument development, as the primary objective at this stage was to examine construct validity and reliability, and to provide a framework for the relative interpretation of performance, rather than definitive diagnosis. Using the mean scores of healthy individuals enables comparison of individual performance against normative patterns and can inform clinicians in language rehabilitation, without yielding a definitive diagnosis.

The present study, similar to comparable studies, had some limitations. Its administration among samples from Tehran city may limit the generalizability of the results to other cities with different cultures and geographic regions. Furthermore, the participants' age range (25-65 years) limits the applicability of the findings to this age group. Higher percentage of female participants was another limitation, due to women's greater willingness to participate in psychological and linguistic research, as well as their greater accessibility and cooperativeness during the sampling process. Although research evidence does not indicate a significant difference between men and women in the comprehension of figurative language, the difference in male to female ratio may exert a marginal effect on certain responses. Therefore, the generalizability of the results should be done with caution. Additionally, in evaluating the instrument's face and content validity, doctoral students were employed as panelists, and clinicians with clinical experience did not participate in this process, which may be a methodological limitation of the study. Another limitation was the absence of a clinical sample for the determination of a diagnostic cutoff score. Consequently, score interpretation was based solely on the performance of healthy individuals, which, although appropriate for the initial stages of instrument development, cannot substitute for definitive diagnostic criteria derived from comparisons with clinical groups.

It is recommended that future studies calculate additional validity and reliability indices, including criterion validity and discriminant validity, for the test. Furthermore, administering the P-RFLT to clinical cases, such as patients with aphasia or neurocognitive disorders, may facilitate the assessment of its sensitivity and clinical utility, as well as the establishment of cutoff scores. Additionally, evaluating the test across diverse age and gender groups, and in other geographic regions, can enhance the instrument's generalizability and external validity.

Conclusion

The P-RFLT is a valid and reliable tool with a coherent and measurable four-factor structure that can be used for assessing figurative language comprehension in Persian-speaking adults in both educational and clinical settings.

Ethical Considerations

Compliance with ethical guidelines

The research has ethical approval from the Ethics Committee of the Institute for Cognitive Science Studies, Shahid Beheshti University, Tehran, Iran (Code: IR.UT.IRICSS.REC.1402.033). All procedures were conducted in accordance with ethical principles in human studies. Prior to data collection, participants were provided with complete and transparent information

regarding the study objectives, procedures, and data usage. Participants were assured that their information would be kept strictly confidential and that participation was entirely voluntary, and they could withdraw at any time without consequences. Furthermore, participants were given access to the research findings so they could benefit from the study results.

Funding

This study was extracted from PhD Dissertation of Zahra Amini's, approved by the Institute for Cognitive Science Studies at Shahid Beheshti University. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Authors' contributions

Conceptualization, methodology, analysis, investigation, review & editing: All authors; Validation, writing original draft: Leila Ghasisin and Zahra Amini; Resources: Leila Ghasisin, Zahra Amini, and Mehdi Pourmohammad; Supervision and project administration: Leila Ghasisin.

Conflict of interest

The authors declared no conflict of interest.

AI tools disclosure statement

In preparing this manuscript, the artificial intelligence tool ChatGPT was used solely for editorial review, to improve textual fluency, enhance coherence and cohesion between paragraphs, and suggest linguistic revisions. The authors carefully reviewed all suggestions provided by the AI tool and edited them where necessary, and assume full responsibility for the final content of the manuscript.

Acknowledgments

The authors would like to thank all participants for their participation in this study.

مقاله پژوهشی

ساخت و بررسی ویژگی‌های روان‌سنجی آزمون ارزیابی درک زبان مجاز فارسی (P-RFLT) در بزرگسالان سالم

زهرا امینی[۱]، مهدی پورمحمد[۲]، علیرضا مرادی[۳]، *لیلا قسیسین[۴]

۱. پژوهشکده علوم شناختی، دانشگاه شهید بهشتی، تهران، ایران.
۲. گروه روان‌شناسی، دانشگاه مک‌اوان، ادمونتون، آلبرتا، کانادا.
۳. گروه روانشناسی تربیتی، دانشگاه آلبرتا، ادمونتون، آلبرتا، کانادا.
۴. گروه گفتاردرمانی، مرکز تحقیقات اسکلتی عضلانی، دانشکده علوم توانبخشی، دانشگاه علوم پزشکی اصفهان ، اصفهان ، ایران.

Use your device to scan and read the article online

Citation Amini Z, Purmohammad M, Mordi A, Ghasisin L. Development and Psychometric Evaluation of the Persian Receptive Figurative Language Test (P-RFLT) in Healthy Adults. Archives of Rehabilitation. 2026; 27(2):198-217. https://doi.org/10.32598/RJ.27.2.141.6

https://doi.org/10.32598/RJ.27.2.141.6

چکیده

هدف زبان مجاز یکی از مؤلفه‌های اساسی توانایی‌های زبانی–شناختی است که نقش مهمی در درک معانی غیرمستقیم، تفکر انتزاعی و تعاملات ارتباطی روزمره ایفا می‌کند و به توانایی فرد در تفسیر معانی ضمنی بستگی دارد و به بافت زبانی وابسته است. درک زبان مجاز مستلزم مهارت‌های پیشرفته زبانی‌شناختی است. در زبان فارسی ابزاری که بتواند به‌صورت جامع مؤلفه‌های زبان مجاز را ارزیابی کند، وجود ندارد؛ لذا هدف از این مطالعه ساخت آزمون درک زبان مجاز فارسی (P-RFLT) و بررسی ویژگی‌های روان‌سنجی آن در بزرگسالان سالم فارسی‌زبان بود.

روش بررسی این مطالعه یک پژوهش روش‌شناختی بود که به‌صورت مقطعی در فاصله زمانی ۱۴۰۲–۱۴۰۳ انجام شد. گویه‌های آزمون بر اساس چارچوب‌های نظری زبان مجاز در زبان‌شناسی شناختی و کاربردشناسی طراحی و چهار مؤلفه استعاره، اصطلاح، ضرب‌المثل و تشبیه در نظر گرفته شد. روایی صوری و محتوایی آزمون توسط گروهی از متخصصان زبان‌شناسی بررسی شد. پس از تعیین گویه‌ها و اجرای پایلوت، نسخه نهایی آزمون بر روی ۳۳۷ فرد بزرگسال سالم فارسی‌زبان در بازه سنی ۲۵ تا ۶۵ سال اجرا گردید.

یافته‌ها نتایج پژوهش نشان داد گویه‌های چهار مؤلفه آزمون درک زبان مجاز فارسی از روایی صوری و محتوایی مناسبی برخوردار بودند. روایی سازه آزمون از طریق تحلیل عاملی اکتشافی (EFA) تأیید شد و مقادیر KMO و بارهای عاملی نشان‌دهنده اندازه‌گیری مستقل هر مؤلفه و توانایی تمایزدهی آزمودنی‌ها بود. پایایی آزمون نیز از دو جنبه بررسی شد؛ همسانی درونی با ضریب آلفای کرونباخ بین ۰/۷۵ تا ۰/۸۰ و پایایی بازآزمون با شاخص ضریب همبستگی درون‌طبقه‌ای (ICC) بین ۰/۷۱ تا ۰/۸۱ نشان داد ابزار از انسجام و ثبات قابل‌قبول برخوردار است.

نتیجه‌گیری آزمون درک زبان مجاز فارسی از روایی و پایایی مناسبی برخوردار است و می‌تواند به‌عنوان ابزاری معتبر و کاربردی در ارزیابی‌های بالینی، توانبخشی شناختی و پژوهش‌های زبان و شناخت مورد استفاده قرار گیرد.

