

روایی آزمون درک خواندن فارسی آموزان غیر ایرانی: تحلیلی بر پایه اصول آزمون‌شناسی*

مریم کارخانه^۱

کارشناسی ارشد مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

محمّدباقر میرزایی حصاریان^۲

استادیار گروه آرفای دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

لیلا گل‌پور^۳

استادیار گروه آرفای دانشگاه بین‌المللی امام خمینی (ره)، قزوین، ایران

چکیده

آزمون‌ها ابزار مهمی جهت ارزیابی یادگیری و توانمندی فارسی‌آموزان هستند. روایی آزمون یکی از مهم‌ترین عواملی است که در ارزشیابی آزمون باید به آن توجه کرد. هدف اصلی این پژوهش بررسی روایی محتوایی و ظاهری آزمون درک خواندن سطح تکمیلی فارسی‌آموزان غیر ایرانی است. فرضیه‌ی این پژوهش بر این اصل استوار است که این آزمون از روایی ظاهری نامطلوب و روایی محتوایی مطلوبی برخوردار است. پرسش‌شوندگان این پژوهش ۱۰ مدرس آموزش زبان فارسی به غیر فارسی‌زبانان و ۲۴ نفر از فارسی‌آموزان دختر و پسر مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) با ملیت‌های گوناگون در بازه سنی ۳۰-۲۰ است. در این پژوهش برای بررسی روایی آزمون‌ها از دو پرسشنامه‌ی آزمون‌های وابسته به ملاک ویرسما و جورس و نیز افزودن شاخص‌های تکمیلی، طبق مدل سامونلز و نظریه سالیوا هاگیس استفاده شده‌است. دو مشخصه‌ی روایی محتوایی و ظاهری هر دو آزمون اندازه‌گیری شدند. نتایج آماری نشان‌دهنده‌ی این موضوع بوده‌اند که این دو آزمون، از روایی محتوایی قابل قبولی برخوردارند اما روایی ظاهری مطلوبی ندارند.

واژه‌های کلیدی: آزمون مهارت درک خوانداری، روایی، فارسی‌آموزان غیر ایرانی.

* تاریخ دریافت: ۱۴۰۱/۰۲/۰۸ | تاریخ پذیرش: ۱۴۰۱/۰۶/۲۱

(نویسنده مسئول)

¹ Email: maryamkarkhane65@gmail.com

² Email: azfaconference1400@gmail.com

³ Email: niyayesh_1@yahoo.com.au



Validity of non-Iranian Students' Reading Comprehension Test: An Analysis Based on the Principles of Testology *

Maryam karkhaneh ¹

M.A. in Persian Language Education Center of Imam Khomeini International University, Qazvin, Iran

Mohammad Bagher Mirzaei Hesarian ²

Assistant Professor, Azfa Department, Imam Khomeini International University, Qazvin, Iran

Laila Golpour ³

Assistant Professor, Azfa Department, Imam Khomeini International University, Qazvin, Iran

Abstract

Tests are an important tool to evaluate the learning and ability of Persian learners. The validity of the test is one of the most important factors that should be considered in the evaluation of the test. The main purpose of this research is to investigate the content and face validity of the supplementary level reading comprehension test of non-Iranian Farsi learners. The hypothesis of this research is based on the principle that this test has unfavorable face validity and favorable content validity. The respondents of this research are 10 teachers of Persian language teaching to non-Persian speakers and 24 male and female Persian learners of the Persian Language Center of Imam Khomeini International University with different nationalities in the age range of 20-30. In this research, to check the validity of the tests, two questionnaires of the tests related to Wiersma and Jours criteria and also adding additional indicators, according to Samuels model and Saliva-Haggis theory, were used. The content and face validity characteristics of both tests were measured. Statistical results have shown that these two tests have acceptable content validity, but they do not have good face validity.

Keywords: Reading skills test, validity, non-Iranian Persian students.

* Received: 2022/04/28 | Accepted: 2022/09/12

¹ Email: maryamkarkhane65@gmail.com (Corresponding Author)

² Email: azfaconference1400@gmail.com

³ Email: niyayesh_1@yahoo.com.au

۱- مقدمه و بیان مسأله

آزمون‌سازی یکی از حوزه‌های مهم در زمینه‌ی زبان‌آموزی است که روز به روز بر میزان اهمیت و اثربخشی آن در گستره‌های مرتبط با زبان افزوده می‌شود. آزمون وسیله‌ای عینی و استاندارد شده‌است که برای اندازه‌گیری نمونه‌ای از رفتار یا خصایص آدمی به کار می‌رود. روشن است که اعتبار، راز ماندگاری و نقش آزمون در تحولات اجتماعی و تأثیر آن بر زندگی فردی حائز اهمیت است و اگر به صورت هدفمند برای اصلاح و بهبود بخشی به آموزش به کار گرفته شود، تأثیر مثبتی بر روند آموزش و یادگیری خواهد داشت.

مهارت خواندن یکی از مهارت‌های اصلی آموزش زبانی است که یادگیری آن مستلزم تمرین و تسلط بر ریزمهارت‌های خواندن است؛ تسلط بر ریزمهارت‌های خواندن نیز مستلزم آگاهی و استفاده از راهکارها و راهبردهای کارآمد است. خواندن موثرترین وسیله برای تحکیم و گسترش دانش و از جمله دانش زبانی است. هریس و سایپی^۱ خواندن را چنین تعریف می‌کنند: «خواندن، تفسیر معنادار نمادهای کلامی نوشته شده‌است». استفرانک اسمیت^۲ خواندن را «کاهش مجهولات» می‌داند. مهارت در خواندن و درک مطلب، از مهم‌ترین نیازهای یادگیری زبان آموزان است. برای ارزشیابی درک خواندن چند مورد وجود دارد که باید در نظر گرفته شود: ۱- سنجش دایره لغات و دستور زبان ۲- انواع پردازش اطلاعات در مهارت خواندن ۳- گونه‌های خواندن (براون، ۲۰۰۴: ۱۸۶).

به باور چاو^۳ (۲۰۱۰) درک خواندن مدت‌هاست که به عنوان یک روش مهم برای سنجش توانایی زبان‌آموزان تلقی می‌شود و توانایی خواندن، زمان استفاده از یک زبان خارجی بدون شک مفید است. وی همچنین عنوان می‌کند که برخی از آزمون‌های خواندن روایی ندارند و آزمون‌های کارآمدی نیستند و طراحان آزمون باید به خوبی از این مسئله آگاه باشند که آزمون خواندن می‌تواند به ما در ارزیابی کیفیت تدریس کمک کند؛ بنابراین روایی یک آزمون خواندن مستقیماً ارزیابی تدریس ما را تحت تأثیر قرار می‌دهد و راهنمای خوبی برای تدریس محسوب می‌شود.

یکی از انواع روایی آزمون، روایی محتوایی است. در بررسی روایی محتوایی به این

¹ Harries & saipy

² Stefrenk smith

³ chao

موضوع توجه می‌شود که تا چه اندازه آزمون همان چیزی را که در پی سنجشش است، می‌سنجد؛ بنابراین بررسی روایی محتوایی آزمون، امری بسیار ضروری و با اهمیت است، چرا که چنانچه آزمونی به اندازه مطلوب از این ویژگی برخوردار نباشد، نتایج حاصل از آن اساساً بی‌اعتبار خواهد بود و اهداف مورد نظر ارزشیابی را محقق نخواهد کرد. بررسی روایی محتوایی آزمون‌های درک خواندن فارسی آموزان غیرایرانی اهمیت دوچندان دارد زیرا این مهارت یکی از دو کانال دریافت اطلاعات (دیدن و شنیدن) محسوب می‌شود و می‌توان گفت: که فارسی‌آموزانی که قصد تحصیل در دانشگاه‌های ایران را دارند بیشتر به مهارت خواندن نسبت به مهارت‌های دیگر زبان فارسی متکی هستند؛ چرا که آنها باید منابع درسی را مطالعه و مطالب اصلی و مهم را شناسایی کنند و با داشتن مهارت کافی در خواندن متون گوناگون، در درک این متون توانمند باشند و مسائل زیادی یاد بگیرند. بنابراین اگر آزمون درک خواندن مراکز زبان فارسی استاندارد نباشد و از ضریب تمیز خوبی برخوردار نباشد موجب بی‌ارزش شدن این مهارت در یادگیری می‌گردد و فارسی‌آموزان نمی‌توانند دوره‌های تحصیلی خود را به خوبی طی کنند و با مشکلات زیادی در امر یادگیری مواجه می‌شوند. در پژوهش حاضر تلاش می‌شود تا روایی ظاهری و محتوایی آزمون درک خواندن سطح تکمیلی مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی^(۹) که ابزار مهمی برای ارزیابی میزان پیشرفت و توانمندی فارسی‌آموزان است، بررسی شود. پرسش‌های پژوهش عبارتند از:

۱- روایی ظاهری آزمون‌های درک خواندن فارسی‌آموزان غیرایرانی سطح تکمیلی چگونه است؟

۲- روایی محتوایی آزمون‌های درک خواندن فارسی‌آموزان غیرایرانی سطح تکمیلی چگونه است؟

متناظر با پرسش‌های فوق فرضیه‌های زیر در نظر گرفته شده است:

۱- روایی ظاهری آزمون‌های درک خواندن فارسی‌آموزان غیرایرانی سطح تکمیلی در سطح مطلوب است.

۲- روایی محتوایی آزمون‌های درک خواندن فارسی‌آموزان غیرایرانی سطح تکمیلی در سطح مطلوب نیست.

