

Developing, Implementing, and Evaluating an Online Placement Test based on Faraz Farsi Educational Collection

Zahra Hajibagheri¹

PhD Candidate of Teaching Persian to Speakers of Non-Persian Students, Shahid Beheshti University of Tehran, Tehran, Iran

Mahinnaz Mirdehghan Farashah²

Associate Professor of Linguistics, Shahid Beheshti University of Tehran, Tehran, Iran

(Corresponding author)

ORCID: 0000-0002-1094-4318

Received: June 24, 2024

Accepted: September 10, 2024

Abstract

Placement test is a test with the purpose of providing information that determines the appropriate level of the learners in an educational course according to their abilities. Due to the rapid development of technology and its application in the field of education, testing has changed from pen and paper-based test to computer-based test. The purpose of this research was to design an online placement test that has the required validity to determine the appropriate level of language learners in educational centers and the web-application under design (for teaching Persian to non-Persian speakers). This test was prepared based on the content of twelve-volume educational collection *Faraz Farsi* (Mirdehghan Farashah et al., 2023) and for the three sub-skills of vocabulary, grammar, and interaction. "Digiform" was selected among the online test makers for the test implementation due to its superior features, including the ability to display the test taker's image and higher security. In order to determine the reliability, the test was administered to 178 non-Persian language learners. The analysis of the obtained data using Coder and Richardson's formula 20, showed that the reliability of the test was 0.93. In order to determine the content validity of the test, a 10-item questionnaire was prepared and given to 20 teachers (18 female and 2 male) of *Faraz Farsi*. The teachers held MA degree or were PhD students in Teaching Persian, and have taught at least for five years. The results showed the desired content validity of the test.

Keywords: Placement Test, Computer-based Test, Content Validity, Reliability, Persian Language Teaching to Non-Persian Speakers.

1. Email: Zahra.hajibagheri@gmail.com

2. Email: Mahinnaz.mirdehghanfarashah@gmail.com



Authors retain the copyright and full publishing rights. This is an open access article distributed under article distributed under [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).



تدوین و اعتبارسنجی آزمون تعیین سطح آنالین زبان فارسی براساس مجموعه آموزشی «فراز فارسی»

زهرا حاجی‌باقری، دانشجوی دکتری آژفا، دانشگاه شهید بهشتی، تهران، ایران^۱

مهین‌ناز میردهقان‌فرشاه، دانشیار گروه زبان‌شناسی دانشگاه شهید بهشتی، تهران، ایران (نویسنده مسئول)^۲

ORCID: 0000-0002-1094-4318

صص ۱۲۹-۱۵۹

چکیده

آزمون تعیین سطح، آزمونی است با هدف فراهم کردن اطلاعات که با توجه به توانایی‌های زبان‌آموزان، سطح مناسب آن‌ها را در یک دوره آموزشی مشخص می‌کند. با توجه به پیشرفت سریع فناوری و کاربرد آن در حوزه آموزش، رویکرد آزمون‌سازی از آزمون مبتنی بر قلم و کاغذ به آزمون مبتنی بر رایانه تغییر یافته است. هدف پژوهش حاضر طراحی آزمون تعیین سطح آنالینی است که از اعتبار لازم برای تعیین سطح مناسب فارسی‌آموزان در مراکز آموزشی و همچنین وب-اپلیکیشن در دست طراحی (برای آموزش زبان فارسی به غیرفارسی‌زبانان) برخوردار باشد. این آزمون براساس محتوای مجموعه دوازده‌جلدی آموزشی «فراز فارسی» (میردهقان‌فرشاه و همکاران، ۱۴۰۲) و برای سه زیرمهارت واژه، دستور و کارکرد تهیه و به‌صورت آزمایشی اجرا شده است. برای اجرای آزمون، از میان محیط‌های آزمون‌ساز آنالین موردبررسی، «دیجی‌فرم» با توجه به ویژگی‌های برتر ازجمله امکان نمایش تصویر آزمون‌دهنده و امنیت بالاتر انتخاب شد. به‌منظور تعیین پایایی، لینک شرکت در آزمون از طریق یک گروه تلگرامی در اختیار ۱۷۸ نفر از فارسی‌آموزان غیرفارسی‌زبان مرکز آژفای دانشگاه شهید بهشتی قرار گرفت. تحلیل داده‌های حاصل، با استفاده از فرمول ۲۰ کودر و ریچاردسون پایایی ۰/۹۳ آزمون را نشان داد. در تعیین روایی محتوایی آزمون، پرسشنامه ۱۰ گویه‌ای تهیه شده در اختیار ۲۰ نفر از مدرسان (۱۸ زن و ۲ مرد) مجموعه آموزشی فراز قرار داده شد. همه مدرسان دارای مدرک کارشناسی ارشد یا دانشجوی دکتری آژفا و با سابقه تدریس بیش از پنج سال بودند. نتایج این پرسشنامه نمایانگر روایی محتوایی مطلوب آزمون می‌باشد.

کلیدواژه‌ها: آزمون تعیین سطح، آزمون مبتنی بر رایانه، روایی محتوایی، پایایی، آموزش زبان فارسی به غیرفارسی‌زبانان.

تاریخ دریافت: ۱۴۰۳/۰۴/تیر - تاریخ پذیرش: ۱۴۰۳/۲۰/شهریور

1. Zahra.hajibagheri@gmail.com

2. Mahinnaz.mirdehghanfarashah@gmail.com

۱. مقدمه

آزمون‌ها به منظور تعیین میزان تحقق اهداف آموزشی، ارائه بازخورد و همچنین تعیین سطوح، مورد استفاده قرار می‌گیرند. آزمون و ارزیابی را می‌توان همانند پلی میان درونداد و برونداد^۱ آموزشی در نظر گرفت که نتایج حاصل از آن‌ها به مدرسان و سیاست‌گذاران حوزه آموزش در مدیریت و درک میزان فرایند یادگیری یاری می‌رساند و همچنین پیش‌بینی عملکرد فراگیران در آینده، نقش تعیین‌کننده‌ای را ایفا می‌کند. بسیاری از مراکز آموزش زبان سعی در طراحی آزمون تعیین سطح دارند تا بتوانند براساس آن سطوح تسلط زبانی زبان‌آموزان را تشخیص داده و آن‌ها را در سطح یادگیری درست قرار دهند (لمب^۲، ۲۰۱۷).