کلیدواژه‌ها زبان مجاز، روان‌سنجی، بزرگسالان، فارسی

تاریخ دریافت: ۱۹ آذر ۱۴۰۴
تاریخ اصلاح: ۰۴ خرداد ۱۴۰۵
تاریخ پذیرش: ۲۰ خرداد ۱۴۰۵

* **نویسنده مسئول:**
دکتر لیلا قسیسین
نشانی: دانشگاه علوم پزشکی اصفهان، دانشکده علوم توانبخشی، مرکز تحقیقات اسکلتی عضلانی، گروه گفتاردرمانی.
تلفن: ۳۲۵۳۷۵۳ (۹۱۳) ۹۸+
رایانامه: ghasisin@rehab.mui.ac.ir

Copyright © 2026 The Author(s);
This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-By-NC 4.0: https://creativecommons.org/licenses/by-nc/4.0/legalcode.en), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

مقدمه

زبان تمثیلی یا مجاز (Figurative Language) ابزاری شناختی و گفتمانی است که فراتر از معنای تحت‌اللفظی عمل می‌کند و به توانایی فرد در تفسیر معانی ضمنی و به بافت زبانی وابسته است. درک زبان مجاز مستلزم مهارت‌های پیشرفته زبانی، شناختی است [۱]. براساس نظریه مفاهیم واژگانی و الگوهای شناختی[1]، پردازش زبان مجاز حاصل تعامل هم‌زمان دانش زبانی و ساختارهای شناختی است؛ به‌گونه‌ای که فرد با بهره‌گیری از واژگان، ساختارهای نحوی و تجربیات پیشین، معنای غیرمستقیم جملات را تفسیر می‌کند. در این چارچوب، استعاره، اصطلاح، ضرب‌المثل و تشبیه جلوه‌هایی متفاوت از یک سازه شناختی زبانی مشترک محسوب می‌شوند [۱، ۲].

گیورا بیان می‌کند که درک زبان مجازی بیش از تمایز میان معنای تحت‌اللفظی و غیرتحت‌اللفظی، به برجستگی معنا وابسته است؛ برجستگی‌ای که تحت تأثیر تجربیات فرد شکل می‌گیرد [۳]. همچنین لیکاف و جانسون، زبان مجاز را بخشی جدایی‌ناپذیر از زبان و تفکر روزمره می‌دانند و معتقدند انسان‌ها برای بیان مفاهیم انتزاعی از آن استفاده می‌کنند [۴].

استعاره از مهم‌ترین مباحث معنی‌شناسی شناختی است و به‌صورت درک یک حوزه مفهومی برحسب حوزه دیگر تعریف می‌شود؛ برای‌مثال، در استعاره «عشق، سفر است»، مفهوم انتزاعی عشق از طریق ویژگی‌های ملموس سفر درک می‌شود [۵]. استعاره‌ها نگاشت‌هایی نظام‌مند میان حوزه مبدأ و حوزه هدف هستند و افزون بر سطح زبان، در سازمان‌دهی ساختارهای مفهومی ذهن نیز نقش دارند. بااین‌حال، استعاره‌ها ازنظر میزان آشنایی و نو بودن متفاوت‌اند [۶]. استعاره‌های متعارف، مانند «وقت طلاست»، بسامد بالا و درک سریع‌تری دارند؛ درحالی‌که استعاره‌های نو و نامتعارف، مانند «غیرت گدازه است»، به پردازش ذهنی بیشتری نیاز دارند [۷، ۸]. همچنین، اگرچه بسیاری از استعاره‌های مفهومی در فرهنگ‌های مختلف مشترک‌اند، زبان و فرهنگ در شکل‌گیری و تفسیر آن‌ها نقش مهمی دارند [۹، ۱۰].

اصطلاحات[2] نیز از جلوه‌های مهم زبان مجاز هستند [۱۱]. این عبارات، مانند «آبغوره نگیر»، معنایی فراتر از مجموع معانی واژه‌های تشکیل‌دهنده خود دارند [۱۲]. پژوهش‌ها نشان داده‌اند میزان آشنایی فرد با اصطلاحات در سرعت و دقت درک آن‌ها نقش اساسی دارد؛ به‌طوری‌که اصطلاحات آشنا سریع‌تر و آسان‌تر فهمیده می‌شوند [۱۳].

ضرب‌المثل‌ها شکل دیگری از زبان مجاز به شمار می‌آیند که نسبت به اصطلاحات، ساختار نحوی ثابت‌تری دارند و به‌صورت عبارات کاملاً ثابت به کار می‌روند، درحالی‌که اصطلاحات می‌توانند ازنظر شخص و زمان صرف شوند [۱۴]. ضرب‌المثل‌ها ساختارهای استعاری و دانش فرهنگی و جمعی را بازتاب می‌دهند و نوعی دانش فشرده‌شده[3] محسوب می‌شوند که مفاهیم انتزاعی و تجربی را در قالب استعاری منتقل می‌کنند [۱۵، ۱۶]. درک ضرب‌المثل‌های آشنا معمولاً به پردازش شناختی کمتری نیاز دارد، درحالی‌که ضرب‌المثل‌های ناآشنا مستلزم استنتاج بیشتری هستند [۱۷]. افزون‌براین، آزمون‌های مبتنی بر ضرب‌المثل‌ها می‌توانند شاخص مناسبی برای ارزیابی عملکردهای شناختی و زبانی باشند [۱۸].

درنهایت، تشبیه‌ها[4] شکلی از بیان مجازی‌اند که ساختاری صریح و مقایسه‌ای دارند. در زبان فارسی، این عبارات زیرگروهی از ضرب‌المثل‌ها محسوب می‌شوند که ساختاری تثبیت‌شده و فرهنگی دارند و برخلاف بسیاری از ضرب‌المثل‌ها، غالباً با استفاده از ادات تشبیه (مانند «مثل»، «انگار»، «همچو»، «چون») مقایسه‌ای صریح میان دو پدیده ایجاد می‌کنند. مانند «معلم‌ها مثل فانوس در شب تاریک‌اند» یک تشبیه محسوب می‌شود [۱۹]. با وجود شفافیت ساختاری بیشتر تشبیه‌ها، درک آن‌ها همچنان به فرایندهای استنتاجی وابسته است، به‌ویژه زمانی که شباهت میان دو مفهوم آشکار نباشد [۲۰].

مطالعات علوم اعصاب شناختی نشان داده‌اند نیم‌کره‌های مغز نقش متفاوتی در پردازش زبان مجاز دارند. نیم‌کره چپ بیشتر در پردازش استعاره‌های آشنا فعال است، درحالی‌که نیم‌کره راست در پردازش استعاره‌های جدید و پیچیده نقش برجسته‌تری دارد [۲۱، ۲۲]. این تفاوت بیشتر به عواملی مانند تازگی، دشواری و نیاز به یکپارچگی اطلاعات مربوط است، نه صرفاً ماهیت استعاری زبان [۲۳]. بررسی این فرایندها ازنظر بالینی نیز اهمیت دارد و می‌تواند به تمایز افراد سالم از مبتلایان به اختلالات شناختی کمک کند [۲۴].

باتوجه‌به نقش مهم زبان مجاز در تفکر انتزاعی و ارتباطات انسانی، ارزیابی آن در پژوهش‌ها و محیط‌های بالینی ضروری است. در این راستا، آزمون تفسیر زبان مجاز[5] برای سنجش درک استعاره، تشبیه، ضرب‌المثل و شخصیت‌بخشی طراحی و در ایالات متحده آمریکا هنجاریابی شده است [۲۵]. همچنین آزمون ارزیابی توانایی‌های کاربردشناختی و زیرساخت‌های شناختی[6] در سال ۲۰۱۶ برای ارزیابی مهارت‌های کاربردشناختی و زبان مجاز، از جمله استعاره و کنایه، در بزرگسالان ایتالیایی تدوین شد [۲۶].