۲- پیشینه پژوهش

قنصولی هزاره (۱۳۸۹) به طراحی و روایی سنجی آزمون بسندگی فارسی پرداخته و برای این منظور از فرآیندی دو مرحله‌ای، شامل تحلیل داده‌های بازنگری سه دانشجوی خارجی دوره دکترای زبان و ادبیات فارسی و بهره‌گیری از نظریه بحث در سنجش روایی بهره گرفته‌است. وی به دلیل ابعاد مهم و تأثیرگذار آزمون در قلمرو سیاسی و فرهنگی کشور، به تولید و روایی سنجی آزمون ملی بسندگی زبان فارسی و اهمیت آزمون سازی زبان و ارتباط آن با دوره‌های آموزش زبان فارسی توجه کرده‌است یافته‌های این تحقیق نشان داد که فارسی آموزان در بخش خواندن و درک مفاهیم بهترین عملکرد را دارند و بخش دستور دشوارترین بخش برای آنهاست در حالی که پایایی آزمون بسندگی طراحی شده ۷۰ درصد بود.

صحرائی و ممقانی (۱۳۹۱) به پژوهشی با هدف ارزیابی پایایی و روایی آزمون زبان انگلیسی وزارت علوم و تحقیقات و فناوری (MSRT) پرداخته‌اند. برای ارزیابی روایی این آزمون از روایی سازه‌ای همسانی درونی و برای ارزیابی پایایی درونی نیز از روش آلفای کرونباخ استفاده کرده‌اند. نتایج این پژوهش نشان داده که آزمون MSRT در سطح کلی دارای روایی و پایایی مناسبی است اما در سطح جزئی از روایی لازم برخوردار نیست.

جباری و خادمی (۱۳۹۲) در پژوهش خود به ساخت و بررسی روایی و پایایی آزمون تشخیصی خواندن و درک مطلب با ۱۱ خرده آزمون که شامل: سلیس خوانی، توسعه واژگان، آگاهی واج شناختی، ایده اصلی متن، ارجاعات محتوایی، حقایق و ایده‌ها در متن، ارتباطات علی و معلولی در متن، توالی وقایع، گرفتن نتایج بر اساس شواهد، تفسیر و هدف نویسنده است، پرداخته‌اند. جامعه آماری این پژوهش ۱۲۰۰ دانش آموز از کلاس‌های سوم، چهارم و پنجم شهر شیراز است که به روش نمونه‌گیری خوشه‌ای انتخاب شده‌اند. یافته‌های این پژوهش بیانگر آن است که روایی ملاکی این آزمون در سطح بالا و پایایی در سطح مورد قبولی است و خرده آزمون‌های سلیس خوانی، واژگان و ایده اصلی متن در این آزمون از روایی بالایی برخوردار است. گل‌پور و وکیلی‌فرد و زندی (۱۳۹۸) در پژوهشی برای طراحی آزمون مهارت (بسندگی) زبان فارسی بر پایه‌ی چهار مهارت زبانی درک شنیداری، درک خوانداری،

صحبت کردن و نگارش، ۱۳۰ فارسی آموز مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی^(ره) را انتخاب کردند؛ همچنین به منظور استانداردسازی آزمون، سؤالات آزمون بسندگی طراحی شده به صورت پیش‌آزمون در آزمون‌های پایان ترم بهار و تابستان ۹۴ مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی^(ره) مورد استفاده قرار دادند. بعد از آسیب شناسی سؤالات و تجزیه و تحلیل، نتایج نشان داد آزمون بسندگی طراحی شده ضریب دشواری و ضریب تمیز بهینه‌ای دارد و از پایایی و روایی قابل قبولی نیز برخوردار است.

مانکسیا^۱ (۲۰۰۸) در پژوهش خود به بررسی ویژگی‌های روایی محتوایی بخش خواندن آزمون‌های جامع، آزمون ملی انگلیسی برای ورود به دانشگاه (NMET) بر اساس چارچوب پیشنهادی پالم (۱۹۹۹) پرداخته است. در این آزمون غنای واژگانی و سرعت و توانایی خواندن با توجه به ویژگی‌های آزمون خواندن ارزیابی شده و با تجزیه و تحلیل داده‌ها چنین نتیجه‌گیری شده که بخش‌های درک مطلب آزمون خواندن NMET روایی محتوایی بالایی دارد.

دونگ^۲ (۲۰۱۱) روایی محتوایی درک خواندن آزمون TEM-8 (آزمون استاندارد برای تخصص در سطح ملی رشته انگلیسی) که یکی از معیارهای قضاوت در زمینه تدریس و یادگیری در چین است را از سال ۲۰۰۸ تا ۲۰۱۰ با توجه به فرایندهای جدیدی براساس چارچوب پالم (۱۹۹۶) مورد بررسی قرار داده است. رویه‌ی تحقیقاتی در سه مرحله انجام شده، مرحله اول چارچوب سنجش ارزیابی خواندن ارتباطی که شامل چارچوب ویژگی‌های آزمون ارتباطی و ماهیت خواندن است و مرحله دوم با تجزیه و تحلیل روایی درک مطلب از سال ۲۰۰۸ تا ۲۰۱۰ از نظر دوره تدریس برای TEM-8 و در مرحله سوم برنامه تدریس TEM-8 را مورد تجزیه و تحلیل قرار داده، یافته‌های این پژوهش نشان داد که بخش درک خواندن در این آزمون دارای روایی بالایی است.

ازر^۳ و همکاران (۲۰۱۳) در پژوهشی به ارزیابی پایایی و روایی محتوایی آزمون انگلیسی به عنوان یک زبان خارجی برای دانش‌آموزان مقطع ابتدایی ترکیه پرداخته‌اند. جامعه آماری این پژوهش ۷۱ دانش‌آموز است که ۳۸ نفر دختر و ۳۳ نفر پسر هستند. معیارهای روایی محتوایی در

¹ Manxia

² Dong

³ Ozer

این آزمون واضح بودن، مرتبط بودن، عدم وجود ابهامات در کلمات و واضح بودن دستورالعمل بوده و برای تعیین روایی محتوایی این آزمون از چهار نفر متخصص استفاده شده است. نتایج آزمون EFL نشان داده که این آزمون ضریب پایایی بالا و روایی محتوایی ناموفقی دارد.

سانتوس^۱ و همکاران (۲۰۱۷) پژوهشی با عنوان روایی آزمون خواندن کلمه برای دانش آموزان ابتدایی پرتغالی انجام داده‌اند. جامعه آماری این پژوهش ۲۸۰ دانش آموز پرتغالی بوده، یافته‌های این پژوهش که از تجزیه و تحلیل فرم‌های TLP (Teste de Leitura de Palavras) آزمون خواندن کلمه) ابزارهای خواندن، حافظه کوتاه مدت و واژگان به دست آمده دارای همبستگی معناداری است و این آزمون روایی محتوایی دارد و نشان می‌دهد که ساختار متغیرهای این آزمون برای ارزیابی خواندن دانش آموزان پرتغالی مناسب است.

۳- مبانی نظری

بنا به تعریف سیف (۱۳۸۵: ۳۳) آزمون وسیله‌ای عینی و استاندارد است که برای اندازه‌گیری و سنجش نمونه‌ای از حالات، رفتار و معلومات به طور کلی رفتار و خصایص فرد یا افراد به کار می‌رود. کرمی (۱۳۸۲) در تعریف آزمون این طور بیان می‌کند که مفهوم آزمون از مفهوم اندازه‌گیری محدودتر است و معمولاً نمونه‌ای از رفتار را اندازه می‌گیرد. هر آزمون از چند سؤال تشکیل شده است که جایگاه آزمودنی را در صفت مورد اندازه‌گیری تعیین می‌کند. زندی (۱۳۸۵) آزمون را وسیله‌ای برای اندازه‌گیری می‌داند که آزمون شونده تکلیف بخصوصی را با آن انجام می‌دهد.

۳-۱- طبقه‌بندی آزمون‌ها

اساساً آزمون‌ها برای دو هدف کلی به کار گرفته می‌شوند: یکی ارزشیابی دریافت یافته‌ها و دیگری پیش‌بینی توان آزمودنی در آینده. هدف از آزمون دریافت یافته‌ها، سنجش مهارت یا توانی است که آزمودنی تا زمان انجام آزمون به دست آورده است؛ در حالی که هدف آزمون پیش‌بینی، تعیین میزان مهارت یا توانی است که آزمودنی در زمان معینی در آینده می‌تواند دارا باشد (جعفرپور، ۱۳۸۴: ۱۱). آزمون‌های دریافت یافته‌ها به سه دسته

¹ Santos

تقسیم می‌شوند: آزمون پیشرفت، آزمون بسندگی و آزمون دانش. آزمون پیشرفت شامل آزمون رشد، آزمون چیرگی و آزمون تشخیصی است و آزمون‌های پیش بین شامل ۳ آزمون تعیین سطح (جایابی)، گزینش و آزمون استعداد است که در هر یک از این آزمون‌ها، آزمون درک خواندن به عنوان بخشی از آن‌ها مورد استفاده قرار می‌گیرد.

۳-۲- آزمون درک خواندن

مهارت خواندن و درک مطلب در همه زمینه‌ها و جنبه‌های آموزش یا حتی زندگی، نقش مهمی دارد. خواندن و درک مطلب هدایت‌شده و یا ساختارمند بخش مهمی از یادگیری این مهارت محسوب می‌شود. خواندن به عنوان فرآیند بیرون کشیدن معنی از متن نوشته شده برای برخی هدفها تعریف شده است. بنا به تعریف سیف (۱۳۸۵: ۳۳) آزمون وسیله‌ای عینی و استاندارد است که برای اندازه‌گیری و سنجش نمونه‌ای از حالات، رفتار و معلومات به طور کلی رفتار و خصایص فرد یا افراد به کار می‌رود.