با استفاده از نتایج این آزمون‌ها، هم زبان‌آموزان هم مدرسان می‌توانند درباره دوره‌های مناسب تصمیم بگیرند. همان‌طور که دوره‌ای با سطح زبانی بالاتر، که متناسب با نتایج آزمون نباشد، با توانایی یادگیری زبان آموز مطابقت نمی‌یابد، دوره‌ای با سطح پایین‌تر نیز زبان‌آموز را دلسرد کرده و انگیزه‌های یادگیری زبانی او را کاهش می‌دهد. آزمون‌های تعیین سطح در انتخاب دوره‌های زبانی درست نقش اساسی دارند. براساس نظر هیتون^۳ (۱۹۸۸) و براون^۴ (۲۰۰۴)، آزمون تعیین سطح برای ارزیابی توانایی زبانی افراد در زبان مقصد استفاده می‌شود تا آن‌ها را به سطوح یادگیری درست هدایت کند. سنجش درست و مؤثر برای موفقیت در یادگیری زبان ضروری است. در دهه‌های اخیر، اثربخشی آزمون تعیین سطح در بین کارشناسان بسیار بحث‌برانگیز شده است. در نتیجه در زمان آزمون تعیین سطح باید دو ملاک اصلی را در نظر گرفت؛ روایی^۵ و پایایی^۶ آزمون. یک آزمون روا و پایا به منظور قرار دادن زبان‌آموزان در سطوح درست یادگیری همان‌قدر مهم است که چنانچه یک آزمون بدون روایی و پایایی درست باشد، می‌تواند آن‌ها را در یک سطح اشتباه قرار دهد و در نتیجه باعث تأثیر منفی در یادگیری آن‌ها و ایجاد دیدگاه منفی نسبت به یادگیری زبان شود (العداوی و همکاران^۷، ۲۰۱۶).

امروزه با توجه به رشد و پیشرفت سریع فناوری و کاربرد آن در حوزه آموزش، رویکرد آزمون‌سازی از آزمون مبتنی بر قلم و کاغذ به آزمون مبتنی بر رایانه تغییر یافته است. استفاده از رایانه به عنوان واسطه برای ارائه آزمون‌ها در مقابل

1. Input & output

2. Lamb

3. Heaton

4. Brown

5. Validity

6. Reliability

7. Al-Adawi et al.

آزمون مبتنی بر قلم و کاغذ، آزمون مبتنی بر رایانه نامیده می‌شود. روش‌های اندازه‌گیری دوستدار محیط‌زیست مانند آزمون‌های مبتنی بر رایانه به‌منظور کاهش مصرف کاغذ استفاده می‌شوند. این رویکرد، جمع‌آوری داده‌هایی که از ارزیابی افراد و کلاس به دست آمده است را آسان‌تر می‌کند و همچنین نیازهای زبان‌آموزان و مراکز را بهتر برآورده می‌کند. این آزمون‌ها می‌توانند از طریق اینترنت یا با استفاده از سیستم‌های کامپیوتری، لپ‌تاپ و تبلت، در یک مرکز به‌صورت آنلاین به‌آسانی اجرا شوند و جمع‌آوری پاسخ‌ها و نمره‌دهی خودکار را تسهیل کنند.

نتایج به دست آمده از مطالعات انجام شده و پژوهش‌هایی که در مرور پیشینه به آن‌ها اشاره خواهد شد، نشان می‌دهد که تا پیش از این در زبان فارسی آزمون تعیین سطح آنلاین طراحی نشده است^۱. با توجه به رشد سریع و چشمگیر فناوری و ورود آن به حیطه آموزش زبان، نیاز به تدوین و طراحی یک آزمون تعیین سطح آنلاین کاملاً مشهود بود.

هدف این پژوهش اعتبارسنجی آزمون تعیین سطح طراحی شده توسط پژوهشگران بر مبنای مجموعه دوازده‌جلدی آموزشی «فراز فارسی» (میردهقان فرامشاه و همکاران، ۱۴۰۲) است. این مجموعه در حال حاضر منبع اصلی تدریس در مراکز آفرای دانشگاه شهید بهشتی، دانشگاه رازی کرمانشاه، دانشگاه عالی دفاع ملی و دانشگاه بغداد است و با توجه به بازخوردهای بسیار مثبتی که از این مراکز دریافت شده است و اعلام نیاز آن‌ها به یک آزمون تعیین سطح، بر آن شدیم تا این آزمون را طراحی کنیم. این آزمون هم به‌منظور استفاده در مراکزی که از این کتاب به‌عنوان منبع آموزشی خود استفاده می‌کنند، هم به‌عنوان آزمون تعیین سطح در نرم‌افزاری که براساس این مجموعه آموزشی در دست تهیه است، به کار گرفته می‌شود. نرم‌افزار مذکور، بخشی از برنامه‌ای جامع در حوزه آفرای است که به‌منظور تدوین منابع تخصصی آموزش زبان فارسی به غیرفارسی‌زبانان برای «اهداف عمومی» (برای رده‌های سنی مختلف بزرگسال، نوجوان، کودک) و «اهداف خاص» (برای اهداف دانشگاهی، اهداف تجاری، زبان‌های مختلف از جمله چینی، هندی و ...) طراحی شده است^۲.

پژوهش حاضر در پی یافتن پاسخ این پرسش‌ها است:

۱. شایان ذکر است که به‌تازگی بانک آزمون‌های زبان فارسی با نام اختصاری (بازفا) برای دانشجویان بین‌المللی بهره‌برداری شده است. آزمون‌های آزمایشی سنجش تحصیلی دانشجویان غیرایرانی در آموزش زبان فارسی، بخشی از این بانک آزمون است که به‌صورت مرتب هر دو هفته یک بار برگزار می‌شود و قرار است اولین اجرای آزمایشی آن در تاریخ ۲ تیر برگزار شود. لینک ثبت‌نام در آزمون

<http://educatech.ir/bazfa/>

۲. به وبگاه برنامه به نشانی Persianprime.org مراجعه شود.

۱. از بین نرم افزارهای آزمون ساز آنالین کدام یک برای برگزاری آزمون تعیین سطح زبان فارسی مناسب تر است و چه ویژگی هایی دارد؟

۲. آزمون تعیین سطح طراحی شده از چه میزان روایی و پایایی برخوردار است؟

۲. پیشینه پژوهش

در این بخش بر سه دسته از پژوهش ها تمرکز داشته ایم: ۱. پژوهش هایی که در حوزه آزمون های مبتنی بر رایانه انجام شده است؛ ۲. پژوهش هایی که در حوزه آزمون های تعیین سطح و اعتبارسنجی روایی و پایایی آن ها صورت گرفته است؛ ۳. پژوهش هایی که آزمون های تعیین سطح آنالین را مورد بررسی قرار داده اند.

۲-۱. آزمون های مبتنی بر رایانه

کارادنیز^۱ (۲۰۰۹) به بررسی تأثیرات آزمون مبتنی بر کاغذ، وب و موبایل بر پیشرفت و موفقیت دانشجویان در آموزش های اینترنتی و همچنین یافتن میزان درک دانشجویان از نحوه ارائه آزمون پرداخت و نشان داد که بین سطح موفقیت دانشجویانی که آزمون مبتنی بر کاغذ، وب و موبایل را انجام دادند و دانشجویانی که فقط آزمون مبتنی بر کاغذ را انجام دادند، تفاوت معناداری وجود ندارد. هرچند مشخص شد که دانشجویان به دلیل سهولت استفاده و بازخورد جامع و فوری، دیدگاه مثبتی نسبت به آزمون مبتنی بر وب و موبایل دارند.