1. The Lexical Concepts and Cognitive Models Theory (LCCM Theory)
2. Idioms
3. Compressed Knowledge
4. Similes
5. Figurative Language Interpretation Test (FLIT)
6. Assessment of Pragmatic Abilities and Cognitive Substrates (APACS)

در ایران نیز ابزارهایی برای سنجش برخی جنبه‌های زبان مجاز طراحی شده‌اند. نجاتی و رامش در سال ۱۳۹۴ آزمون ضرب‌المثل‌ها را برای دانش‌آموزان تهرانی تدوین کردند. این آزمون برای سنجش کارکردهای اجرایی و زبان در این دامنه سنی کاربرد دارد [۲۷]. همچنین مطالعه قوامی و همکاران در سال ۱۳۹۴، نسخه فارسی مجموعه آزمون متداول کارکردهای اجرایی دنیس کاپلان را در دامنه سنی ۱۶ تا ۴۰ سال تهیه نمودند [۲۸]. رحیمی‌فر و همکاران نیز در سال ۱۳۹۹ «آزمون مهارت‌های زبانی سطح بالا» را برای بیماران نورولوژیک و بزرگسالان طراحی کردند که این آزمون شامل خرده‌آزمون شامل تکرار جملات طولانی، جمله‌سازی، استنتاج، درک جملات پیچیده دستوری، درک جملات مبهم، تعریف کلمه و درنهایت ضرب‌المثل‌ها است [۲۹].

باوجوداین ابزارها، همچنان کمبود آزمون‌های جامع و استاندارد برای ارزیابی هم‌زمان استعاره، اصطلاح، ضرب‌المثل و تشبیه در زبان فارسی احساس می‌شود. این کمبود، ارزیابی دقیق توانایی‌های زبان مجاز را، به‌ویژه در بیماران مبتلا به اسکیزوفرنی، آسیب مغزی تروماتیک و بیماری‌های عصبی مانند زوال عقل فرونتوتمپورال و اسکلروز جانبی آمیوتروفیک[۷]، محدود کرده است [۳۰، ۳۱].

بنابراین، هدف پژوهش حاضر طراحی و بررسی ویژگی‌های روان‌سنجی «آزمون درک زبان مجاز فارسی[۸]» در بزرگسالان سالم فارسی‌زبان ۲۵ تا ۶۵ سال بود. این محدوده سنی به‌دلیل تأثیر تغییرات رشدی پیش از ۲۵ سالگی و فرایند پیری پس از ۶۵ سالگی بر پردازش زبان مجاز انتخاب شد [۳۲]. پژوهش حاضر با فراهم کردن ابزاری جامع برای سنجش زبان مجاز، می‌تواند در ارزیابی‌های بالینی و پژوهش‌های آینده کاربرد گسترده‌ای داشته باشد.

روش مطالعه

مطالعه حاضر یک پژوهش روش‌شناختی است که به‌صورت مقطعی با هدف توسعه و ارزیابی آزمون درک زبان مجاز فارسی طراحی شد. در این مطالعه، ابتدا گویه‌های زبان مجاز در چهار حوزه استعاره، ضرب‌المثل، اصطلاحات و تشبیه استخراج شد و سپس ویژگی‌های روان‌سنجی آن تعیین شد. مطالعه در سه مرحله صورت گرفت: ۱. طراحی گویه‌های آزمون، ۲. تعیین روایی محتوایی و صوری آن‌ها، ۳. اجرای نهایی آزمون.

جامعه آماری پژوهش

جامعه آماری این پژوهش، ۴۱۸ نفر فرد بزرگسال ۲۵ تا ۶۵ ساله بودند که ۴۱ نفر در مرحله تعیین سطح آشنایی گویه‌ها، ۴۰ نفر در مطالعه آزمایشی و ۳۳۷ نفر در مرحله نهایی مطالعه شرکت داشتند. حجم نمونه براساس جدول مورگان مشخص شد [۳۳]. نمونه‌گیری به شیوه در دسترس در استان تهران در پارک‌ها، محل کار، پاساژها، شرکت‌ها به‌مدت ۶ ماه در سال‌های ۱۴۰۲ و ۱۴۰۳ انجام گرفت.

معیارهای ورود به مطالعه عبارت بودند از قرار گرفتن در محدوده سنی ۲۵ تا ۶۵ سال، تک‌زبانه فارسی، داشتن حداقل تحصیلات ابتدایی به‌منظور پاسخگویی به آزمون، و عدم وجود اختلالات بینایی یا شنوایی غیرقابل‌اصلاح با وسایل کمکی که تمام موارد فوق براساس پرسش‌نامه اطلاعات فردی جمع‌آوری می‌شد. همچنین داشتن سلامت شناختی شرکت‌کنندگان که با استفاده از نسخه فارسی آزمون رتبه‌بندی بالینی زوال عقل (CDR-P)[۹] تأیید گردید [۳۴].

روش اجرا

مرحله اول: طراحی گویه‌های آزمون

نسخه فارسی آزمون درک زبان مجاز براساس آزمون چشلیسکا[۱۰] و همکاران توسعه یافت. این آزمون چهار مؤلفه اصلی زبان مجازی شامل استعاره، اصطلاح، ضرب‌المثل و تشبیه را دربر می‌گیرد [۳۵]. در مرحله نخست، عبارات مرتبط با این مؤلفه‌ها از منابع و متون مختلف به‌صورت گسترده و با روش قیاسی استخراج شدند. پس از استخراج، این عبارات در قالب جملات کوتاه و ساده به‌عنوان گویه‌های آزمون صورت‌بندی شدند.

به‌منظور کنترل متغیرهای زبانی مداخله‌گر، ویژگی‌های ساختاری گویه‌ها ازجمله تعداد کلمات، طول جمله و الگوی دستوری تا حد امکان یکسان نگهداشته شد؛ به‌گونه‌ای که تمامی گویه‌ها در قالب جملات ساده و کوتاه (۴ تا ۷ کلمه‌ای) تنظیم شدند. ازآنجایی‌که در مرحله تدوین آزمون بهتر است تعداد گویه‌ها ۲ تا ۳ برابر تعداد موردنیاز نهایی باشد [۳۶]؛ در مرحله اولیه ۸۶ استعاره، ۱۱۲ اصطلاح، ۴۸ ضرب‌المثل و ۴۸ تشبیه انتخاب شدند. دو مقوله استعاره‌ها و اصطلاحات بخش گسترده و بنیادی نظام مفهومی و واژگانی زبان را تشکیل می‌دهند، بنابراین به تعداد بیشتری انتخاب شدند [۴].

مرحله دوم: تعیین روایی صوری و محتوایی و میزان آشنایی

در این مرحله گویه‌ها در اختیار ۱۰ دانشجوی دکتری زبان‌شناسی شناختی قرار گرفت که ۵ نفر مدرس دانشگاه، ۳ نفر دارای سابقه کار در پژوهشکده‌های مختلف و ۲ نفر سابقه درمان در حوزه گفتار و زبان را داشتند تا گویه‌ها ازنظر تناسب با سازه، شفافیت، میزان نمایندگی و تناسب با هدف آزمون ارزیابی شوند. سپس شاخص نسبت روایی محتوایی[۱۱] (CVR) براساس قانون لاوشه[۱۲] محاسبه شد و گویه‌هایی که مقدار CVR آن‌ها کمتر از ۰/۶۲ بود، حذف شدند.