جعفرپور (۱۳۸۴: ۹۸-۱۰۰) ویژگی‌های یک آزمون خوب را اینگونه بیان می‌کند: «آزمون خوب دست کم دارای چهار ویژگی روایی، پایایی، عملی بودن و ارتباط است که این چهار ویژگی را می‌توان اصول آزمون‌سازی معرفی کرد». برای ارزشیابی مهارت درک خواندن چند مورد وجود دارد که باید در نظر گرفته شود ۱- سنجش دایره لغات و دستور زبان ۲- انواع پردازش اطلاعات در مهارت خواندن ۳- گونه‌های خواندن (براون، ۲۰۰۴: ۱۹۰-۱۸۶). شایان ذکر است که آزمون‌های مرکزهای زبان ابزار مهمی جهت ارزیابی یادگیری و توانمندی فارسی‌آموزان در مهارت‌های زبان است؛ بنابراین آزمون درک خواندن نیز اگر ویژگی‌های آزمون خوب را نداشته باشد و از ضریب تمیز خوبی برخوردار نباشد اثرات منفی زیادی در آموزش و یادگیری فارسی‌آموزان به دنبال خواهد داشت و ممکن است فارسی‌آموزان نتوانند دوره‌های تحصیل خود را به خوبی طی کنند و حتی ممکن است از اعتبار مراکز زبان هم کاسته شود.

۳-۳- روایی

عباس زادگان و فتوت (۱۳۸۴: ۹۵) روایی را این چنین بیان می‌کنند: توافق بین نمره آزمون با صفت یا خصیصه‌ای که آزمون برای اندازه‌گیری آن ساخته شده است. به عبارتی دیگر مقصود از روایی آن است که وسیله اندازه‌گیری در واقع همان خصیصه مورد نظر را

اندازه بگیرد، نه خصیصه دیگری را. بنابراین اگر معلوم شود که آزمون چه چیزی را اندازه می‌گیرد و نمره آزمون توافق وجود داشته باشد، نتایج حاصل از اجزای آزمون می‌تواند اطلاعاتی به دست دهد که در امر تصمیم‌گیری مفید واقع شوند. مقصود از روایی یک وسیله اندازه‌گیری، مناسب بودن، مربوط بودن، با معنا بودن و مفید بودن استنباط‌های خاصی است که از روی نمره‌های آن به دست می‌آید. به باور سیف (۱۳۹۵: ۵۳۴) آزمونی دارای روایی است که برای اندازه‌گیری آنچه مورد نظر است، کافی و مناسب باشد. برای مثال یک آزمون پیشرفت تحصیلی ریاضی کلاس پنجم دبستان در صورتی یک آزمون رواست که محتوا و هدف‌های آن درس را به خوبی اندازه‌گیری کند و به جز یادگیری دانش‌آموزان از آن درس چیز دیگری را شامل نباشد. پس روایی یک امر نسبی است، نه یک ویژگی که هست یا نیست. روایی به معنی آن است که آیا روش انتخاب شده برای سنجش موضوع مورد نظر مناسب است و آن را می‌سنجد یا نه و یا به عبارتی دیگر ابزار اندازه‌گیری تا چه حد خصیصه مورد نظر را می‌سنجد.

۳-۳-۱- روایی سازه‌ای

روایی سازه‌ای ارزشمندترین و در عین حال دشوارترین راه سنجش ابزار پیمایش است و مقصود آن است که نمرات حاصل از اجرای ابزار چقدر به مفاهیم یا سازه‌های نظریه مورد نظر مربوط است. روایی سازه از نظر درک، اندازه‌گیری و اعلام دشوار است این نوع روایی، اندازه‌ای است مربوط به معنی دار بودن مقیاس یا ابزار پیمایش در کاربرد عملی آن عباس زادگان و فتوت (۱۳۸۴: ۱۰۶). روایی سازه بیشتر از روایی محتوایی و ملاکی جنبه نظری دارد (سیف، ۱۳۹۵: ۵۴۶). روایی سازه به این نکته توجه دارد که آزمون تهیه شده، تا چه اندازه ویژگی یا توانایی یا سازه‌ی ویژه‌ای را اندازه می‌گیرد (جعفر پور، ۱۳۷۵: ۸۷).

۳-۳-۲- روایی سنجه‌ای

منظور از روایی سنجه‌ای یا (ملاکی) میزان ارتباط بین نمرات حاصل از یک آزمون با نمرات حاصل از یک آزمون دیگر یا وسیله اندازه‌گیری دیگر است. بنابراین هر زمان که از نمرات یک آزمون برای پیش‌بینی عملکرد افراد در آزمونی دیگر استفاده کنیم، با روایی

ملاکی سروکار داریم. آزمون دوم که عملکرد فرد در آن پیش‌بینی می‌شود، ملاک نام دارد.

۳-۳-۳- روایی ظاهری

در روایی ظاهری به جای توجه به ماهیت محتوای آزمون به ویژگی‌های ظاهری آن مانند کلمات و جملات آزمون، صورت ظاهری آزمون و قابلیت پذیرش و معقول بودن آزمون برای آزمودنی‌ها توجه می‌شود. در این نوع روایی بررسی می‌شود که آیا کلمات و جملات به کار رفته در آزمون برای آزمودنی‌ها قابل درک و فهم است یا خیر؛ چون اگر مطالب قابل درک نباشند، انگیزه پاسخگویی را در آزمودنی‌ها ایجاد نمی‌کند. شکل ظاهری هر آزمون مانند انتخاب عناوین مناسب و هماهنگ با اسم آزمون، نوع تایپ و تصاویر به کار رفته در آزمون و حتی صفحه‌آرایی دفترچه سؤالات می‌تواند انگیزه پاسخگویی به سؤال‌ها را افزایش دهد. در ضمن به کار بردن کلماتی در سؤال‌ها که با مقوله مورد اندازه‌گیری مطابقت داشته باشد، نقش موثرتری در این زمینه دارد.

اهمیت انجام روایی ظاهری در این است که آزمودنی با رغبت به آزمون پاسخ می‌دهد. به عبارت دیگر روایی ظاهری، گرچه ضامن اندازه‌گیری دقیق نیست، ولی ممکن است با انگیزش آزمودنی به پاسخگویی بر روایی نمرات حاصل از آزمون تأثیر داشته باشد (کرمی، ۱۳۸۲: ۲۰۱). از نظر عباس زادگان و فتوت (۱۳۸۴) روایی ظاهری منطقی بودن، جالب بودن و تناسب ظاهری ابزار پژوهش است و زمانی اهمیت می‌یابد که ظاهر سؤال‌ها انگیزش، رغبت و تمایل آزمودنی را به همکاری و پاسخ دادن ایجاد نماید. روایی ظاهری مبتنی بر مرور سریع سؤال‌ها توسط داوران ناکار آزموده مانند خواهر، دوست یا همکلاسی است.

سنجش روایی ظاهری ممکن است فقط شامل نشان دادن پیمایش به چند فرد غیر مجرب باشد. این نوع روایی کمتر از دیگر انواع روایی علمی است و غالباً با روایی محتوایی اشتباه می‌شود. گرچه این دو به یکدیگر شبیه هستند، اما روایی ظاهری سنجش بسیار غیر رسمی‌تر از متناسب بودن سؤال‌هاست. به اعتقاد سیف (۵۱۳۹۵: ۵۳۶) یکی از مشتقات روایی محتوایی، روایی ظاهری است که بیانگر آن است که سؤال‌های آزمون تا چه حد در ظاهر

شبیبه به موضوعی هستند که برای اندازه‌گیری آن تهیه شده‌اند. روایی ظاهری یک ویژگی مطلوب آزمون است؛ زیرا ظاهر آزمون بر انگیزش آزمون‌شوندگان اثر می‌گذارد. اگر شکل ظاهری و نوع ماده‌های آزمونی به نظر آزمودنی‌ها بهینه نباشد، این برداشت ممکن است در کارکرد آنها تأثیر منفی بگذارد و هدف آزمونگر از انجام آزمون که تعیین تولدایی که حق اوست، برآورده نخواهد شد. اگر چه کیفیت به نظر رسیدن، جنبه‌عینی ندارد، اما به دلیل نقشی که ظاهر آزمون می‌تواند در کارکرد آزمودنی داشته باشد، روایی ظاهری از اهمیت ویژه‌ای برخوردار است (جعفر پور، ۱۳۸۴: ۸۵).

۳-۳-۴- روایی محتوایی

این نوع روایی بیانگر آن است که نمونه سؤال‌ها، تکالیف یا پرسش‌های ابزار سنجش تا چه حد معرف مهارت‌ها، توانایی‌ها، درک مفاهیم و دیگر رفتارهاست. برای این منظور باید سؤال‌های آزمون بر اساس هدف‌های رفتاری و محتوای مورد اندازه‌گیری، تدوین شود. روایی محتوایی عبارت است از اندازه‌های ذهنی در خصوص میزان متناسب بودن سؤال‌ها از نظر دسته‌ای از افراد که در مورد سؤال‌ها اطلاعاتی دارند و در آنها تجدید نظر می‌کنند، سنجش روایی محتوایی نوعاً شامل تجدید نظری سازمان یافته در محتوای پیمایش است تا اطمینان یابیم واجد تمامی چیزهایی که باید و فاقد تمامی چیزهایی که نباید است (عباس زادگان و فتوت، ۱۳۸۴: ۹۷-۹۶).

از آنجا که همه آزمون‌های پیشرفت تحصیلی در یک زمان محدود به کار می‌روند و از آنجا که هیچ آزمونی نمی‌تواند تمام محتوا و کلیه هدف‌های یک درس را شامل شود پس یک آزمون از مجموعه‌ای سؤال که به عنوان نمونه‌ای از تمامی محتوا و هدف‌های آموزشی درس انتخاب شده است تشکیل می‌یابد روایی محتوایی به این مطلب اشاره می‌کند که نمونه سؤال‌های مورد استفاده در یک آزمون تا چه حد معرف تمامی محتوا یا کل جامعه سؤال‌های ممکن است که می‌توان از محتوا یا موضوع مورد نظر تهیه کرد. هر چه آزمون از این لحاظ بهتر باشد دارای روایی بیشتری است (سیف، ۱۳۹۴: ۵۳۴).