در پژوهشی نیکوپور و همکاران (۲۰۱۱) با در نظر گرفتن نگرش زبان آموزان ایرانی نسبت به رایانه و یادگیری زبان از طریق رایانه، به مقایسه عملکرد آن ها در آزمون کاغذی و آزمون مبتنی بر رایانه پرداخته اند. آزمون مورد نظر شامل بخش های شنیدن، خواندن و کاربرد زبان بود که در دو اجرای مختلف برگزار شد. نتایج نشان داد اگرچه زبان آموزان ایرانی نگرش مثبتی نسبت به آزمون های مبتنی بر رایانه و یادگیری دیجیتال زبان داشتند، اما در آزمون های کاغذی بهتر عمل کردند.

حسینی و همکاران (۲۰۱۴) دوروش آزمون دهی کاغذی^۲ و رایانه ای را مورد مقایسه قرار دادند. از آزمون شوندگان آزمون چهار گزینه ای مهارت خواندن گرفته شد که نتایج حاکی از اولویت آزمون رایانه ای در مقایسه با آزمون کاغذی بوده است. نتایج پژوهش، مزایای آزمون رایانه ای را این گونه نشان می دهد: سرعت در اجرا، نتایج عینی، سهولت در اجرا، دقت بیشتر، بی واسطه بودن و گزارش دهی و امتیازدهی فوری. نتایج نهایی پژوهش بیانگر این واقعیت است که

1. Karadeniz

2. Paper and Pencil Tests (PPT)

نگرش دانشجویان نسبت به آزمون رایانه‌ای مثبت است اما در آزمون کاغذی موفقیت بیشتری دارند.

حیدرزاده و همکاران (۱۳۹۸) دو نوع آزمون رایانه‌ای و کاغذی را مورد مقایسه قرار داده‌اند. در این پژوهش پنج دوره از آزمون زبان وزارت بهداشت، درمان و آموزش پزشکی^۱ که هم‌زمان در هر دو نوع کاغذی و مبتنی بر رایانه برگزار شده است، انتخاب و مورد بررسی قرار گرفت. مهم‌ترین تفاوت نوع کاغذی این آزمون با نوع رایانه‌ای آن در مهارت شنیدن است. نتایج کلی پژوهش نشان‌دهنده این واقعیت است که در کل، نمرات گروه آزمون رایانه‌ای بر نمرات گروه آزمون کاغذی برتری داشته است.

با توجه به مطالب بیان شده در این قسمت، می‌توان نتیجه گرفت که آزمون‌دهندگان ایرانی نسبت به آزمون‌های مبتنی بر رایانه نگرش مثبتی دارند هرچند در مواردی در آزمون‌های کاغذی دارای عملکرد بهتری هستند و نمرات بهتری در این آزمون‌ها به دست می‌آورند.

۲-۲. آزمون‌های تعیین سطح

هدف آزمون تعیین سطح ارزیابی توانایی‌های زبانی فعلی زبان‌آموزان به منظور تطبیق آن‌ها با دوره‌های زبانی مناسب برای سطح آن‌ها است. تعیین سطح دقیق اولیه یک عامل کلیدی در یادگیری موفقیت‌آمیز یک زبان خارجی است، بنابراین زبان‌آموزان باید در کلاسی که مناسب سطحشان است ثبت‌نام کنند.^۲ در اینجا به بررسی چند آزمون تعیین سطح که در ایران انجام شده است، می‌پردازیم.

صحرائی و رضاپور (۱۳۹۸) داشتن روایی و پایایی مناسب را از ویژگی‌های مهم یک آزمون موفق می‌دانند. ایشان در پژوهش خود به بررسی روایی و پایایی آزمون تعیین سطح زبان فارسی بنیاد سعدی پرداخته‌اند. این آزمون از سه بخش اصلی دانش زبانی، نگارش و صحبت کردن تشکیل شده است که بخش دانش زبانی شامل سه زیربخش دستور، واژه و خواندن است. با توجه به میزان پایایی قابل قبول مشخص شد که آزمون در بخش‌ها و زیر بخش‌های گوناگون و کلیت آزمون، پایایی مناسب و قابل قبولی داشت.

کامیابی‌گل و همکاران (۱۳۹۸) آزمون تعیین سطح مرکز آفای دانشگاه فردوسی مشهد را اعتبارسنجی و روایی‌سنجی کرده‌اند. برای به دست آوردن روایی آزمون از تحلیل گویه شامل درجه تشخیص و درجه سختی استفاده شده و همچنین در یک مصاحبه کتبی نظرات و پاسخ‌های مدرسان و فارسی‌آموزان درباره سطح آزمون، سختی آزمون

1. MHLE

2. Stanford Language Center School of Humanities and Sciences

و کافی بودن زمان آزمون بررسی شده است. روایی محتوایی آزمون توسط اساتید مرکز «خوب» ارزیابی شده است. پس از آسیب‌شناسی آزمون نتیجه را می‌توان تا حدودی متأثر از مدت‌زمان یادگیری زبان فارسی، وقفه افتادن بین دوره‌های آموزشی و کمبود زمان برای پاسخگویی به آزمون دانست.

استاجی و رستمی (۱۴۰۱) در پژوهشی به بررسی زوایای مختلف آزمون بسندگی طراحی شده‌ای با نام آسپا و میزان پایایی و روایی آن پرداخته‌اند. یافته‌های این آزمون نشان داد که بخش واژگان و دستور از سطح آسانی بیشتر و بخش نوشتن از سطح دشواری بالاتری در جامعه آماری موردبررسی برخوردار است. نتایج این بررسی نمایانگر پایایی و روایی محتوایی مناسب آزمون بود. در این آزمون مهارت‌های دستور و نوشتن بالاترین میزان و مهارت خواندن کمترین میزان روایی محتوایی را داشته‌اند.

۲-۳. آزمون‌های تعیین سطح آنالین

نقوی^۱ و همکاران (۲۰۲۳) در پژوهشی روایی و پایایی یک آزمون تعیین سطح آنالین در یک موسسه آموزشی در کشور عمان را بررسی کرده‌اند. روایی محتوایی از طریق تعادل درونی و بیرونی پرسش‌ها و کلید پاسخ‌ها بررسی شد درحالی‌که روایی هم‌زمان معیارهای مرتبط از طریق معیار نمره‌دهی^۲ آیلنس موردبررسی قرار گرفت. ثبات درونی و پایایی آیت‌های آزمون خواندن با استفاده از آلفای کرونباخ اندازه‌گیری شد درحالی‌که برای آزمون شنیدن آمار توصیفی محاسبه شد. برای ارزیابی روایی درونی آزمون‌های صحبت کردن و نوشتن، از تی-تست استفاده شد. تحلیل داده‌ها نشان داد که آزمون برای قرار دادن زبان‌آموزان در سطوح مختلف مؤثر بوده و تحلیل توصیفی برای بررسی روایی و پایایی نتایج مثبتی را برای آزمون نشان داد.