7. Amyotrophic Lateral Sclerosis (ALS)
8. Persian Figurative Receptive Language Test (P-RFLT)
9. Clinical Dementia Rating (CDR-P)
10. Cieślicka
11. Content Validity Ratio (CVR)
12. Lawshe

همچنین روایی صوری گویه‌ها ازنظر وضوح، قابل‌فهم بودن، سادگی نگارش، خوانایی، چیدمان و نظم کلی ارزیابی شد و براساس بازخوردهای دریافتی، اصلاحاتی ازجمله رفع ابهام، ساده‌سازی جملات و بازآرایی ساختار برخی گویه‌ها انجام گردید. در پایان این مرحله، ۲۸ استعاره، ۲۷ ضرب‌المثل، ۳۱ تشبیه و ۵۹ اصطلاح در قالب گویه‌های آزمون باقی ماند.

در ادامه، به‌منظور جلوگیری از بروز اثر سقف و کف، سطح دشواری گویه‌ها براساس میزان آشنایی[13] کنترل شد [۳۷]. به‌منظور سنجش این شاخص، نسخه اولیه آزمون در اختیار ۴۱ فرد بزرگسال واجد شرایط قرار گرفت و از آنان خواسته شد میزان آشنایی خود را با هر گویه براساس مقیاس لیکرت ۵ درجه‌ای (از «خیلی کم» تا «خیلی زیاد») گزارش کنند. گویه‌هایی با میانگین بالاتر از ۳ به‌عنوان آشنا و کمتر از ۳ به‌عنوان ناآشنا طبقه‌بندی شدند [۳۵].

درنهایت، براساس نتایج تحلیل آشنایی و نظر تیم پژوهشی، نسخه نهایی آزمون شامل ۱۸ استعاره، ۳۲ اصطلاح، ۱۵ ضرب‌المثل و ۱۵ تشبیه تدوین شد که به‌صورت پیوسته از سطح آسان به دشوار سازمان‌دهی شدند.

پس از این مراحل، برای سنجش پردازش شناختی زبان مجازی، گویه‌های آزمون در دو قالب طراحی شدند:

پرسش‌های چندگزینه‌ای (سه‌گزینه‌ای)

در این قالب، هر گویه به‌صورت یک واحد شامل جمله محرک (حاوی عبارت مجازی) و سه گزینه پاسخ طراحی شد: ۱. پاسخ صحیح با معنای مجازی یا مورد انتظار، ۲. گزینه انحرافی مبتنی بر معنای تحت‌اللفظی، و ۳. گزینه کاملاً نامرتبط. شرکت‌کنندگان موظف بودند، گزینه صحیح را انتخاب کنند. نمره‌دهی به‌صورت (۰) و (۱) انجام شد. این نوع پرسش‌ها برای سنجش درک معنای استعاره‌ها، اصطلاحات و بخشی از ضرب‌المثل‌ها و تشبیه‌ها به‌کار رفتند. به‌عنوان‌نمونه، در گویه «عموی من از دار دنیا رفت»، پاسخ صحیح «عموی من مرد» بود.

پرسش‌های تکمیل جمله

بخشی از گویه‌های مربوط به ضرب‌المثل‌ها و تشبیه‌ها به‌صورت تکمیل جمله طراحی شد. در این قالب، جمله به‌صورت ناقص ارائه می‌شود و شرکت‌کنندگان می‌بایست آن را در حد یک یا دو واژه تکمیل کنند. این نوع پرسش‌ها به‌منظور سنجش پردازش عمیق‌تر معنایی، دسترسی به معانی ذخیره‌شده در حافظه بلندمدت و توانایی استنتاج رابطه معنایی به‌کار رفتند [۳۸]. نمره‌دهی در این بخش نیز به‌صورت (۰) و (۱) انجام شد و اشتباهات نگارشی در ارزیابی لحاظ نشد. به‌عنوان‌مثال، در ضرب‌المثل «پایان شب...»، پاسخ مورد انتظار «سیه سپید است» بود.

13. Familiarity

نمونه سؤالات طراحی‌شده در طی یک مطالعه آزمایشی بر روی ۴۰ شرکت‌کننده واجد شرایط اجرا شد. پیش از اجرا، از شرکت‌کنندگان درخواست شد درباره اندازه و نوع قلم نوشتاری و همچنین فاصله خطوط اظهار نظر کنند تا آزمون از نظر خوانایی بهینه شود. تمامی شرکت‌کنندگان نسخه آماده‌شده را تأیید کردند.

پس از اجرای آزمایشی، فراوانی و درصد انتخاب هر گزینه محاسبه و میزان جذابیت گزینه‌های انحرافی ارزیابی شد. همچنین، همبستگی انتخاب هر گزینه با نمره کل آزمون بررسی گردید تا قدرت تمییزدهندگی گزینه‌ها مشخص شود. گزینه‌هایی که فراوانی انتخاب بسیار پایین یا عملکرد تمییزدهنده مناسبی نداشتند و از گزینه‌های انحرافی قابل‌تفکیک نبودند، شناسایی و اصلاح شدند و نسخه نهایی آزمون تهیه شد.

مرحله سوم: اجرای نهایی آزمون

نسخه نهایی آزمون به‌منظور بررسی همسانی درونی، پایایی بازآزمون و روایی سازه بر روی ۳۳۷ شرکت‌کننده واجد معیارهای ورود اجرا شد. اجرای آزمون در محیطی آرام انجام گرفت. در ابتدا، هدف پژوهش به‌صورت شفاهی برای شرکت‌کنندگان توضیح داده شد و در صورت تمایل، رضایت‌نامه آگاهانه تکمیل و سپس پرسش‌نامه اطلاعات فردی تکمیل می‌شد. به‌منظور اطمینان از سلامت شناختی، آزمون رتبه‌بندی بالینی زوال عقل (CDR-P) اجرا گردید. آزمون اصلی به‌صورت مداد-کاغذی خوداجرا[14] در اختیار شرکت‌کنندگان قرار گرفت.

به‌منظور آشنایی شرکت‌کنندگان با نحوه پاسخ‌دهی، دو گویه نمونه (یکی چندگزینه‌ای و دیگری تکمیل جمله) در ابتدای هر بخش از آزمون قرار داده شد.

به‌منظور کاهش خستگی شرکت‌کنندگان، گویه‌ها به دو بخش مساوی تقسیم شدند و بین دو بخش، استراحت کوتاهی در نظر گرفته شد. پس از پایان اجرای آزمون، نمره‌دهی گویه‌ها توسط یک ارزیاب مستقل که در فرایند نمونه‌گیری دخالتی نداشت، انجام شد. ورود داده‌ها به نرم‌افزار SPSS نیز توسط فردی مستقل از فرایند نمونه‌گیری و نمره‌دهی صورت گرفت تا دقت داده‌ها افزایش یابد. به‌منظور بررسی پایایی آزمون‌بازآزمون، آزمون مجدداً با فاصله زمانی دو هفته بر روی ۶۷ نفر از شرکت‌کنندگان اجرا شد.

روش تحلیل داده‌ها

جهت تحلیل داده‌ها از نرم‌افزار SPSS با نسخه ۲۰ و شاخص‌های توصیفی و استنباطی استفاده شد. آمار توصیفی برای محاسبه شاخص‌هایی مانند میانگین، انحراف‌معیار، و میانه بود که برای توصیف ویژگی‌های عملکرد آزمودنی‌ها در خرده‌آزمون‌ها استفاده شد. روایی صوری به‌صورت کیفی براساس نظر متخصصان و

14. Paper-and-pencil, Self-administered

جدول ۱. فراوانی گویه‌های دارای نسبت روایی محتوایی (CVR) و میزان آشنایی با آن‌ها در آزمون درک زبان مجاز

مؤلفه‌ها	تعداد گویه‌های CVR ≤ ۰/۶۲	گویه‌های آشنا (میانگین ≤ ۳)	گویه‌های ناآشنا (میانگین ≥ ۳)	گویه‌های آشنا انتخاب‌شده	گویه‌های ناآشنا انتخاب‌شده	گویه‌های نهایی انتخاب‌شده
اصطلاحات	۵۹	۳۹	۲۰	۱۷	۱۵	۳۲
استعاره‌ها	۲۸	۱۲	۱۶	۹	۹	۱۸
ضرب‌المثل‌ها	۲۷	۱۷	۱۰	۱۰	۵	۱۵
تشبیه	۳۱	۲۲	۹	۱۰	۵	۱۵

توانبخشی

شرکت‌کنندگان ، روایی محتوایی براساس CVR، همسانی درونی از طریق محاسبه ضریب آلفای کرونباخ. روایی بازآزمون از طریق محاسبه ضریب همبستگی درون‌طبقه‌ای (ICC) و روایی سازه به کمک تحلیل عاملی اکتشافی انجام شد.