بررسی محتوای ابزار در تعیین روایی محتوایی آزمون یکی از مهم‌ترین بخش‌های ارزیابی روایی یک ابزار است. هدف از این ارزیابی پاسخ به این سؤال است که آیا محتوای

ابزار، قابلیت اندازه‌گیری هدف تعریف شده را دارد یا خیر؟ (حاجی‌زاده و اصغری، ۱۳۹۴: ۳۹۹-۳۹۸). این که آزمون تا چه اندازه نمونه‌هایی از هدف‌های رفتاری مورد سنجش را در برمی‌گیرد، به روایی محتوایی آن مربوط است؛ یعنی محتوای آزمون تا چه اندازه مطالب مورد اندازه‌گیری را در بردارد (جعفرپور، ۱۳۷۵: ۸۵).

۳-۴- روش تعیین روایی ظاهری

در تعیین کیفی روایی ظاهری موارد سطح دشواری، میزان عدم تناسب و ابهام مورد بررسی قرار می‌گیرد و اصلاح می‌شود. در گام بعدی جهت کاهش و حذف آیتم‌های نامناسب و تعیین اهمیت هر یک از آیتم‌ها از روش کمی نمرات تأثیر آیتم استفاده خواهد شد (حاجی‌زاده و اصغری، ۱۳۹۴: ۴۰۲-۴۰۱). برای بررسی نمرات تأثیر، ابتدا از شرکت‌کنندگان خواسته می‌شود تا میزان اهمیت هر یک از سؤالات آزمون را در یک طیف لیکرتی ۵ قسمتی مشخص نمایند؛ سپس نمرات تأثیر از طریق فرمول زیر محاسبه می‌شود:

$$\text{نمره تأثیر} = \text{فراوانی} \times \text{میزان اهمیت}$$

منظور از فراوانی، درصد مشارکت‌کنندگانی است که هر آیتم را بررسی کرده‌اند و میزان اهمیت، متوسط پاسخ مشارکت‌کنندگان به گزینه‌های مربوط به اهمیت هر گویه است. برای پذیرش روایی ظاهری هر گویه، نمره تأثیر آن نبایستی کمتر از ۱/۵ باشد و فقط سؤالاتی از لحاظ روایی ظاهری قابل قبول هستند که نمره آن‌ها بالاتر از ۱/۵ باشد.

۳-۵- روش تعیین روایی محتوایی

اولین گام در تعیین روایی آزمون، بررسی روایی محتوایی است. روایی محتوایی به تحلیل منطقی محتوای یک آزمون بستگی داشته و تعیین آن بر اساس قضاوت ذهنی و فردی است. برای تعیین روایی محتوایی از دو روش کیفی و کمی استفاده می‌شود. در بررسی کیفی محتوا، پژوهشگر از متخصصان درخواست می‌کند تا پس از بررسی کیفی ابزار بر اساس معیارهای رعایت دستور زبان، استفاده از کلمات مناسب و امتیازدهی مناسب قرارگیری آیتم‌ها در جای مناسب خود، بازخورد لازم را ارائه دهند که بر اساس آن موارد اصلاح خواهند شد.

برای بررسی روایی محتوایی به شکل کمی از دو ضریب نسبت روایی محتوایی (CVR)^۱ و شاخص روایی محتوایی (CVI)^۲ استفاده می‌شود. برای تعیین نسبت روایی محتوایی، که توسط لاوشه^۳ (۱۹۷۵) طراحی شده است، عمدتاً از قضاوت داوران صاحب نظر استفاده می‌شود؛ بدین صورت که از افراد متخصص درخواست می‌شود تا هر آیت‌م آزمون را بر اساس طیف سه قسمتی (ضروری است، مفید است ولی ضروری نیست، ضرورتی ندارد) بررسی نمایند، سپس پاسخ‌ها بر اساس فرمول محاسبه می‌گردند. تبدیل قضاوت کیفی داوران درباره‌ی روایی محتوایی به کمیت که ویژه آزمون‌های وابسته به ملاک یا ملاکی است، به روش لاوشه^۴ شهرت دارد (سیف، ۱۳۹۵: ۵۵۱).

$$CVR = ne - N/2$$

ne تعداد افراد متخصصی است که به گزینه ضروری پاسخ داده‌اند و N تعداد کل افراد متخصص است. مقدار CVR محاسبه شده از فرمول با مقدار متناظر آن از جدول مقایسه و تصمیم‌گیری خواهد شد. چنانچه CVR محاسبه شده از CVR جدول بیشتر باشد، اعتبار محتوای آن آیت‌م تایید می‌گردد.

جدول ۱: حداقل مقدار CVR قابل قبول بر اساس تعداد متخصصین نمره‌گذار

حداقل مقدار CVR قابل قبول بر اساس تعداد متخصصین نمره‌گذار					
تعداد متخصصین	مقدار CVR	تعداد متخصصین	مقدار CVR	تعداد متخصصین	مقدار CVR
۵	۰/۹۹	۱۱	۰/۵۹	۲۵	۰/۳۷
۶	۰/۹۹	۱۲	۰/۵۶	۳۰	۰/۳۳
۷	۰/۹۹	۱۳	۰/۵۴	۳۵	۰/۳۱
۸	۰/۷۵	۱۴	۰/۵۱	۴۰	۰/۲۹
۹	۰/۷۸	۱۵	۰/۴۹		
۱۰	۰/۶۲	۲۰	۰/۴۲		

برای بررسی شاخص روایی محتوا (CVI) از روش والتز و باسل استفاده شد. بدین صورت که متخصصان «مرتبط بودن»، «واضح بودن» و «ساده بودن» هر گویه را بر اساس یک

¹ Content Validity ratio

² Content validity index

³ Lawshe

⁴ Lawshe

طیف لیکرتی ۴ قسمتی مشخص کردیم. متخصصان مربوط بودن هر گویه را از نظر خودشان از ۱ «مربوط نیست»، ۲ «نسبتاً مربوط است»، ۳ «مربوط است»، تا ۴ «کاملاً مربوط است» تعیین کردند. ساده بودن گویه نیز به ترتیب از ۱ «ساده نیست»، ۲ «نسبتاً ساده است»، ۳ «ساده است»، تا ۴ «کاملاً ساده است» و واضح بودن گویه نیز به ترتیب از ۱ «واضح نیست»، ۲ «نسبتاً واضح است»، ۳ «واضح است»، تا ۴ «کاملاً واضح است» مشخص شد. حداقل مقدار قابل قبول برای شاخص CVI برابر با ۰/۷۹ است و اگر شاخص CVI گویه‌ای کمتر از ۰/۷۹ باشد آن گویه بایستی حذف شود (حاجی زاده و اصغری، ۱۳۹۴: ۴۰۰).

۴- روش پژوهش

پژوهش حاضر از نوع توصیفی - تحلیلی است.

۴-۱- جامعه آماری پژوهش

پرسش‌شوندگان اول این پژوهش را ۱۰ مدرس آموزش زبان فارسی به غیر فارسی‌زبانان، که ۳ نفر دارای مدرک تحصیلی دکتری و ۷ نفر آقا و خانم کارشناسی ارشد در این رشته و متخصص در امر آموزش به غیرفارسی‌زبانان بودند، تشکیل داده‌اند. پرسش‌شوندگان دوم تعداد ۲۴ نفر از فارسی‌آموزان غیر ایرانی ۱۴ دختر و ۱۰ پسر با ملیت‌های مختلف اعم از چینی، عربی، ژاپنی و بنگلادشی سطح تکمیلی مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی^(د) قزوین تشکیل داده است.

۴-۲- ابزار پژوهش

دو بسته آزمون درک خواندن سطح تکمیلی مرکز آموزش زبان فارسی امام خمینی سال تحصیلی ۹۶-۹۷ آزمون اول شامل ۳۰ پرسش است و آزمون دوم ۳۵ پرسش دارد. پرسش‌های این دو آزمون به صورت چهارگزینه‌ای و برگرفته از چند متن است. در هر دو آزمون سؤال اول از نوع کلوزتست هست با متنی با موضوع علوم زیستی که ۱۵ پرسش برای این متن طراحی شده و در آزمون دوم ۱۱ پرسش و در هر دو آزمون این سؤالات خرده آزمون غنای واژگانی را در بر می‌گیرد. در هر دو آزمون چهار متن داده شده است. متن‌ها کوتاه و بلند هستند در آزمون اول تعداد سؤالات طراحی شده برای هر متن ۵ پرسش است

در آزمون دوم برای هر متن ۴ پرسش در نظر گرفته شده است که پرسش‌های این آزمون، خرده آزمون‌های (توسعه واژگان، آگاهی واج شناختی، ایده اصلی متن، ارجاعات محتوایی، حقایق و ایده‌ها در متن، ارتباطات علی و معلولی در متن، توالی وقایع، گرفتن نتایج بر اساس شواهد تفسیر و هدف نویسنده) را در بر می‌گیرد.