آخیدنور^۳ (۲۰۱۹) به منظور مشخص کردن سطح زبانی زبان‌آموزان ثبت‌نام شده در دوره‌های زبان انگلیسی، راهبردهای جدید برای آزمون تعیین سطح آنالین را بررسی کرده است. او در پژوهش خود عملکرد آزمون پایانی دانش‌آموزان را به منظور تعیین اثربخشی آزمون مورد مقایسه قرار داده است. این پژوهش از رویکرد سه‌گانه نظرسنجی، پرسشنامه و مصاحبه‌های نیمه‌ساختار یافته بهره برده است؛ بنابراین مطالعه درباره نقش‌های آزمون تعیین سطح آنالین، اهمیت فناوری را در تقویت عملکرد یادگیری به‌ویژه برای زبان‌آموزانی که در عصر دیجیتال زندگی می‌کنند آشکار

1. Naqvi

2. Benchmark

3. Akhidenor

می‌سازد.

پولیت^۱ (۲۰۰۹) آزمون تعیین سطح آنلاین آکسفورد را ابزاری می‌داند که برای سنجش توانایی شرکت‌کنندگان در آزمون طراحی شده است و عملکرد ارتباطی را در سطوح مختلف مهارت زبان انگلیسی با توجه به چارچوب مرجع مشترک اروپا^۲ بررسی می‌کند. او توضیح می‌دهد که چگونه این آزمون توانایی‌های زبان‌آموزان را اندازه‌گیری می‌کند و نتایج چگونه براساس چارچوب مرجع مشترک اروپا تفسیر می‌شوند. یک ابزار ارزیابی مؤثر باید این اطمینان را به وجود آورد که همه پرسش‌ها یکپارچه هستند، اطلاعات موثقی درباره سطح مهارت زبان‌آموزان می‌دهند و در هر بار استفاده، نتایج دقیق و قابل‌اعتماد به ما ارائه می‌کند. آزمون تعیین سطح آنلاین آکسفورد کاربرد زبان (واژه، دستور و کارکرد) و مهارت شنیداری را مورد ارزیابی قرار می‌دهد.

نتایج به دست آمده از مطالعات انجام شده و پژوهش‌هایی که در این بخش به آن‌ها اشاره شد، نشان می‌دهد که تا پیش از این در زبان فارسی آزمون تعیین سطح آنلاین طراحی نشده است. یکی از مهم‌ترین ویژگی‌های این آزمون تعیین سطح، طراحی آن براساس محتوای جامع‌ترین و به‌روزترین منبع آموزش زبان فارسی و همچنین اجرای آن در محیط آزمون‌ساز آنلاین است. همان‌گونه که پیش‌تر اشاره شد، چون یکی از مهم‌ترین اهداف طراحی آزمون مزبور استفاده از آن در اپلیکیشن آموزش زبان فارسی در دست تهیه است، آزمون به‌صورت آنلاین طراحی و تهیه شده است. با توجه به رشد سریع و چشمگیر فناوری و ورود آن به حیطه آموزش زبان، نیاز به تدوین و طراحی یک آزمون تعیین سطح آنلاین کاملاً مشهود بود. هدف ما طراحی آزمونی مشابه با آزمون تعیین سطح آنلاین آکسفورد برای زبان فارسی بود اما به دلیل عدم پشتیبانی آزمون‌سازهای آنلاین موجود از بارگذاری فایل‌های شنیداری (۲-۳)، تنها از قالب کاربرد زبانی^۳ این آزمون بهره گرفتیم. شایان ذکر است که آزمون طراحی شده (واژه، دستور و کارکرد) همسو با تقسیم‌بندی موجود در چارچوب مرجع فارسی مصوب وزارت علوم (میردهقان فراشاه و همکاران، ۱۳۹۵) است. این چارچوب نه تنها برای مؤلفان کتاب‌های درسی بلکه برای آزمونگران در امر ارزشیابی چارچوبی مرجع و جامع به شمار می‌رود و سطوح تسلط هر سطح را به‌خوبی بازگو می‌کند.

1. Politt

2. Common European Framework of Reference (CEFR)

3. Use of language

۳. مبنای نظری پژوهش

آزمون‌های مبتنی بر رایانه، آزمون‌هایی هستند که به کمک رایانه و نرم‌افزارهای مخصوص برای پاسخ به پرسش‌ها و ارزیابی دانش و مهارت‌های زبان‌آموزان استفاده می‌شوند. این نوع آزمون‌ها امکاناتی مانند زمان‌بندی خودکار، امکان انتخاب گزینه‌ها، تصاویر و صدا را برای ارزیابی بهتر فراهم می‌کنند (هاگ^۱، ۲۰۲۳). این آزمون‌ها انواع مختلفی دارند که در این بخش به توضیح آن‌ها می‌پردازیم.

۳-۱. انواع مدل‌های اجرای آزمون‌های مبتنی بر رایانه

مهم‌ترین تصمیم یک طراح آزمون مبتنی بر رایانه، انتخاب مدل اجرای آزمون است چراکه این مدل، آیتم‌ها و ترتیبی را که این آیتم‌ها ارائه می‌شوند، کنترل می‌کند. چهار مدل اجرای آزمون عبارتند از:

۱. مدل خطی یا ثابت^۲: ساده‌ترین نوع آزمون مبتنی بر رایانه که اساساً مدل اجرای آزمون‌های کاغذی رایج را تکرار می‌کند این مدل است. به هر زبان‌آموز مجموعه‌ی مشابهی از پرسش‌ها داده می‌شود که ممکن است فقط ترتیب آن‌ها متفاوت باشد. این مدل مانند آزمون‌های رایج تهیه و نمره‌دهی می‌شود و نمرات با جمع کردن تعداد پاسخ‌های درست یا از طریق نظریه‌ی پرسش - پاسخ^۳ محاسبه می‌شود.

۲. مدل تصادفی: در این مدل هر زبان‌آموز با مجموعه‌ای از پرسش‌ها مواجه می‌شود که از یک بانک سؤالات استخراج شده است. این پرسش‌ها براساس قواعد آماری از بانک سؤالات استخراج می‌شوند تا در آزمون‌های مختلف از یکسان بودن محتوا، سطح دشواری و پایایی آن‌ها اطمینان حاصل شود. اگرچه در این مدل نیز نمرات آزمون را می‌توان با مجموع تعداد پاسخ‌های درست محاسبه کرد، اما نظریه‌ی پرسش - پاسخ ترجیح داده می‌شود.