یافته‌ها

آزمون در مراحل مختلف بر روی ۴۱۸ فرد بزرگ‌سال سالم در دامنه سنی ۲۵ تا ۶۵ سال در سال ۱۴۰۲ اجرا شد. از میان شرکت‌کنندگان، ۷۰ درصد (۲۹۳ نفر) زن و ۳۰ درصد (۱۲۵نفر) مرد بودند

از نظر توزیع سنی، ۱۸ درصد در گروه سنی ۲۵ تا ۳۴ سال، ۵۱/۶ درصد در گروه سنی ۳۵ تا ۴۴ سال، ۲۳ درصد در گروه سنی ۴۵ تا ۵۴ سال و ۷/۴ درصد در گروه سنی ۵۵ تا ۶۵ سال قرار داشتند. درخصوص سطح تحصیلات، ۱ درصد از شرکت‌کنندگان دارای تحصیلات زیر دیپلم، ۱۶/۶ درصد دیپلم، ۷/۴ درصد فوق‌دیپلم، ۴۷/۴ درصد کارشناسی، ۲۲/۶ درصد کارشناسی ارشد و ۵ درصد دارای مدرک دکتری بودند.

روایی محتوایی و میزان آشنایی با گویه‌ها

برای تعیین روایی محتوایی آزمون از نظر متخصصان و محاسبه CVR استفاده شد و میزان آشنایی گویه‌ها، با یک مقیاس لیکرت ۵ درجه‌ای مشخص شد که نتایج حاصل در جدول شماره ۱ نشان داده شده است.

پایایی آزمون زبان مجاز

جهت تعیین پایایی آزمون زبان مجاز از دو روش همسانی درونی و پایایی بازآزمون استفاده شد.

همسانی درونی زیرآزمون‌های زبان مجاز با استفاده از ضریب آلفای کرونباخ در جدول شماره ۲ نشان می‌دهد آزمون زبان مجاز در هر بخش دارای همسانی درونی مناسبی است (α>۰/۷).

پایایی بازآزمون، با گرفتن آزمون مجدد پس از ۲ هفته از ۶۷ فرد که در مرحله اول شرکت داشتند، انجام شد. برای بررسی پایایی بازآزمون از ضریب همبستگی درون طبقه‌ای (ICC) با فاصله اطمینان ۹۵ درصد (ICC) براساس مدل دوطرفه مختلط[15] و تعریف توافق مطلق[16] استفاده شد (جدول شماره ۳). نتایج نشان داد مقادیر ICC برای چهار مؤلفه بین ۰/۷۱ تا ۰/۸۱ بود که بیانگر پایایی بازآزمون مطلوب است.

روایی سازه آزمون

تعیین روایی سازه با استفاده از تحلیل عاملی اکتشافی نشان داد در هر مؤلفه، به‌طور جداگانه داده‌ها ساختار مورد نظر را پشتیبانی می‌نمایند، به‌طوری‌که مقدار KMO برای هر مؤلفه به‌ترتیب، ۰/۷۳۹، ۰/۷۷۸، ۰/۷۴ و ۰/۸۰۴ به دست آمد. سطح معناداری آزمون بارتلت در همه موارد و برای هر مؤلفه کمتر از ۰/۰۵ به دست آمد (P<۰/۰۵). ازآنجاکه مقادیر بار عاملی هر

15. Two-Way Mixed
16. Absolute Agreement

جدول ۲. همسانی درونی زیر آزمون‌ها آزمون‌های زبان مجاز

زیرآزمون	آلفای کرونباخ
اصطلاحات	۰/۸۰
استعاره‌ها	۰/۷۸
ضرب‌المثل‌ها	۰/۷۵
تشبیهات	۰/۷۸

توانبخشی

جدول۳. آمار توصیفی و پایایی پایایی آزمون- بازآزمون طی دو مرحله اندازه‌گیری

متغیر		میانگین±انحراف‌معیار	فاصله اطمینان (۹۵ ٪)
نمره کل آزمون	شروع	۶۶/۱±۵/۷	۰/۶۸ (۰/۷۵، ۰/۶)
	۲ هفته بعد	۶۵/۹±۴/۸	
اصطلاحات	شروع	۲۶/۹۴±۲/۳۸	۰/۷۱ (۰/۸۳، ۰/۶۷)
	۲ هفته بعد	۲۶/۲۳±۲/۶۹	
استعاره‌ها	شروع	۱۴/۶۳±۱/۹۶	۰/۷۹ (۰/۸۸، ۰/۶۹)
	۲ هفته بعد	۱۵±۱/۵۹	
ضرب‌المثل	شروع	۹/۱۶±۰/۹۳	۰/۷۷ (۰/۸۵، ۰/۷۱)
	۲ هفته بعد	۹±۱/۲۶	
تشبیهات	شروع	۱۰/۹۴±۰/۸۷	۰/۸۱ (۰/۹۳، ۰/۶۵)
	۲ هفته بعد	۱۱/۲۲±۰/۹۸	

توانبخشی

سؤال در هر مؤلفه، بالای ۰/۴ یا نزدیک ۰/۴ به دست آمد، نشان از مناسب بودن سؤالات مربوطه هر حوزه داشت.

همچنین برای دامنه مرجع عملکرد نمونه سالم، کمترین نمره، بیشترین نمره، میانگین و انحراف‌معیار محاسبه شد که در جدول شماره ۴ قابل‌مشاهده است.

باتوجه‌به مقادیر میانگین و میانه به‌دست‌آمده، می‌توان نتیجه گرفت دامنه مرجع عملکرد نمونه سالم برای آزمون ۳۲ گویه‌ای اصطلاحات ۲۶، برای آزمون ۱۸ گویه‌ای استعاره‌ها ۱۵، برای آزمون ۱۰ گویه‌ای ضرب‌المثل‌ها ۹، برای آزمون ۱۰ گویه‌ای تشبیهی ۹، برای آزمون ۵ گویه‌ای تکمیل جمله ضربالمثل، ۴ و برای آزمون ۵ گویه‌ای تکمیل جمله تشبیهی ۳ می‌باشد.

بحث

این مطالعه با هدف توسعه و اعتبارسنجی آزمون درک زبان مجاز فارسی چهار مؤلفه‌ای استعاره‌ها، اصطلاحات، ضرب‌المثل و تشبیهات در بزرگسالان فارسی‌زبان طراحی شد. یافته‌ها نشان داد آزمون طراحی‌شده از روایی محتوایی، صوری و سازه‌ای مطلوب و شاخص‌های پایایی آن از انسجام درونی و ثبات زمانی مناسبی برخوردار است. تحلیل‌ها نشان دادند گویه‌های آزمون به‌طور منسجم ابعاد مختلف زبان مجاز را پوشش می‌دهند و ساختار چهار مؤلفه‌ای به‌دست‌آمده در تحلیل عاملی اکتشافی با چارچوب نظری شناختی و کاربردشناختی همخوانی دارد. از این منظر، تفکیک آزمون به چهار مؤلفه در پژوهش حاضر با این پیش‌فرض نظری هم‌راستاست که هر نوع زبان مجاز، اگرچه در سطح زبانی متفاوت است، اما در سطح شناختی به شبکه‌های مشترک پردازش معنا وابسته است. یافته‌های تحلیل عاملی اکتشافی نشان داد گویه‌ها به‌طور منسجم در چهار عامل متمایز قرار می‌گیرند. همخوانی ساختار عاملی با چارچوب نظری نشان داد داده‌های تجربی از مدل مفهومی زبان مجاز در زبان‌شناسی شناختی و کاربردشناختی حمایت می‌کند [۴، ۵].