ابزار دوم این پژوهش پرسشنامه فرم درجه‌بندی برای سنجش روایی محتوایی آزمون‌های وابسته به ملاک ویرسما و جورس (۱۹۹۰) و شاخص‌های تکمیلی که طبق مدل نظریه سامونلز و سالیوا هاگیس (۲۰۰۰) بود، به این پرسشنامه افزوده شد که برای تعیین کمی روایی محتوایی سؤالات آزمون از این پرسشنامه استفاده شد. این پرسشنامه شامل خرده آزمون درک خواندن (توسعه واژگان، آگاهی واج شناختی، ایده اصلی متن، ارجاعات محتوایی، حقایق و ایده‌ها در متن، ارتباطات علی و معلولی در متن، توالی وقایع، گرفتن نتایج بر اساس شواهد، تفسیر و هدف نویسنده بود) که از طیف ۱ تا ۵ بررسی شد و عدد ۱ کمترین و عدد ۵ بیشترین روایی را نشان می‌داد. در پرسشنامه دوم برای به دست آوردن نمره تأثیر هر یک از ۶۵ سؤال آزمون‌ها طیف لیکرتی ۵ قسمتی با امتیاز (کاملاً مهم ۵)، (مهم است ۴)، (به طور متوسط مهم است ۳)، (اندکی مهم است ۲) و (اصلاً مهم نیست ۱) در نظر گرفته شده که از این پرسشنامه برای تعیین کمی روایی ظاهری استفاده شد.

۴-۳- شیوه اجرای پژوهش

جهت تعیین روایی محتوایی آزمون درک خواندن سال تحصیلی نیمه دوم ۹۶-۹۷ پرسشنامه به ده نفر از متخصصین در رشته آموزش زبان فارسی به غیر فارسی زبانان داده شد و از آنان درخواست شد در خصوص هر یک از ۳۰ سؤال آزمون اول درک خواندن و همچنین ۳۵ سؤال آزمون دوم، به صورت تعیین هدف و شاخص اصلی مربوط به هر سؤال و تعیین میزان ضروری بودن این شاخص از عدد ۱ تا ۳ (عدد ۱ به معنای «ضروری»، عدد ۲ به معنای «مفید ولی ضروری نیست» و عدد ۳ به معنای «ضروری نیست») را انتخاب کنند. پاسخ‌ها براساس فرمول CVR محاسبه و با جدول لاوشه انطباق داده شد. پس از تعیین و محاسبه CVR برای ارزیابی کمی روایی محتوایی و جهت اطمینان از این که مهم‌ترین و صحیح‌ترین محتوا (ضرورت سؤال انتخاب شده است) انتخاب شده، از نسبت روایی محتوایی و برای اطمینان از

این که سؤالات به بهترین نحو جهت اندازه‌گیری محتوا طراحی شده، از شاخص روایی محتوایی استفاده شد. بررسی CVI بر اساس شاخص روایی محتوایی باسل^۱ و والتز^۲ (۱۹۸۳) صورت گرفت. بدین منظور مجدداً از آنان خواسته شد که در مورد هر یک از سؤال‌های آزمون سه معیار مرتبط، واضح و ساده بودن بر اساس طیف لیکرتی چهار قسمتی اظهار نظر نمایند. به وسیله مجموع امتیازات موافق برای هر آیت‌م که رتبه ۳ و ۴ کسب کردند بر تعداد کل رای دهندگان CVI محاسبه شد.

برای بررسی روایی ظاهری از روش کمی استفاده شد. برای تعیین کمی روایی ظاهری نمره تأثیر هر سؤال محاسبه شده است. ابتدا برای هر یک از ۶۵ سؤال آزمون طیف لیکرتی ۵ قسمتی در نظر گرفته شده، امتیاز (کاملاً مهم ۵)، (مهم است ۴)، (به طور متوسط مهم است ۳)، (اندکی مهم است ۲) و (اصلاً مهم نیست ۱) سپس پرسشنامه جهت تعیین روایی ظاهری در اختیار ۲۴ زبان آموز سطح تکمیلی از چند کشور، آقا و خانم و در رده سنی مختلف قرار داده شد.

۵- ارائه و واکاوی آزمون‌ها

۵-۱ واکاوی روایی ظاهری آزمون یک

طبق داده‌های جدول (۱) در آزمون ۱ از تعداد ۳۰ سؤال، روایی ظاهری ۲۰ سؤال این آزمون از عدد ۱ بیشتر است و روایی ظاهری ۱۰ سؤال از عدد ۱ کمتر است. نمره تأثیر دو سؤال ۳ و ۲۲ عدد یکسان ۱/۲۲ و دو سؤال ۸ و ۱۹ نمره تأثیر یکسان ۱/۵۸ دارند و سؤالات ۱۶ و ۲۱، دو سؤال ۷ و ۲۵ و سؤال ۲۷ و ۳۰ دو به دو نمره تأثیر یکسانی دارند. بیشترین نمره تأثیر ۱/۹۸ مربوط به سؤال ۲۳ و کمترین نمره تأثیر عدد ۰/۲۹ مربوط به سؤال ۱۴ است؛ بنابراین سؤال ۲۳ بیشترین روایی ظاهری و سؤال ۱۴ کمترین روایی را دارد. سؤال ۲۳ که دارای روایی بیشتری نسبت به بقیه سؤالات است، در طبقه‌بندی سؤالات از نوع سؤالات درک مطلب که شاخص حقایق و ایده‌ها در متن را نشان می‌دهد، است و سؤال ۱۴ که کمترین روایی ظاهری را دارد، از نوع سؤال درک مطلب و دارای شاخص غنای واژگانی است. برای نمونه در آزمون ۱ سؤال ۱۱ که از نوع سؤال کلوز تست است با توجه به داده‌ها نمره تأثیر این سؤال کمتر از ۱/۵ بوده و

¹ Waltz.

² Bausell

همچنین ضریب روایی محتوایی آن کمتر از ۰/۶۲؛ بنابراین این سؤال از لحاظ روایی ظاهری و روایی محتوایی قابل قبول در نظر گرفته نشده است. در این آزمون از بین ۳۰ سؤال ۵ سؤال دارای روایی ظاهری هستند و نمره تأثیر ۴ سؤال به روایی ظاهری نزدیک است.

*آلودگی هوا همچنین به اجسام، ساختمان‌ها و اموال مردم آسیب می‌رساند و رفاه و آسایش عمومی تعادل طبیعی محیط زیست را مختل.....

الف) می‌بخشد ب) می‌گردد ج) می‌سازد د) می‌رسد

(آزمون یک، پرسش ۱۱)

در مورد گویه واضح بودن CVI. ۲۳ سؤال صد در صد و از بین ۳۵ سؤال، پرسش ۱ که نمونه‌ای از سؤالات کلوز تست بود کمترین CVI ۰/۶ را داشت.

* با دانستن اجزا و مواد تشکیل‌دهنده نوشابه و معایب آن از این پس دیگر این نوشیدنی‌های پرطرفدار نمی‌روید.

الف) نازنجک ب) سراغ پ) هیاهو ت) ردیف

(آزمون یک، پرسش ۱۴)

جدول ۲: فراوانی نمره تأثیر آزمون ۱

شماره	نمره	روایی	شماره	نمره	روایی	شماره	نمره	روایی
۱	۰/۸۷	ندارد	۱۱	۱/۴۱	ندارد	۲۱	۱/۲۱	ندارد
۲	۰/۹۸	ندارد	۱۲	۰/۸۷	ندارد	۲۲	۱/۲۲	ندارد
۳	۱/۲۲	ندارد	۱۳	۱/۲۹	ندارد	۲۳	۱/۹۸	دارد
۴	۰/۸۹	ندارد	۱۴	۰/۲۹	ندارد	۲۴	۱/۳۷	ندارد
۵	۱/۱۴	ندارد	۱۵	۰/۴۷	ندارد	۲۵	۱/۱۶	ندارد
۶	۰/۶۸	ندارد	۱۶	۱/۲۱	ندارد	۲۶	۱/۳۵	ندارد
۷	۱/۱۶	ندارد	۱۷	۱/۳۸	ندارد	۲۷	۰/۹۵	ندارد
۸	۱/۵۸	دارد	۱۸	۱/۶۰	دارد	۲۸	۱/۵۴	دارد
۹	۱/۰۹	ندارد	۱۹	۱/۵۸	دارد	۲۹	۱/۴۶	ندارد
۱۰	۰/۴۱	ندارد	۲۰	۱/۳۵	ندارد	۳۰	۰/۹۵	ندارد

۵-۲ واکاوی روایی ظاهری آزمون دو

طبق داده‌های جدول ۲، نمره تأثیر ۱۶ سؤال آزمون ۲، از عدد ۱/۵ بیشتر است و ۷ سؤال این آزمون نمره تأثیرش کمتر از عدد ۱ است. روایی ظاهری ۵ سؤال از عدد ۱/۵ بیشتر است؛ بنابراین ۱۶ سؤال روایی ظاهری دارند. سؤال ۱ و ۱۹ با هم نمره تأثیر یکسانی دارند و همچنین سؤال ۱۲ و ۱۵ نیز نمره تأثیرشان یکی است. بیشترین روایی ظاهری با نمره تأثیر ۲/۶۸ مربوط به سؤال ۲۱ از نوع سؤالات درک مطلب است که شاخص غنای واژگانی را بررسی می‌کند و همچنین کمترین روایی مربوط به سؤال ۱۱ است که نمره تأثیر آن ۰/۵ است. سؤال ۲۸ از آزمون دوم که از نوع سؤال درک مطلب است دارای روایی ظاهری ۱/۷۸ و روایی محتوایی ۱ است. طبق داده‌های گروه متخصصان و فارسی‌آموزان صد درصد در روایی محتوایی و روایی ظاهری این سؤال اتفاق نظر دارند.