۳. مدل چندمرحله‌ای^۴: این مدل ویژه‌ی آزمون‌های انطباقی است و از اولین و ساده‌ترین آن‌ها به شمار می‌رود. در مرحله‌ی اول این مدل ۱۰ پرسش یا بیشتر به زبان‌آموز ارائه می‌شود. وقتی که زبان‌آموز به این پرسش‌ها پاسخ داد، نمره‌اش محاسبه می‌شود و بر آن اساس تصمیم‌گیری می‌شود. این تصمیم برای انتخاب میان آزمون‌های موجود در مرحله‌ی دوم است؛ یعنی براساس عملکرد زبان‌آموز در مرحله‌ی قبل، کدام یک برای او مناسب‌تر است. با دنبال کردن اصول استانداردهای آزمون انطباقی، به زبان‌آموزی که در مرحله‌ی اول بهتر عمل کرده است، در مرحله‌ی دوم آزمونی دشوارتر به زبان‌آموز

1. Tobias Haug
2. Linear or fixed form
3. Item response theory (IRT)
4. Multi-stage form

ضعیف آزمون آسان‌تری داده می‌شود. با کامل کردن مرحله دوم آزمون، یک نمره نهایی محاسبه می‌شود که عملکرد زبان‌آموز در هر دو مرحله را نشان می‌دهد.

۴. مدل آیتم- انطباقی^۱: این مدل نمره هر پرسش را محاسبه می‌کند و تصمیم می‌گیرد بعداً چه چیزی را ارائه دهد. در این مدل پرسش‌ها براساس سطح عملکرد یک زبان‌آموز در پاسخ‌دهی به پرسش‌هایی که قبلاً در آزمون به او داده شده است، از بانک سؤالات انتخاب می‌شود (اجیم^۲، ۲۰۱۸).

۳-۲. مزایای آزمون مبتنی بر رایانه

اساسی‌ترین مؤلفه‌ای که در آزمون‌ها باید به آن توجه ویژه‌ای کرد، امنیت آزمون است. یکی از چالش‌های بزرگ آزمون‌های مبتنی بر قلم و کاغذ، حفظ و نگهداری پرسش‌های آزمون است، اما از نظر کیفی یکی از محاسن آزمون‌های مبتنی بر رایانه، مصون ماندن پرسش‌های آزمون است. مسأله دیگر، محیط امن برگزاری آزمون‌ها است. ناسازگاری محیط آزمون مبتنی بر قلم و کاغذ، توانایی واقعی آزمون‌دهندگان را نمایان نمی‌سازد. در مقایسه، آزمون‌های مبتنی بر رایانه با تضمین کنترل محیط، نتایج قابل اعتماد و سودمندی را در بر دارند زیرا مراکز که برای آزمون مبتنی بر رایانه در نظر گرفته می‌شوند، به گونه‌ای طراحی می‌شوند که فضای ایمن و مناسبی را برای برگزاری آزمون ایجاد می‌کنند. اجرای آزمون مبتنی بر رایانه پیشرفته‌تر از اجرای آزمون مبتنی بر قلم و کاغذ است. این امر بر آزمون‌دهندگان و بر اعتبار برنامه آزمون، تأثیرات مثبتی دارد.

در آزمون‌های مبتنی بر رایانه نتایج و دستاوردها به صورت امن ذخیره و پشتیبان‌گیری می‌شوند و یافته‌ها می‌توانند همیشه در دسترس باشند، در این صورت امکان مفقود شدن برگه‌های آزمون وجود ندارد. براون (۲۰۰۴) مزایای آزمون مبتنی بر رایانه را در سه طبقه فهرست می‌کند:

۱. مزایای طراح آزمون، ۲. مزایای مجری آزمون، ۳. مزایای آزمون‌دهنده/ داوطلب آزمون

در طبقه مزایای طراح آزمون، می‌توان به نمره‌دهی صحیح‌تر و واقعی‌تر، هدفمندسازی سطح توانایی خاص، کاهش تقلب، تأثیر مثبت آزمون بر یادگیری^۳ و روایی تقویت‌شده اشاره کرد.

در طبقه مزایای مجری آزمون، می‌توان به جمع‌آوری، انتقال و ذخیره‌سازی داده‌ها، نمره‌دهی از راه دور، آزمون‌های

1. Item- adaptive form

2. Ejim

3. Washback

کوتاه‌تر با مدت زمان کمتر و انعطاف اجرایی اشاره کرد. گردآوری داده‌ها در آزمون‌های مبتنی بر رایانه نسبتاً آسان است. در این مرحله، گردآوری اطلاعات و داده‌های اجرایی، درباره آزمون‌دهندگان و نمراتشان و اطلاعات پژوهشی درباره تعداد دفعاتی که هر داوطلب می‌تواند به پرسش‌ها پاسخ دهد و همچنین تغییرات ایجاد شده در پاسخ‌ها به کار گرفته می‌شود. داده‌ها پس از گردآوری به مکان‌هایی برای ذخیره و نمره‌دهی انتقال داده می‌شوند. یکی دیگر از مزایای آزمون‌های مبتنی بر رایانه ظرفیت ذخیره‌سازی داده‌ها است. با توجه به ظرفیت بالای درایوها و محیط ذخیره‌سازی ابری، هیچ محدودیتی جهت ذخیره‌سازی داده‌های گردآوری شده وجود ندارد. گردآوری، انتقال و ظرفیت ذخیره‌سازی داده‌ها در آزمون‌های مبتنی بر رایانه نمره‌دهی مؤثر از راه دور را امکان‌پذیر می‌کند. گرچه نمره‌دهی آزمون‌های ادراکی با فرمت‌هایی همچون چندگزینه‌ای، صحیح-غلط، وصل‌کردنی و غیره راحت است، ولی هنگام ارتباط با نمونه‌های تولیدی زبان مانند نوشتن و یا مصاحبه‌های ضبط‌شده ضروری است داده‌ها برای نمره‌دهی توسط نیروی انسانی و یا نمره‌دهی خودکار به یک مکان مرکزی انتقال یابد. برگزاری آزمون‌های مبتنی بر رایانه اعم از فردی و گروهی، سازماندهی اجرای آزمون از نظر برنامه‌ریزی کلاس‌ها، استخدام و آموزش آزمونگران را آسان‌تر و منعطف‌تر می‌کند. همچنین این آزمون‌ها از نظر زمان و جغرافیا نسبت به آزمون‌های مبتنی بر قلم و کاغذ انعطاف بیشتری دارند. (همان،

(۲۰۰۴)

طبقه سوم مزایای آزمون‌های مبتنی بر رایانه، مربوط به مزایای آزمون‌دهنده است. اکثر آزمون‌دهندگان از رایانه و فرایند آزمون مبتنی بر رایانه لذت می‌برند و این آزمون‌ها از منظر آنان جذاب است. آزمون‌دهندگان می‌توانند آزمون را با سرعت خود تنظیم نمایند و در نتیجه آزمون برایشان کمتر اضطراب‌آور و ترسناک می‌شود. آزمون‌های مبتنی بر رایانه بازخوردهای سریع‌تری را در مقایسه با آزمون‌های مبتنی بر کاغذ و قلم ارائه و گاهی افزون‌بر نتایج آماری، توضیحاتی در ارتباط با عملکرد آزمون‌دهندگان نیز گزارش می‌دهند.