مقایسه آزمون حاضر با آزمون (FLIT) [۲۵] نشان داد هر دو ابزار به‌طور مستقیم توانایی پردازش زبان مجاز شامل استعاره، تشبیه و ضرب‌المثل را ارزیابی می‌کنند و ازنظر هدف طراحی هم‌راستا

جدول ۴. توزیع نمرات زیرآزمون‌ها (n=۳۳۷)

زیر آزمون	کمترین نمره	بیشترین نمره	میانه	میانگین±انحراف‌معیار
اصطلاحات	۱۶	۳۲	۲۶/۲۳	۲۶/۰۹±۲/۷
استعاره‌ها	۸	۱۸	۱۵	۱۴/۹۲±۲/۱
ضرب‌المثل	۴	۱۰	۸/۹۵	۸/۹±۱/۰۱
تشبیه	۳	۱۰	۸/۶۶	۸/۵±۱/۴
نمره کل آزمون زبان مجاز	۳۹	۷۶	۶۶/۲۴	۶۶/۱±۵/۷

توانبخشی

هستند. در هر دو آزمون، روایی صوری و محتوایی از طریق نظر متخصصان و بازبینی گویه‌ها تأیید شده است و شاخص‌های روان‌سنجی مشابه ازجمله آلفای کرونباخ گزارش شده است. بااین‌حال، تفاوت اصلی در آن است که آزمون (FLIT) عمدتاً بر سنجش کلی توانایی پردازش آیتم‌های استعاری و غیرمستقیم تمرکز دارد و نتایج آن در سطح نمره کل یا زیرمقیاس‌های محدود ارائه می‌شود، درحالی‌که در پژوهش حاضر، زبان مجاز به‌صورت یک ساختار چندمؤلفه‌ای تفکیک شده و هر مؤلفه به‌طور مستقل تحلیل روان‌سنجی شده است.

یافته‌های این پژوهش با نتایج آکرا و بامبینی [۲۶] ازنظر شاخص‌های پایایی همسو است. همچنین از نظر ساختاری، نتایج تحلیل عاملی با مطالعه APACS همخوانی دارد که در آن نیز توانایی‌های مرتبط با زبان مجاز در قالب مؤلفه‌های قابل‌تفکیک و منسجم قابل‌اندازه‌گیری گزارش شده است. این امر نشان می‌دهد سازمان‌دهی شناختی زبان مجاز در هر دو مطالعه قابلیت مدل‌سازی ساختاری دارد. بااین‌حال، در مطالعه APACS از مدل‌های پیشرفته‌تر روان‌سنجی و در چارچوب استانداردسازی ابزار در زبان انگلیسی استفاده شده است، درحالی‌که در پژوهش حاضر تحلیل عاملی اکتشافی با هدف استخراج ساختار چهارمؤلفه‌ای در بافت زبانی– فرهنگی فارسی انجام شده است.

در مقایسه با مطالعه قوامی و همکاران [۲۸] باید گفت هر دو پژوهش بر اهمیت زبان مجاز و نقش بافت فرهنگی در درک آن تأکید دارند. بااین‌حال، تفاوت اصلی با آزمون حاضر در رویکرد پژوهشی است؛ به‌طوری‌که مطالعه مذکور بر یک بعد خاص از زبان مجاز و تحلیل‌های کاربردشناختی تمرکز داشته، درحالی‌که در پژوهش حاضر زبان مجاز به‌صورت چندمؤلفه‌ای و در قالب یک ابزار روان‌سنجی ساختاریافته بررسی شده است.

تفاوت در شاخص‌های روان‌سنجی و به‌ویژه میزان پایایی بین مؤلفه های مختلف را می‌توان از منظر پردازش معنا و بار شناختی تبیین کرد. انواع مختلف زبان مجاز از نظر میزان شفافیت معنایی و نیاز به استنتاج شناختی با یکدیگر تفاوت دارند؛ به‌طوری‌که تشبیه‌ها به‌دلیل وجود نشانه‌های صریح مقایسه از شفافیت معنایی بیشتری برخوردار بوده و با بار اجرایی کمتر و پردازش مستقیم‌تر همراه هستند. درمقابل، استعاره‌ها به‌ویژه در صورت ناآشنا بودن، نیازمند پردازش استنباطی و نگاشت‌های مفهومی پیچیده‌تر هستند. همچنین میزان آشنایی فرهنگی با گویه‌ها نقش مهمی در عملکرد آزمودنی‌ها دارد، به‌گونه‌ای که آیتم‌های آشنا (به‌ویژه اصطلاحات و ضرب‌المثل‌های رایج) با پردازش خودکارتر و بار شناختی کمتر پردازش می‌شوند؛ درحالی‌که آیتم‌های ناآشنا بار اجرایی بیشتری بر سیستم شناختی تحمیل می‌کنند [۷، ۲۰].

از منظر عصبی نیز این تفاوت‌ها می‌توانند با میزان درگیری شبکه‌های زبانی کلاسیک در نیمکره چپ، شبکه‌های مرتبط با پردازش معنایی انتزاعی در نیمکره راست و شبکه اجرایی فرونتال–پریتال مرتبط باشند. بنابراین تفاوت در آشنایی، شفافیت معنایی و نوع ساختار زبانی می‌تواند به تفاوت در دشواری گویه‌ها و در نتیجه تغییر در شاخص‌های پایایی و الگوی پاسخ‌دهی منجر شود که این امر بخشی از تفاوت‌های مشاهده‌شده میان مطالعات را تبیین می‌کند [۲۱].

در این مطالعه، تمرکز اصلی بر ویژگی‌های روان‌سنجی آزمون بود و نه بر کارکرد تشخیصی آن. بنابراین به‌جای تعیین نقطه برش، از میانگین عملکرد افراد سالم به‌عنوان شاخص مرجع استفاده شد. این رویکرد با مراحل اولیه ابزارسازی همخوان است، زیرا هدف اصلی در این مرحله بررسی روایی سازه، پایایی و ارائه چارچوبی برای تفسیر نسبی عملکرد است، نه تشخیص قطعی. استفاده از میانگین عملکرد افراد سالم امکان مقایسه عملکرد فردی با الگوی طبیعی را فراهم می‌کند و می‌تواند به تفسیر بالینی در حوزه توانبخشی زبانی کمک کند، بدون آنکه به تشخیص قطعی منجر شود.

نتیجه‌گیری

درمجموع، یافته‌های این پژوهش نشان می‌دهد آزمون ارزیابی درک زبان مجاز در بزرگسالان فارسی‌زبان از ساختاری منسجم و قابل‌اندازه‌گیری برخوردار است و می‌توان آن را در قالب مؤلفه‌های متمایز اما مرتبط تبیین کرد. الگوی به‌دست‌آمده بیانگر آن است که پردازش انواع زبان مجاز تحت تأثیر میزان شفافیت معنایی، آشنایی فرهنگی و بار شناختی متفاوت قرار دارد. از منظر کاربردی، نتایج نشان می‌دهد آزمون طراحی‌شده می‌تواند ابزار مناسبی برای ارزیابی توانایی‌های مرتبط با درک زبان مجاز در بافت زبان فارسی باشد و در موقعیت‌های آموزشی و بالینی مورد استفاده قرار گیرد. همچنین، هم‌راستایی ساختار عاملی با چارچوب‌های نظری زبان‌شناسی شناختی از کفایت نظری ابزار حمایت می‌کند.