* از آمارهای سنی و جنسی جمعیت به چه منظور استفاده نمی‌شود؟

الف) بهبود وضعیت جوانان

ب) ارتقای سطح سلامت جامعه

پ) جلوگیری از پیری جمعیتی، تحقق برنامه‌های تأمین اجتماعی

ت) تأمین شغل برای حال و آینده

(پرسش ۱۱- آزمون دو)

جدول ۳: درصد فراوانی نمره تأثیر آزمون ۲

شماره سؤال	نمره تأثیر	روایی ظاهری	شماره سؤال	نمره تأثیر	روایی ظاهری
۱	۱/۴۸	ندارد	۱۹	۱/۴۸	ندارد
۲	۱/۱۲	ندارد	۲۰	۱/۶۲	دارد
۳	۱/۰۹	ندارد	۲۱	۲/۶۸	دارد
۴	۰/۸۹	ندارد	۲۲	۱/۲۲	ندارد
۵	۱/۵۶	دارد	۲۳	۰/۷۴	ندارد
۶	۰/۷۱	ندارد	۲۴	۱/۶۰	دارد
۷	۰/۶۲	ندارد	۲۵	۲/۲۳	دارد
۸	۲/۱۸	دارد	۲۶	۲/۳۱	دارد
۹	۰/۸۷	ندارد	۲۷	۲/۶۵	دارد

شماره سؤال	نمره تأثیر	روایی ظاهری	شماره سؤال	نمره تأثیر	روایی ظاهری
۱۰	۱/۰۳	ندارد	۲۸	۱/۷۸	دارد
۱۱	۰/۵۰	ندارد	۲۹	۰/۹۶	ندارد
۱۲	۱/۵۸	دارد	۳۰	۰/۸۰	ندارد
۱۳	۱/۴۸	ندارد	۳۱	۱/۹۸	دارد
۱۴	۱/۶۲	دارد	۳۲	۰/۸۲	ندارد
۱۵	۱/۵۸	دارد	۳۳	۱/۵۳	دارد
۱۶	۱/۸۶	دارد	۳۴	۱/۰۱	ندارد
۱۷	۱/۱۵	ندارد	۳۵	۰/۷۵	ندارد
۱۸	۱/۵۶	دارد			

۳-۵ واکاوی روایی محتوایی آزمون یک

برای بررسی محاسبه روایی محتوایی از دو شاخص CVI و CVR استفاده شده و داده‌های به دست آمده با جدول لاوشه مقایسه شده است. با توجه به این که تعداد ارزیابان ۱۰ نفر بودند، شاخص روایی باید از ۰/۶۲ بیشتر باشد. ضریب روایی محتوایی سؤالات ۱، ۲، ۳، ۱۰، ۱۱ و ۱۴ کمتر از ۰/۶۲ است. در مجموع از ۳۰ سؤال، ارزیابان در روایی محتوایی ۶ سؤال توافق نظر ندارند. ضریب روایی محتوایی ۱۸ پرسش یک است و صد درصد ارزیابان در روایی محتوایی این سؤال‌ها توافق نظر دارند. در ۶ سؤال بیست درصد ارزیابان مخالف روایی محتوایی آنها هستند که ضریب روایی محتوایی ۰/۸ دارند و در روایی سؤالات ۲ و ۱۴ ۴۰٪ ارزیابان توافق نظر دارند و ۴۰٪ ارزیابان در روایی محتوایی سؤال‌های ۱ و ۱۰ مخالفند و کمترین ضریب روایی محتوایی ۰/۲ مربوط به سؤال ۳ و ۱۱ است که شاخص غنای واژگانی را مورد بررسی قرار داده‌است.

جدول ۴: درصد فراوانی ضریب روایی محتوایی آزمون ۱

شماره	ضریب روایی	شماره	ضریب روایی	شماره	ضریب روایی
۱	۰/۶	۱۱	۰/۲	۲۱	۱
۲	۰/۴	۱۲	۱	۲۲	۱
۳	۰/۲	۱۳	۱	۲۳	۱
۴	۱	۱۴	۰/۴	۲۴	۱
۵	۰/۸	۱۵	۰/۸	۲۵	۱

شماره	ضریب روایی	شماره	ضریب روایی	شماره	ضریب روایی
۶	۱	۱۶	۱	۲۶	۰/۸
۷	۱	۱۷	۱	۲۷	۰/۸
۸	۱	۱۸	۰/۸	۲۸	۱
۹	۱	۱۹	۱	۲۹	۱
۱۰	۰/۶	۲۰	۱	۳۰	۰/۸

۵-۴ واکاوی روایی محتوایی آزمون دو

طبق داده‌های جدول و نمودار (۴) سؤالات ۱، ۵، ۱۲، ۱۳، ۱۸، ۲۰ و ۳۱ دارای ضریب روایی محتوایی کمتری از عدد ۰/۶۲ نسبت به جدول لاوشه هستند. از بین ۳۵ سؤال ۱۸ سؤال ضریب روایی محتوایی آنها ۱ است. صد درصد ارزیابان بر روایی محتوایی آنها اتفاق نظر داشتند و ۶ سؤال ضریب روایی محتوایی ۰/۶ دارند که ۴۰ درصد ارزیابان با روایی محتوایی این سؤالات مخالف بودند و ۱۰ سؤال نظر ۸۰ درصد ارزیابان را به خود جلب کرده است و کمترین ضریب روایی محتوایی ۰/۲ مربوط به سؤال ۱ که شاخص غنای واژگانی را در آزمون ارزیابی می‌کند. طبق نظر ارزیابان از تعداد ۳۵ سؤال ۲۸ سؤال دارای روایی محتوایی است.

جدول ۵: درصد فراوانی ضریب روایی محتوایی آزمون ۲

شماره سؤال	ضریب روایی	شماره سؤال	ضریب روایی
۱	۰/۲	۱۹	۱
۲	۰/۸	۲۰	۰/۶
۳	۰/۸	۲۱	۰/۸
۴	۰/۸	۲۲	۰/۸
۵	۰/۶	۲۳	۱
۶	۰/۸	۲۴	۱
۷	۰/۸	۲۵	۰/۸
۸	۱	۲۶	۱
۹	۱	۲۷	۱
۱۰	۱	۲۸	۱
۱۱	۱	۲۹	۱
۱۲	۰/۶	۳۰	۱
۱۳	۰/۶	۳۱	۰/۶

شماره سؤال	ضریب روایی	شماره سؤال	ضریب روایی
۱۴	۰/۸	۳۲	۱
۱۵	۰/۸	۳۳	۱
۱۶	۱	۳۴	۱
۱۷	۱	۳۵	۱
۱۸	۰/۶		

۵-۵ واکاوی روایی محتوایی آزمون یک براساس شاخص «مرتبط بودن»

برای محاسبه شاخص روایی محتوایی CVI طبق مدل باسل و والتز (۱۹۸۳) شاخص روایی محتوایی «مرتبط بودن» نباید کمتر از ۰/۷۹ باشد. طبق داده‌های (جدول ۵) سؤال ۱ کمترین عدد شاخص روایی محتوایی ۰/۷ را دارد. ارزیابان از بین ۳۰ سؤال در روایی سؤال شماره ۱ اتفاق نظر ندارند. ارزیابان در شاخص روایی محتوایی مرتبط بودن سؤال‌ها، ۱۸ سؤال صد در صد و ۸ سؤال ۹۰٪ و ۳ سؤال ۸۰٪ اتفاق نظر دارند. شاخص روایی محتوایی محاسبه شده در سؤال‌های ۳، ۵، ۶، ۱۱، ۱۲، ۲۱ و ۹۰٪ و سؤالات ۲، ۲۴ و ۳۰، ۸۰٪ مرتبط هستند. بین ۳۰ سؤال فقط سؤال ۱ نامرتبط شمرده شده است و بقیه سؤال‌ها همگی مرتبط هستند.

جدول ۶: درصد فراوانی شاخص روایی محتوایی (مرتبط بودن) آزمون ۱

شماره سؤال	مرتبط بودن	شماره سؤال	مرتبط بودن	شماره سؤال	مرتبط بودن
۱	۰/۷	۱۱	۰/۹	۲۱	۰/۹
۲	۰/۸	۱۲	۰/۹	۲۲	۱
۳	۰/۹	۱۳	۱	۲۳	۱
۴	۱	۱۴	۱	۲۴	۰/۸
۵	۰/۹	۱۵	۱	۲۵	۱
۶	۰/۶	۱۶	۱	۲۶	۱
۷	۱	۱۷	۰/۹	۲۷	۱
۸	۱	۱۸	۰/۹	۲۸	۱
۹	۱	۱۹	۱	۲۹	۱
۱۰	۱	۲۰	۱	۳۰	۰/۸

۵-۶ واکاوی روایی محتوایی آزمون یک بر اساس شاخص «واضح بودن»

طبق داده‌های (جدول ۶) که مربوط به شاخص روایی واضح بودن سؤالات آزمون ۱

است، سؤال ۱ با ۰/۳ و سؤال ۱۰ با ۰/۷ روایی کمتری نسبت به بقیه سؤال‌ها دارند. ۱۹ سؤال صد درصد روایی محتوایی شاخص واضح بودن را دارند و ۱۰٪ ارزیابان در شاخص روایی گویه واضح بودن ۷ سؤال اتفاق نظر ندارند و ۹۰٪ ارزیابان با واضح بودن سؤالات ۳، ۵، ۶، ۱۵، ۲۱، ۲۴ و ۳۰، موافق هستند؛ بنابراین از ۳۰ سؤال، ۲۸ سؤال دارای روایی محتوایی واضح بودن است.