۳-۳. معایب / کاستی‌های آزمون مبتنی بر رایانه

براون (۲۰۰۴) مشکلات آزمون‌های مبتنی بر رایانه را در دو گروه قرار می‌دهد:

۱. مشکلات فیزیکی

- مشکلات فیزیکی شامل میزان دسترسی به تجهیزات کامپیوتری، کیفیت تجهیزات، محدودیت در انواع پرسش‌ها، نیاز به تعداد زیادی پرسش و هزینه‌های آموزش متخصص می‌شود.

- به دلیل عدم دسترسی همگان به اینترنت و تجهیزات کامپیوتری در سراسر جهان، این مساله می‌تواند عملکرد آزمون‌دهندگان را به روشی منفی تحت تأثیر قرار دهد. تفاوت در کیفیت تجهیزات نیز منجر به پدیدار شدن چالش‌هایی در آزمون می‌شود.
- یکی دیگر از اشکالات و معایب آزمون‌های مبتنی بر رایانه، عدم وجود تنوع در نوع پرسش‌ها است و همچنین به دلیل نمردهی از طریق سیستم، بیشتر متکی بر ارزیابی مهارت‌های ادراکی‌اند تا مهارت‌های تولیدی؛ بنابراین به‌منظور جلوگیری از به خطر افتادن آزمون خصوصاً در طول اجرای آزمایشی آزمون، نیاز به تولید پرسش‌های بیشتری است.
- هزینه‌های ناشی از آموزش متخصصان، راه‌اندازی و نگهداری آزمایشگاه‌های کامپیوتری و تجهیزات و نرم‌افزارها یکی دیگر از مشکلات فیزیکی آزمون‌های مبتنی بر رایانه است.

۲. مشکلات عملکردی

- ازجمله معایب عملکردی آزمون‌های مبتنی بر رایانه را می‌توان عدم آشنایی با رایانه و نگرانی‌های موجود، پایایی و روایی دانست.
- در نظر داشته باشید فرصت انتخاب بین آزمون مبتنی بر رایانه و آزمون مبتنی بر کاغذ و قلم به آزمون‌دهنده داده شود و اطمینان حاصل کنید نمرات حاصل از هرکدام از آزمون‌ها قابل مقایسه باشد.
- عدم آشنایی با رایانه در طول آزمون، فرد را دچار اضطراب می‌کند و در پی آن نمرات تحت تأثیر قرار می‌گیرد.
- از دیگر کاستی‌های آزمون‌های مبتنی بر رایانه می‌توان به پایایی نمرات آزمون اشاره کرد. پایایی می‌تواند تحت تأثیر عوامل محیطی و اجرایی، شیوه‌های نمردهی، مشکلات فرمت آزمون یا مشکلات شخصی آزمون‌دهندگان قرار بگیرد.
- به دلیل تغییر آموزش زبان از رویکرد ساختاری، موقعیتی و مهارت‌محور به رویکردهای ارتباطی، ارزیابی‌ها نیز باید متناسب با رویکردهای ارتباطی باشد. ولی متأسفانه آزمون‌های مبتنی بر رایانه هنوز نتوانسته است با این پیشرفت‌ها سازگار شود.

برگزاری آزمون‌های مبتنی بر رایانه در مقایسه با آزمون‌های مبتنی بر قلم و کاغذ کمتر طبیعی به نظر می‌رسد؛ زیرا در اکثر مواقع مشاهده بیش از یک پرسش بر صفحه رایانه سخت است و نگاه کلی به آزمون برای تعیین راهبرد ترجیحی برای پاسخ دادن امکان‌پذیر نیست. در برخی موارد، آزمون‌های مبتنی بر رایانه به دلیل نشستن طولانی مدت جلوی رایانه مانند خستگی چشم، سردرد، گردن‌درد و دست‌درد منجر به نارضایتی آزمون‌دهندگان شده است.

در این قسمت از پژوهش مفهوم آزمون‌های مبتنی بر رایانه، مزایا و معایب آن‌ها و انواع مدل‌های این گروه از

آزمون‌ها بررسی و با توجه به مزایا، معایب و مدل‌های موجود، مدل خطی یا ثابت برای اجرای آزمون تعیین سطح آنالین در نظر گرفته شد.

۴. روش‌شناسی پژوهش

پژوهش حاضر از نوع کاربردی است و برای انجام آن از روش میدانی استفاده شده است. به منظور تجزیه و تحلیل داده‌های پژوهش، شیوه کمی به کار رفته است. پس از بررسی آزمون‌های تعیین سطح داخلی و خارجی، آزمونی براساس مجموعه آموزشی «فراز فارسی» طراحی شد. داده‌های پژوهش از طریق برگزاری آزمون آنالین در قالب آزمون تعیین سطح از دانشجویان دانشگاه شهید بهشتی متقاضی دوره فارسی آموزی جمع‌آوری شد.

۴-۱. شرکت‌کنندگان پژوهش

جامعه آماری پژوهش شامل ۱۷۸ نفر (زن و مرد) از دانشجویان دانشگاه شهید بهشتی متقاضی دوره فارسی آموزی با بازه سنی ۱۸ تا ۳۰ سال و ۲۰ نفر (۲ مرد و ۱۸ زن) از مدرسانی بودند که این فارسی آموزان پس از تعیین سطح در کلاس آن‌ها قرار گرفتند. همه مدرسان دارای مدرک کارشناسی ارشد یا دانشجوی دکتری آفا و با سابقه تدریس بیش از پنج سال بودند. مدرسان در بخش پاسخ به پرسش‌نامه برای بررسی روایی محتوایی آزمون شرکت داشتند.

۴-۲. ابزارهای پژوهش

برای انجام پژوهش از ابزارهای مختلفی استفاده شده است:

۱. برنامه آزمون‌ساز آنالین که بتوان پرسش‌های آزمون را در آن قرار داد و به نتایج موردنظر دست یافت.
۲. آزمون تعیین سطح که براساس مجموعه «فراز فارسی» طراحی شده است.
۳. استفاده از نرم‌افزار SPSS نسخه ۲۶ برای محاسبه ضریب دشواری و روش کودر و ریچاردسون^۱ به منظور اندازه‌گیری پایایی آزمون.
۴. پرسش‌نامه‌ای که به منظور بررسی روایی محتوایی آزمون در اختیار مدرسان قرار گرفت.

۴-۲-۱. آزمون‌ساز

از آنجایی که متقاضیان در کشورهای خودشان (عراق، سوریه، یمن، پاکستان) حضور داشتند و از طریق یک موسسه رابط معرفی شده بودند، برای اجرای این پژوهش به آزمون‌ساز آنالینی نیاز بود که بتوان پرسش‌های آزمون را در آن قرار

1. Kuder & Richardson (KR20)

داده و به نتایج موردنظر دست یافت. بدین منظور ابتدا چند محیط آزمون‌ساز آنلاین را موردبررسی قرار دادیم. نتایج مقایسه و ویژگی‌های این محیط‌ها در جدول (۱) قابل مشاهده است.