محدودیت‌ها

این مطالعه نیز مانند پژوهش‌های مشابه دارای محدودیت‌هایی است. اجرای آن در شهر تهران ممکن است تعمیم‌پذیری نتایج را به سایر فرهنگ‌ها و مناطق جغرافیایی محدود کند. همچنین، دامنه سنی شرکت‌کنندگان (۲۵ تا ۶۵ سال) موجب می‌شود نتایج صرفاً به این گروه سنی قابل‌تعمیم باشد. از نظر ترکیب نمونه، عدم توازن جنسیتی (غلبه نسبی زنان) نیز از محدودیت‌ها محسوب می‌شود که احتمالاً ناشی از تمایل بیشتر زنان به مشارکت در پژوهش‌های روان‌شناختی و زبان‌شناختی و همچنین همکاری و دسترسی بالاتر آنان در فرآیند نمونه‌گیری بوده است. اگرچه شواهد پژوهشی تفاوت معناداری بین زنان و مردان در درک زبان مجاز نشان نمی‌دهد، بااین‌حال این عدم توازن ممکن است

بر برخی الگوهای پاسخ‌دهی تأثیر جزئی داشته باشد. بنابراین تعمیم‌پذیری نتایج باید با احتیاط انجام شود. همچنین، در بررسی روایی صوری و محتوایی ابزار، از دانشجویان مقطع دکتری استفاده شد و متخصصان دارای تجربه بالینی در این فرایند مشارکت نداشتند که این امر می‌تواند به‌عنوان یکی از محدودیت‌های روش‌شناختی مطالعه در نظر گرفته شود. درنهایت، یکی دیگر از محدودیت‌های مطالعه، عدم استفاده از نمونه بالینی برای تعیین نقطه برش تشخیصی بود؛ ازاین‌رو تفسیر نمرات ابزار صرفاً براساس عملکرد افراد سالم انجام شد که اگرچه برای مراحل اولیه ابزارسازی مناسب است، اما نمی‌تواند جایگزین معیارهای تشخیصی قطعی مبتنی بر مقایسه با گروه‌های بالینی باشد.

پیشنهادها

پیشنهاد می‌شود در مطالعات آتی، سایر شاخص‌های روایی و پایایی، ازجمله روایی ملاکی و افتراقی نیز برای آزمون محاسبه گردد. همچنین، اجرای این آزمون بر روی گروه‌های بالینی، مانند بیماران مبتلا به زبان‌پریشی یا اختلالات عصبی-شناختی، می‌تواند به بررسی حساسیت و کاربرد بالینی آن و تعیین نقاط برش کمک کند. به‌علاوه، بررسی آزمون در جمعیت‌های مختلف سنی و جنسیتی و همچنین سایر مناطق نیز می‌تواند به تعمیم‌پذیری و اعتبار بیرونی ابزار بیفزاید.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش

طرح این پژوهش پیش از اجرا، در کمیته اخلاق در پژوهش دانشگاه شهید بهشتی تصویب شد و دارای کد تأییدیه اخلاقی از کمیته اخلاق مؤسسه آموزش عالی علوم شناختی با شناسه اخلاق (IR.UT.IRICSS.REC.1402.033) می‌باشد. تمامی مراحل تحقیق مطابق با اصول اخلاقی پذیرفته‌شده در مطالعات انسانی انجام گرفت. پیش از شروع گردآوری داده‌ها، اطلاعات کامل و شفافی درخصوص هدف، مراحل اجرا و شیوه استفاده از داده‌ها به شرکت‌کنندگان ارائه شد. به آن‌ها اطمینان داده شد که اطلاعاتشان به‌صورت محرمانه نگهداری خواهد شد و مشارکت در پژوهش کاملاً داوطلبانه است؛ به این معنا که هر زمان تمایل داشته باشند، می‌توانند بدون هیچ پیامدی از ادامه همکاری انصراف دهند و سپس از آن‌ها در دو نسخه رضایت اخلاقی گرفته شد که یک نسخه نزد پژوهشگران و دیگری نزد شرکت‌کننده بود. همچنین، امکان دسترسی به نتایج کلی پژوهش برای داوطلبان فراهم شده است تا از یافته‌های تحقیق بهره‌مند شوند.

حامی مالی

مطالعه حاضر بخشی از یافته حاصل از رساله دکتری زهرا امینی می‌باشد که در پژوهشکده علوم شناختی دانشگاه شهید بهشتی تصویب شده است. برای انجام این پژوهش کمک مالی دریافت نشده و تمامی هزینه‌های آن به‌صورت شخصی توسط پژوهشگران تأمین شده است.

مشارکت نویسندگان

مفهوم‌سازی، روش‌شناسی، تحلیل، تحقیق و بررسی، ویراستاری و نهایی‌سازی نوشته و تأمین مالی: همه نویسندگان؛ اعتبارسنجی و نگارش پیش‌نویس: لیلا قسیسین و زهرا امینی؛ منابع: لیلا قسیسین، زهرا امینی و مهدی پورمحمد؛ نظارت و مدیریت پروژه: لیلا قسیسین.

تعارض منافع

بنابر اظهار نویسندگان، این مقاله تعارض منافع ندارد.

بیانیه استفاده از ابزارهای هوش مصنوعی

در تهیه این مقاله از ابزار هوش مصنوعی(ChatGPT) صرفاً به‌منظور بازبینی نگارشی، بهبود روانی متن، ارتقای انسجام و پیوستگی بین پاراگراف‌ها، و پیشنهاد اصلاحات زبانی استفاده شده است. تمامی ایده‌های پژوهش، طراحی مطالعه، گردآوری و تحلیل داده‌ها، تفسیر یافته‌ها و نتیجه‌گیری‌های علمی توسط نویسندگان انجام شده است. نویسندگان تمامی پیشنهادهای ارائه‌شده توسط ابزار هوش مصنوعی را به‌دقت بررسی و در صورت لزوم ویرایش کرده‌اند و مسئولیت کامل محتوای نهایی مقاله را بر عهده دارند.

تشکر و قدردانی

نویسندگان مقاله از تمامی شرکت‌کنندگان در مراحل مختلف این پژوهش تشکر و قدردانی می‌کنند.

References

[1] Evans V. Figurative language understanding in LCCM Theory. Cognitive linguistics. 2010; 2110:601-62. [DOI:10.1515/cogl.2010.020]

[2] Colston H, Gibbs RW. Figurative Language Communicates Directly Because It Precisely Demonstrates What We Mean. Canadian Journal of Experimental Psychology. 2021; 75(2):228-33. [DOI:10.1037/cep0000254] [PMID]

[3] Giora R. Literal vs. figurative language: Different or equal? Journal of Pragmatics. 2002; 34(4):487-506. [DOI:10.1016/S0378-2166(01)00045-5]

[4] Lakoff G, Johnson Mark. Metaphors We Live By. Chicago: The University of Press; 2003. [DOI:10.7208/chicago/9780226470993.001.0001]

[5] Kövecses: Z. Metaphor: A Practical Introduction. Second ed. Oxford: Oxford: Oxford University Press; 2010. [Link]

[6] Beaty RE, Silvia PJ. Metaphorically speaking: Cognitive abilities and the production of figurative language. Memory & Cognition. 2013; 41(2):255-67. [DOI:10.3758/s13421-012-0258-5] [PMID]