جدول ۷: درصد فراوانی شاخص روایی محتوایی (واضح بودن) آزمون ۱

شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن
۱	۰/۳	۱۱	۱	۲۱	۰/۹
۲	۰/۸	۱۲	۱	۲۲	۱
۳	۰/۹	۱۳	۱	۲۳	۱
۴	۱	۱۴	۱	۲۴	۰/۹
۵	۰/۹	۱۵	۰/۹	۲۵	۱
۶	۰/۹	۱۶	۱	۲۶	۱
۷	۱	۱۷	۱	۲۷	۱
۸	۱	۱۸	۱	۲۸	۱
۹	۱	۱۹	۱	۲۹	۱
۱۰	۰/۷	۲۰	۱	۳۰	۰/۹

۵-۷ واکاوی روایی محتوایی آزمون یک بر اساس شاخص «ساده بودن»

طبق داده‌های جدول ۸ درصد فراوانی شاخص روایی محتوایی ساده بودن، سؤال ۱ عددی کمتر از ۰/۷۹ است و ارزیابان اتفاق نظر در مورد روایی محتوایی ساده بودن این سؤال ندارند. در ۲۱ سؤال، صد درصد اتفاق نظر دارند. ۸۰٪ ارزیابان در روایی محتوایی سؤال ۲ موافق هستند و سؤالات ۳، ۶، ۱۲، ۲۱، ۲۸، ۲۹ و ۳۰، ۹۰٪ روایی محتوایی دارند.

جدول ۸: درصد فراوانی شاخص روایی محتوایی ساده بودن آزمون ۱

شماره سؤال	ساده بودن	شماره سؤال	ساده بودن	شماره سؤال	ساده بودن
۱	۰/۶	۱۱	۱	۲۱	۰/۹
۲	۰/۸	۱۲	۰/۹	۲۲	۱
۳	۰/۹	۱۳	۱	۲۳	۱

شماره سؤال	ساده بودن	شماره سؤال	ساده بودن	شماره سؤال	ساده بودن
۴	۱	۱۴	۱	۲۴	۱
۵	۱	۱۵	۱	۲۵	۱
۶	۰/۹	۱۶	۱	۲۶	۱
۷	۱	۱۷	۱	۲۷	۱
۸	۱	۲۸	۱	۲۸	۰/۹
۹	۱	۱۹	۱	۲۹	۰/۹
۱۰	۱	۲۰	۱	۳۰	۰/۹

۵-۸ واکاوی روایی محتوایی آزمون دو بر اساس شاخص «مرتبط بودن»

طبق درصد فراوانی‌های به دست آمده در (جدول ۸) بین سؤالات فقط سؤال ۵ و ۱ شاخص روایی محتوایی مرتبط بودن ۰/۶ است و کمتر از ۰/۷۹ است. از تعداد ۳۵ سؤال ۱۵ سؤال صد درصد مرتبطند. هشتاد درصد ارزیابان در روایی محتوایی مرتبط بودن، ۷ سؤال موافق هستند و در روایی محتوایی ۱۱ سؤال فقط ده درصد ارزیابان مخالفند و در سؤال ۵ هفتاد درصد ارزیابان در شاخص روایی مرتبط بودن سؤالات موافقتند. بنابراین ۳۳ سؤال دارای شاخص روایی محتوایی مرتبط بودن هستند.

جدول ۹: درصد فراوانی شاخص روایی محتوایی مرتبط بودن آزمون ۲

شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن
۱	۰/۶	۸	۰/۸	۱۵	۰/۸	۲۲	۰/۹	۲۹	۱
۲	۰/۹	۹	۰/۹	۱۶	۱	۲۳	۱	۳۰	۱
۳	۱	۱۰	۰/۹	۱۷	۱	۲۴	۰/۹	۳۱	۰/۹
۴	۱	۱۱	۰/۹	۱۸	۱	۲۵	۱	۳۲	۱
۵	۰/۷	۱۲	۰/۸	۱۹	۰/۹	۲۶	۱	۳۳	۱
۶	۰/۸	۱۳	۰/۹	۲۰	۰/۹	۲۷	۱	۳۴	۱
۷	۰/۸	۱۴	۰/۸	۲۱	۰/۸	۲۸	۰/۹	۳۵	۱

۵-۹ واکاوی روایی محتوایی آزمون دو بر اساس شاخص «واضح بودن»

طبق داده‌های جدول ۹ همه سؤالات گویه واضح بودن را دارند جز سؤال ۹ که عددی کمتر از ۰/۷۹ است. ۲۳ سؤال صد درصد واضح هستند و ارزیابان در روایی محتوایی ۶ سؤال

ده درصد اختلاف نظر دارند و ۸۰٪ در روایی محتوایی سؤالات ۷، ۸، ۱۲، ۱۴ و ۳۱ موافق هستند. ۳۴ سؤال از تعداد ۳۵ سؤال دارای روایی محتوایی واضح بودن هستند.

جدول ۱۰: درصد فراوانی شاخص روایی محتوایی واضح بودن آزمون ۲

شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن	شماره سؤال	واضح بودن
۱	۱	۲۲	۱	۱۵	۰/۹	۸	۰/۸	۲۹	۱
۲	۱	۲۳	۱	۱۶	۱	۹	۰/۷	۳۰	۱
۳	۱	۲۴	۱	۱۷	۱	۱۰	۰/۹	۳۱	۰/۸
۴	۱	۲۵	۱	۱۸	۱	۱۱	۱	۳۲	۱
۵	۰/۹	۲۶	۱	۱۹	۰/۸	۱۲	۰/۸	۳۳	۱
۶	۰/۹	۲۷	۱	۲۰	۰/۹	۱۳	۰/۹	۳۴	۱
۷	۰/۸	۲۸	۱	۲۱	۰/۸	۱۴	۰/۸	۳۵	۱

۵-۱۰ واکاوی روایی محتوایی آزمون دو بر اساس شاخص «ساده بودن» طبق داده‌های (جدول ۱۰) که شاخص روایی ساده بودن سؤالات آزمون ۲ را نشان می‌دهد شاخص روایی محتوایی (ساده بودن) سؤال ۹ از بین تعداد ۳۵ سؤال کمتر از ۰/۷۹ است. ۲۲ سؤال صد درصد شاخص روایی محتوایی را دارند و ارزیابان در روایی محتوایی (ساده بودن) ۷ سؤال نود درصد و ۵ سؤال هشتاد درصد موافق هستند.

جدول ۱۱: درصد فراوانی شاخص روایی محتوایی (ساده بودن) آزمون ۲

شماره سؤال	ساده بودن	شماره سؤال	ساده بودن	شماره سؤال	ساده بودن	شماره سؤال	ساده بودن	شماره سؤال	ساده بودن
۱	۱	۲۲	۱	۱۵	۰/۹	۸	۰/۹	۲۹	۱
۲	۱	۲۳	۱	۱۶	۱	۹	۰/۷	۳۰	۱
۳	۱	۲۴	۱	۱۷	۱	۱۰	۰/۸	۳۱	۰/۸
۴	۱	۲۵	۱	۱۸	۱	۱۱	۱	۳۲	۰/۹
۵	۰/۹	۲۶	۱	۱۹	۰/۸	۱۲	۰/۸	۳۳	۱
۶	۰/۹	۲۷	۱	۲۰	۰/۸	۱۳	۰/۸	۳۴	۱
۷	۰/۹	۲۸	۱	۲۱	۰/۸	۱۴	۰/۸	۳۵	۱

۶ - مقایسه روایی ظاهری آزمون‌های یک و دو

در آزمون یک از مجموع ۳۰ پرسش ۵ پرسش نمره تأثیر قابل قبولی داشتند و این

پرسش‌ها از روایی ظاهری برخوردار بودند؛ به عبارت دیگر، ۱۶/۶ درصد سؤالات این آزمون از روایی ظاهری مطلوبی برخوردار بودند و می‌توان گفت که آزمون از لحاظ شاخص روایی ظاهری، ضعیف است، در حالی که از ۳۵ پرسش آزمون دو ۱۶ پرسش دارای روایی ظاهری بودند که به عبارتی ۴۵/۷ پرسش‌های این آزمون از روایی ظاهری مطلوبی برخوردار بودند.

۷- مقایسه روایی محتوایی آزمون‌های یک و دو

در آزمون یک با توجه به داده‌های جدول از مجموع ۳۰ پرسش ۲۴ پرسش دارای روایی محتوایی بودند؛ به عبارتی، ۸۰ درصد سؤالات این آزمون روایی محتوایی داشتند و در آزمون دو از مجموع ۳۵ پرسش ۲۴ پرسش روایی محتوایی داشتند که در این آزمون نیز ۸۰ درصد پرسش‌ها از روایی محتوایی برخوردار بودند. بنابراین هر دو آزمون درصد یکسانی از روایی محتوایی برخوردار بودند.

۸- بحث

هدف اصلی این پژوهش بررسی روایی آزمون درک خواندن فارسی آموزان غیرایرانی مرکز زبان دانشگاه بین‌المللی امام خمینی^(ره) است. با توجه به یافته‌های به دست آمده از بررسی روایی آزمون درک خواندن در مورد روایی ظاهری آزمون ۱، از بین ۳۰ سؤال ۵ سؤال دارای روایی ظاهری است و کمترین عدد روایی ۰/۲۹ است. تعدادی از سؤالات آزمون یک به خوبی این هدف را پوشش نداده است. در آزمون دو کمترین روایی ظاهری عدد ۰/۵ است، از تعداد کل ۳۵ سؤال ۱۶ سؤال دارای روایی ظاهری است.

به نظر می‌رسد که از نظر روایی بهترین خرده مهارت‌های خواندن مربوط به سؤالات درک مطلب از نوع شاخص ارجاعات محتوایی و ایده و حقایق‌ها در متن این آزمون است و روایی ظاهری در آزمون ۲ بیشتر مشاهده می‌شود. شایان ذکر است که پرسشنامه‌های مربوط به روایی ظاهری که در اختیار فارسی آموزان قرار گرفته شد، در مورد هرگونه پرسشنامه به آنها توضیح شفاهی داده شده و ممکن است چون این نظر سنجی بعد از اتمام آزمون از فارسی آموزان خواسته شده و نه در زمان برگزاری آزمون، از نظر فارسی آموزان سؤالات دارای روایی ظاهری کمتری باشند و آن‌ها ضرورتی در لازم بودن سؤالات ندانند و همچنین اکثریت

فارسی‌آموزانی که از آنها نظر سنجی شده، عرب زبان بودند که نسبت به ملیت‌های دیگر غنای واژگانی آنها به زبان فارسی نزدیکتر بود. ممکن است اگر از افراد بیشتری با ملیتهای مختلف نظر سنجی می‌شد یافته‌های دیگری به دست می‌آمد.