جدول ۱. مقایسه آزمون‌سازهای آنلاین

فرم‌افزار	گوگل فرم	دیجی فرم	پرسال	
*	*	*	*	تصحیح خودکار
*	*	*	*	بازه زمانی برای آزمون
*	*	*	*	نمایش یا ارسال نتیجه آزمون
*	*	*	*	خروجی نمرات به صورت اکسل
*	*	*	*	تصادفی سازی گزینه‌ها
*	*	*	*	انتشار آزمون
*		*	*	ارزیابی و گزارش گیری
*	*	*	*	نمایش در تمامی دستگاه‌ها
*	*	*	*	قالب‌بندی و زیباسازی آزمون
*		*	*	دریافت وجه جهت ساخت آزمون
*			*	دریافت وجه جهت شرکت در آزمون
*		*		ضبط صدا حین آزمون
*	*	*	*	تصادفی سازی پرسش‌ها
*	*	*	*	طراحی پرسش‌های تصویری
	*			ثبت نام در گوگل و ثبت ایمیل برای آزمون‌ساز

فرم افزار	گوگل فرم	دیجی فرم	پرسال	
	*			ثبت نام در گوگل و ثبت ایمیل برای آزمون دهنده
*			*	امکان دریافت صوت
*		*	*	نمایش نمودارها و پاسخ ها به صورت لحظه ای
*		*	*	دریافت لوکیشن
*	*	*	*	جلوگیری از ورود مجدد آزمون دهنده به آزمون
*		*	*	امکان خصوصی سازی آزمون برای گروهی خاص
*		*	*	تنظیمات متناسب با زبان فارسی
*		*	*	اتصال به درگاه های پرداخت
		*		امکان نمایش تصویر آزمون دهنده در ابتدا و هنگام آزمون

بررسی این چهار محیط آزمون ساز نشان داد که پرسال، دیجی فرم و فرم افزار تقریباً از ویژگی های یکسانی برخوردار هستند و برای ساخت آزمون در هر سه محیط نیاز به پرداخت وجه است، اما هزینه دیجی فرم نسبت به دو محیط دیگر پایین تر است. از سوی دیگر، در محیط پرسال و فرم افزار علاوه بر آزمون گر، آزمون دهنده نیز باید وجهی را پرداخت کند. به همین دلیل، برای اجرای آزمون خود محیط دیجی فرم را انتخاب کردیم. برخلاف محیط های دیگر، دیجی فرم امکان نمایش تصویر آزمون دهنده در ابتدا و هنگام آزمون را فراهم می کند که این موضوع می تواند احتمال تقلب را کاهش دهد و در نتیجه امنیت آزمون بالا رود.

همان طور که اشاره شد در این پژوهش به دنبال اجرای آزمون تعیین سطح مبتنی بر رایانه در یک محیط آزمون ساز آنالاین بودیم تا مزایا، معایب و چالش های آزمون سازهای آنالاین را بررسی کنیم. پس از بررسی چند محیط آزمون ساز آنالاین، آزمون ساز دیجی فرم انتخاب شد. این محیط نتایج آزمون را در قالب یک فایل اکسل ارائه می کند که در آن برای هر آزمون دهنده نمره، تعداد پاسخ درست، نادرست و بی پاسخ را نشان می دهد؛ اما اطلاعاتی درباره اینکه چه تعداد از آزمون دهندگان به یک پرسش پاسخ درست داده اند یا هر گزینه چند بار توسط آزمون دهندگان انتخاب شده

است ارائه نمی‌کند و آزمون‌گر برای انجام محاسبات آماری باید در یک فایل اکسل جداگانه این آمار را وارد کند. دیجی‌فرم برای هر پرسش نموداری ارائه می‌کند که نشان می‌دهد هر یک از گزینه‌های پرسش‌ها چند بار توسط آزمون‌دهندگان انتخاب شده است. به عنوان مثال در پرسش یک، ۲۱ نفر گزینه اول، ۳۳ نفر گزینه دوم و ۸۸ نفر گزینه سوم را انتخاب کرده‌اند. هر یک از این تعداد در نمودار با رنگ متفاوتی نشان داده می‌شود.

با توجه به اهمیت امنیت در آزمون‌های آنلاین، این محیط آزمون‌ساز به این مهم بسیار خوب توجه کرده است. در این محیط، موارد تکرار تقلب به صورت سیستمی مشخص می‌شود که عبارتند از:

- کنارهم بودن آزمون‌دهنده‌ها: چنانچه آزمون‌دهنده‌ها از یک سیستم استفاده کنند یا فاصله آن‌ها کمتر از ۵۰ متر باشد، تقلب در نظر گرفته می‌شود.
- دو بار شرکت کردن در آزمون: چنانچه آزمون‌گر محدودیت هر فرد یک بار را، فعال کرده باشد، افرادی که دو بار شرکت کرده باشند مشخص می‌شوند.
- تصویر مخدوش: چنانچه تصویر کاربر حین آزمون مخدوش باشد نیز تقلب ثبت می‌شود.
- استفاده از سایر اپلیکیشن‌ها حین آزمون: حین آزمون اجازه استفاده از اپلیکیشن‌ها و پیام‌رسان‌ها در دیجی‌فرم وجود ندارد.

در بخش نتایج برای هر یک از آزمون‌دهندگان اطلاعاتی را ارائه می‌دهد که شامل نمره، تعداد پاسخ‌های درست، نادرست و بی‌پاسخ است. علاوه بر این، زمان شروع پاسخگویی به آزمون، زمان ثبت آن، مدت زمان آزمون و تعداد آی‌پی سیستم کاربر را نشان می‌دهد. چنانچه آزمون‌دهنده حین آزمون تقلب کرده باشد یا بیش از یک بار وارد محیط آزمون شده باشد، نشان داده می‌شود. اگر آزمون‌دهنده از اپلیکیشن‌های دیگر استفاده کرده باشد، تعداد دفعات آن مشخص می‌شود. به عنوان مثال، در این آزمون یکی از آزمون‌دهندگان حین آزمون از ۳۵ اپلیکیشن دیگر استفاده کرده بود.

در پایان، با توجه به ویژگی‌ها و نقاط قوت و ضعف محیط آزمون‌ساز آنلاین دیجی‌فرم، این محیط را برای تولید آزمون‌های مبتنی بر رایانه پیشنهاد می‌دهیم.