[7] Mon SK, Nencheva M, Citron FMM, Lew-Williams C, Goldberg AE. Conventional metaphors elicit greater real-time engagement than literal paraphrases or concrete sentences. Journal of Memory and Language. 2021; 121:104285. [DOI:10.1016/j.jml.2021.104285]

[8] Yao Z, Huang X, Chai Y, Zhang J. Conventionality and context jointly modulate the effect of inhibitory control on L2 metaphor comprehension. Humanities and Social Sciences Communications. 2024; 11:1460. [DOI:10.1057/s41599-024-03977-4]

[9] Chen J, Lv J, Chen B. Crossing the cultural bridge: The role of inhibitory control during second language metaphor comprehension. Bilingualism: Language and Cognition. 2025; 28(5):1393-409. [DOI:10.1017/S1366728924001081]

[10] Kabra A, Liu E, Khanuja S, Aji A, Winata G, Aremu A, et al. Multi-lingual and Multi-cultural Figurative Language Understanding. California: Association for Computational Linguistics; 2023. [DOI:10.18653/v1/2023.findings-acl.525]

[11] Wiliński J. Metaphodioms: Connecting metaphor and idioms. Brno studies in English. 2022; 48(1):117-35. [DOI:10.5817/BSE2022-1-6]

[12] Evans V, Green M. Cognitive Linguistics: An Introduction. second ed. London: Routledge; 2018. [Link]

[13] Alasgarova R, Mahmudova I, Rzayev J. Cross-linguistic processing of idioms: The role of cultural familiarity and Construction Grammar in idiom comprehension. 2025; 8:2025. [DOI:10.2478/lf-2025-0009]

[14] Langlotz A. Idiomatic creativity: A cognitive-linguistic model of idiom-representation and idiom-variation in English. Amsterdam: John Benjamins Publishing Company; 2006. [Link]

[15] Kljajevic V. Older and Wiser: Interpretation of Proverbs in the Face of Age-Related Cortical Atrophy. Frontiers in Aging Neuroscienc. 2022; 14:919470. [DOI:10.3389/fnagi.2022.919470] [PMID]

[16] Woźniak J. [Rec.] Sadia Belkhir. Proverbs within cognitive linguistics. State of the art. Amsterdam / Philadelphia: John Benjamins Publishing Company. 2024. Pp. 351. Glottodidactica. 2025; 52:207-11. [DOI:10.14746/gl.2025.52.12]

[17] Lauro LJ, Tettamanti M, Cappa SF, Papagno C. Idiom Comprehension: A Prefrontal Task? Cerebral Cortex. 2008; 18(1):162-70. [DOI:10.1093/cercor/bhm042] [PMID]

[18] Amanzio M, Geminiani G, Leotta D, Cappa S. Metaphor comprehension in Alzheimer's disease: Novelty matters. Brain and Language. 2008; 107(1):1-10. [DOI:10.1016/j.bandl.2007.08.003] [PMID]

[19] Kosimov KA. Investigating the Similarities and Differences Between Metaphor and Simile: A Cognitive Linguistics Perspective. Journal of Ethics and Diversity in International Communication. 2023; 3(6):23-7. [Link]

[20] Kao CY. Similes, metaphors, and creativity. Educational Psychology. 2021; 42(2):163–84. [DOI:10.1080/01443410.2021.1929850]

[21] Schmidt GL, DeBuse CJ, Seger CA. Right hemisphere metaphor processing? Characterizing the lateralization of semantic processes. Brain and Language. 2007; 100(2):127-41. [DOI:10.1016/j.bandl.2005.03.002] [PMID]

[22] Cardillo ER, Watson CE, Schmidt GL, Kranjec A, Chatterjee A. From novel to familiar: Tuning the brain for metaphors. NeuroImage. 2012; 59(4):3212-21. [DOI:10.1016/j.neuroimage.2011.11.079] [PMID]

[23] Tang X, Qi S, Jia X, Wang B, Ren W. Comprehension of scientific metaphors: Complementary processes revealed by ERP. Journal of Neurolinguistics. 2017; 42:12-22. [DOI:10.1016/j.jneuroling.2016.11.003]

[24] Yang J, Huang L. Comprehension of metaphors in patients with mild cognitive impairment: Evidence from behavioral and ERP data. Acta Psychologica. 2023; 235:103894. [DOI:10.1016/j.actpsy.2023.103894] [PMID]

[25] Smith EH, Palmer BC. Smith/Palmer Figurative Language Interpretation Test. Palo Alto, CA: Consulting Psychologists Press; 1979. [Link]

[26] Arcara G, Bambini V. A Test for the Assessment of Pragmatic Abilities and Cognitive Substrates (APACS): Normative Data and Psychometric Properties. Frontiers in Psychology. 2016; 7:70. [Link]

[27] Nejati V, Ramesh S. [Proverb Comprehension Test: Development and Evaluation of psychometric properties (Persian)]. Journal of Modern Rehabilitation. 2016; 9(3):106-14. [Link]

[28] Ghawami H, Raghibi M, Tamini B, Dolatshahi B, Rahimi-Movaghar V. Cross-Cultural Adaptation of Executive Function Tests for Assessments of Traumatic Brain Injury Patients in Southeast Iran. Behavioral Psychology/Psicologia Conductual. 2016; 24(3):531-54. [Link]

[29] Rahimifar P, Soltani M, Latifi S, Majdinasab N, Moradi N. Reliability, validity, and normative investigation of Persian version of a High-Level Language Test (BESS). Applied Neuropsychology. Adult. 2020; 27(6):540-8. [DOI:10.1080/23279095.2019.1575221] [PMID]

[30] Liégeois FJ, Mahony K, Connelly A, Pigdon L, Tournier JD, Morgan AT. Pediatric traumatic brain injury: Language outcomes and their relationship to the arcuate fasciculus. Brain and Language. 2013; 127(3):388-98. [DOI:10.1016/j.bandl.2013.05.003] [PMID]

[31] Geraudie A, Battista P, García A, Allen I, Miller Z, Gorno-Tempini ML, et al. Speech and language impairments in behavioral variant frontotemporal dementia: A systematic review. 2021. [Preprint]. [DOI:10.1101/2021.07.10.21260313]

[32] Uekermann J, Thoma P, Daum I. Proverb interpretation changes in aging. Brain and Cognition. 2008; 67(1):51-7. [DOI:10.1016/j.bandc.2007.11.003] [PMID]

[33] Chuan CL, Penyelidikan J. Sample size estimation using Krejcie and Morgan and Cohen statistical power analysis: A comparison. Jurnal Penyelidikan ipbl. 2006; 7(1):78-86. [Link]

[34] Lotfi M-S, Tagharrobi Z, Sharifi K, Abolhasani J. [Diagnostic Accuracy of Persian Version of Clinical Dementia Rating (P-CDR) for Early Dementia Detection in the Elderly (Persian)]. Journal of Rafsanjan University of Medical Sciences. 2015; 14(4):283-98. [Link]

[35] Cieślicka A, Rataj K, Jaworska-Pasterska D. Figurative language impairment in aphasic patients. Neuropsychiatria i Neuropsychologia/Neuropsychiatry and Neuropsychology. 2011; 6:1-10. [Link]

[36] DeVellis RF, Thorpe CT. Scale Development: Theory and Applications. Thousand Oaks, CA: Sage Publications; 2021. [Link]

[37] Hambleton RK, Swaminathan, H., & Rogers, H. J. Fundamentals of item response theory. Thousand Oaks, CA, US: Sage Publications, Inc; 1991. [Link]

[38] Papagno C, Caporali A. Testing idiom comprehension in aphasic patients: The effects of task and idiom type. Brain and Language. 2007; 100(2):208-20. [DOI:10.1016/j.bandl.2006.01.002] [PMID]

This Page Intentionally Left Blank