در بررسی شاخص روایی CVR در آزمون ۱ با مقایسه با جدول لاوشه (۱۹۷۵) باید توجه داشت که آنچه در جدول آمده است شاخص درجه توافقی میان داوران است نه مشخص‌کننده روایی محتوایی آزمون؛ از تعداد کل سؤالات، ۵ سؤال عددی پایین‌تر از ۰/۷۹ داشتند و در ۱۲ سؤال به صورت صد در صد توافقی میان ارزیابان وجود داشت و در آزمون دو تعداد ۶ سؤال CVR محاسبه شده آنها کمتر از ۰/۷۹ و از بین ۳۵ سؤال تعداد ۱۸ سؤال CVR آنها صد درصد است. شاخص CVI در سه سطح مرتبط بودن و واضح بودن و ساده بودن طبق مدل باسل و والتز (۱۹۸۳) محاسبه شد. گویه مرتبط بودن در آزمون ۱ از بین ۳۰ سؤال یک سؤال CVI کمتری دارد و ۱۸ سؤال صددر صد است و گویه واضح بودن این آزمون دو سؤال CVI کمتری دارند و ۲۰ سؤال از کل سؤال‌ها صد درصد ارزیابان در مرتبط بودن سؤال‌ها توافقی نظر داشتند و همچنین در گویه ساده بودن ۲۱ سؤال CVI صد درصد به دست آوردند و دو سؤال CVI آنها به حد نصاب نرسید. در بررسی شاخص CVI آزمون ۲ گویه مرتبط بودن، یک سؤال CVI کمتری داشت و از بین ۳۵ سؤال ۱۴ سؤال ارزیابان صددر صد توافقی نظر داشتند. همچنین در گویه ساده بودن این آزمون در ۲۲ سؤال، ارزیابان صد درصد توافقی نظر داشتند و یک سؤال از تعداد کل سؤال‌ها CVI کمتری داشت. طبق یافته‌های به دست آمده هر دو آزمون دارای روایی محتوایی هستند. می‌توان گفت در آزمون ۱ روایی ظاهری چندان مطلوب نبوده است و مقایسه روایی ظاهری آزمون‌ها در دو دوره سال تحصیلی ۹۷ آزمون دوره دوم روایی ظاهری بیشتری نسبت به آزمون دوره اول دارد و روایی ظاهری در طی دوره دوم روبه بهبود است و همچنین آزمون درک خواندن دو دوره دارای روایی محتوایی یکسان است و روایی محتوایی در سؤالات درک مطلب این دو آزمون نسبت به خرده مهارت‌های دیگر خواندن بیشتر دیده می‌شود. نتایج نشان می‌دهد که محتوای این آزمون در مورد خواندن و میزان درک مطلب اطلاعات لازم را داراست و می‌تواند در جهت میزان پیشرفت تحصیلی فارسی‌آموزان کاربرد داشته باشد.

نتیجه‌گیری

یافته‌های این پژوهش همسو با پژوهش جباری و خادمی در سال ۱۳۹۲ است. در هر دو پژوهش روایی خرده مهارت‌های خواندن بررسی شده است. در پژوهش جباری علاوه بر روایی، پایایی آزمون نیز بررسی شده است ولی در پژوهش حاضر فقط روایی محاسبه شده و جامعه آماری در پژوهش جباری دانش‌آموزان ابتدایی هستند و شاخص‌های سلیس خوانی و واژگان و ایده اصلی متن دارای روایی بالایی نسبت به خرده مهارت‌های دیگر است و در پژوهش حاضر جامعه آماری فارسی‌آموزان سطح تکمیلی در رده سنی ۲۰-۳۰ سال هستند و بالاترین روایی مربوط به خرده مهارت‌های ارجاعات محتوایی و وقایع است. نتایج هردو پژوهش نشان می‌دهد که محتوای این آزمون‌ها در مورد خواندن و میزان درک مطلب اطلاعات لازم را داراست و می‌تواند در جهت میزان پیشرفت تحصیلی کاربرد داشته باشد. برای بررسی روایی محتوایی در این پژوهش همانند پژوهش ازر و همکاران در سال ۲۰۱۳ از پرسشنامه که گویه‌های واضح بودن، مرتبط بودن و عدم وجود ابهامات در کلمات بود استفاده شده که در پژوهش ازر، شاخص روایی محتوایی پایین ولی این پژوهش از شاخص روایی محتوایی مطلوبی برخوردار است.

تعارض منافع: طبق گفته نویسنده‌گان، پژوهش حاضر فاقد هرگونه تعارض منافع است.

منابع و مأخذ

- از کیا، مصطفی. (۱۳۸۲). روش‌های کاربردی تحقیق. تهران: انتشارات کیهان.
- جباری، سوسن؛ خادمی، محسن. (۱۳۹۲). ساخت آزمون تشخیصی خواندن و درک مطلب جهت دانش‌آموزان دوره‌ی ابتدایی. پژوهش‌های برنامه‌درسی. دوره ۳. شماره ۲. صص: ۵۱-۳۳.
- جعفرپور، عبدالجواد. (۱۳۷۵). مقدمه‌ای بر آزمون‌شناسی زبان. شیراز: مرکز نشر دانشگاه شیراز.
- حاجی زاده، ابراهیم؛ اصغری، محمد. (۱۳۹۰). روش‌ها و تحلیل‌های آماری با نگاه به روش تحقیق در علوم زیستی و بهداشتی. چاپ دوم. تهران: انتشارات جهاد دانشگاهی.
- زندی، بهمن و همکاران. (۱۳۹۸). «طراحی و اعتبار بخشی آزمون استاندارد زبان فارسی (بر پایه چهار مهارت زبانی)». جستارهای زبانی. سال دهم. شماره ۶. صص: ۱۷۱-۱۳۹.
- سرمد، زهره. (۱۳۹۶). روش‌های تحقیق در علوم رفتاری. تهران: نشر آگه.
- سنگری، محمدرضا و همکاران. (۱۳۸۵). آموزش خواندن (روش‌های مطالعه و خلاصه‌ای از برگزاری آزمون پرلز. تهران: ابوعطا.
- سیف، علی‌اکبر. (۱۳۹۵). اندازه‌گیری سنجش ارزشیابی زبان. چاپ چهل و سوم. تهران: نشر دوران.
- فتوت، احمدرضا؛ عباس زادگان، محمدرضا. (۱۳۸۴). کاربرد پایایی و روایی در پژوهش. تهران: نشر دفتر پژوهش‌های فرهنگی.
- قنسولی هزاره، بهزاد. (۱۳۸۹). «طراحی و روایی سنجی آزمون بسندگی زبان فارسی». پژوهش زبان‌های خارجی. شماره ۵۷. صص: ۱۲۹-۱۱۵.
- کرمی، ابوالفضل. (۱۳۹۰). آشنایی با آزمون‌سازی و آزمون‌های روانی. چاپ نهم. تهران: مرکز نشر روانسنجی.
- مرادی، علیرضا و همکاران. (۱۳۹۵). «روان‌شناسی بررسی اعتبار و روایی عاملی آزمون خواندن و نارساخوانی». تازه‌های علوم شناختی. سال ۱۸. شماره ۹۶. صص: ۳۴-۲۲.
- ممقانی، هانیه. (۱۳۹۰). ارزیابی روایی و پایایی آزمون وزارت علوم تحقیقات و فناوری (MSRT) با هدف دستیابی به یک آزمون مهارت برای زبان فارسی. رساله کارشناسی ارشد. به راهنمایی رضامراد صحرائی. دانشکده ادبیات و علوم انسانی. دانشگاه علامه طباطبائی.
- Brown, H. Douglas. (2004). language eassesmen principles and classroom practices, library of Congress Cataloging –in-publicationdata.

- Chao, C. (2017) .On Reading Test and Its Validity. Published by Canadian Center of Science and Education .
- Chen, Ch. (2010). On Reading Test and Its Validity’, Published by Canadian Center of Science and Education.
- Dong, B. (2011). A Content Validity Study of TEM-8 Reading Comprehension (2008 -2010). Kristianstad University Sweden.
- Ipek, O. Shawn M, F. Ebed, S. Diana, G. (2013). Reliability and content validity of an English as Foreign Language (EFL) grade-level test for Turkish primary grade students’ , Published by Elsevier Ltd.S
- Kim, A. (2011). Reliability and Validity of a Test of Reading Achievement and Reading Psychological Processes: Focusing on Gyeonggi-do and Chungcheongnam-do, The Korean Academy of Speech-Language Pathology and Audiology.
- Lawsh, CH. (1975). A quantitative approach to content validity’ Personnel Psych.
- Manxia, D. (2008). Content Validity Study on reading Comprehension Test of NMET ‘CELEA Journal Krishnan41y
- Santos, S , Chaves-Sousa, S., Ribeiro, I., Viana, F. L., Vale, A. P.,, & Cadime, I. (2017). Validity evidence of the Test of Word Reading for Portuguese elementary students. *European Journal of Psychological Assessment*, 33(6), 460–466
- Siddiek, G.(2010). The Impact of Test Content Validity on Language Teaching and Learning’, Published by Canadian Center of Science and Education.20
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry*, 45, 2-40.
- Waltz, CF. Bausell ,RB.(1983). *Nursing Research: Design, Statistics, and Computer Analysis’* , Philadelphia: W.B. Saunders.