۲-۲-۴. آزمون تعیین سطح

مهم‌ترین و اصلی‌ترین ابزار این پژوهش، آزمون تعیین سطح است. آزمون‌های تعیین سطح (پایه‌گزینی) آزمون‌هایی هستند که هدف آن‌ها فراهم ساختن اطلاعاتی است که با توجه به توانایی‌های زبان‌آموزان، مناسب‌ترین سطح برای

آن‌ها را در یک دوره آموزشی مشخص می‌کند (فالچر و دیویدسون^۱، ۲۰۰۷). موفق‌ترین آزمون‌های تعیین سطح آزمون‌هایی هستند که برای یک موقعیت و هدف خاص طراحی شوند. آزمون تعیین سطح مجموعه آموزشی «فراز فارسی» به منظور تعیین سطح فارسی آموزان در بدو ورود به دوره‌های فارسی آموزی مرکز آذفای دانشگاه شهید بهشتی و همچنین آزمون تعیین سطح در نرم‌افزاری که براساس این مجموعه آموزشی در دست تهیه است، طراحی و به صورت آزمایشی اجرا شده است. مخاطبان این آزمون بزرگسالانی هستند که با اهداف دانشگاهی و/یا تجاری و با آگاهی محدود از فرهنگ و زبان فارسی از ملیت‌های مختلف به ایران می‌آیند. از آنجاکه نیاز به آزمونی که بتواند فارسی آموزان را به لحاظ سطح توانش زبانی در جایگاه درستی قرار دهد، ضروری است، بر آن شدیم آزمون تعیین سطحی را طراحی کنیم که پاسخگوی این نیاز باشد و نتایج آن بتواند برای تعیین سطح فارسی آموزان در سه سطح پایه، میانی و پیشرفته مورد استفاده قرار بگیرد. برای تهیه آزمون تعیین سطح مراحل زیر انجام شد:

گام اول: برنامه‌ریزی‌های کلی (بررسی مفهوم آزمون تعیین سطح، بررسی پیشینه آزمون‌های تعیین سطح و بررسی آزمون‌های تعیین سطح موجود)

گام دوم: نوشتن ویژگی‌های آزمون

گام سوم: تهیه و تدوین آزمون، اجرای نسخه آزمایشی پرسش‌ها و بررسی پرسش‌نامه‌های تکمیل شده توسط مدرسان

گام چهارم: اجرای آزمایشی آزمون و تحلیل آماری نتایج

گام پنجم: فراهم کردن شکل نهایی آزمون براساس نتایج تحلیل آزمون آزمایشی و سطح‌بندی پیشنهادی

۴-۲-۱. ویژگی‌های آزمون تعیین سطح مجموعه آموزشی «فراز فارسی»

ویژگی‌های این آزمون در جدول (۲) قابل مشاهده است.

جدول ۲. مشخصات هدف و بافت آزمون

اجزا	توضیحات
هدف آزمون	تعیین سطح فارسی‌آموزان برای حضور در کلاس‌های متناسب با سطحشان (پایه، میانی، پیشرفته)
سازه‌های مورد ارزیابی و تعریف آن‌ها	واژه، دستور، کارکرد
دامنه کاربرد زبان مقصد و انواع فعالیت‌های رایج/مهم	این آزمون به‌طور غیرمستقیم توانایی آزمون‌دهندگان را برای کاربرد ارتباطی زبان فارسی در سطوح مختلف مورد ارزیابی قرار می‌دهد.
مشخصات آزمون‌دهندگان	بزرگسال، دارای ملیت‌های مختلف، دارای مدرک تحصیلی دیپلم یا بالاتر، با اهداف دانشگاهی و/یا تجاری، آگاهی محدود از فرهنگ و زبان فارسی

۴-۲-۲. تهیه و تدوین آزمون

به‌منظور سنجش توانایی فارسی‌آموزان برای کاربرد ارتباطی زبان فارسی، این آزمون در سه بخش کارکرد، دستور و واژه مبتنی بر مجموعه فراز فارسی تهیه و تدوین شد. برای شناسایی پرسش‌هایی که از نظر محتوایی یا ساختاری قابل قبول نیستند، این پرسش‌ها را در اختیار ۱۰ گویشور فارسی‌زبان قرار دادیم. این گویشوران کسانی بودند که از نظر سن، تحصیلات و سابقه کلی شبیه به آزمون‌دهندگان بودند. هیوز (۲۰۱۲) معتقد است بهتر است این افراد متخصص نباشند، زیرا احتمالاً افراد متخصص مانند افراد عادی رفتار نمی‌کنند. از نظر وی پرسش‌هایی که برای گویشور بومی مشکل‌ساز هستند، به‌طور قطعی نیاز به تجدیدنظر یا جایگزینی دارند؛ بنابراین براساس نتایج به دست آمده از گویشوران بومی، پرسش‌هایی را اصلاح یا حذف کردیم. مشخصات ساختار کلی آزمون شامل اجزا، فرمت پرسش‌ها و تعداد هر بخش در جدول (۳) مشاهده می‌شود.

جدول ۳. مشخصات ساختار کلی آزمون

سازه	اجزا، فرمت، تعداد
کارکرد	توانایی: - پرسیدن ساعت، پرسیدن قیمت و حساب کردن، بیان علاقه و تفر، سفارش غذا، درخواست صحبت تلفنی یا موکول کردن تماس به وقت دیگر، بیان حالت بیماری و ابراز همدردی، درخواست مؤدبانه (فراز پایه ۱ و ۲). - بیان مشکل و رسیدگی به مشکل، بیان اجبار یا ممنوعیت، دادن پیشنهاد و قبول آن، واکنش به خبر خوب و بد، عذرخواهی یا قبول عذرخواهی، درخواست یا ارائه راهکار (فراز میانی ۱ و ۲). - واکنش نشان دادن، حدس زدن، توصیه کردن و هشدار دادن، موافقت یا مخالفت با کاری، عوض کردن بحث، بیان اطمینان یا احتمال، موافقت با اکراه و نارضایتی (فراز پیشرفته ۱ و ۲). ۲۰ پرسش سه گزینه‌ای شامل ۷ پرسش از سطح پایه، ۷ پرسش از سطح میانی، ۶ پرسش از سطح پیشرفته. - پرسش‌های این بخش در قالب گفت‌وگوهای کوتاه خودساخته بین دو نفر ارائه می‌شود. امید می‌رود که ماهیت گفتاری آزمون به صورت غیرمستقیم نشان‌دهنده توانایی‌های گفتاری زبان‌آموزان باشد.
دستور	نکات دستوری: - حال اخباری، ضمیرهای پیوسته، چندم/چندمین، قیده‌های تکرار، پس‌از/قبل‌از، صفت برتر و برترین، واژه‌های پرسشی، فعل امر، حروف اضافه، حال ناتمام، حال التزامی: باید/نباید، حروف ربط ساده، گذشته ساده (فراز پایه ۱ و ۲). - گذشته استمراری، اجبار در گذشته، گذشته نقلی، قیده‌های گذشته نقلی، گذشته دور، ترتیب زمانی، مجهول حال اخباری، هنوز، فعل‌های یک‌شخصه، فعل‌های غیرشخصی، صفت مبهم، ضمیر مبهم (فراز میانی ۱ و ۲). - حرف اضافه مرکب، عبارت‌های موصولی، مجهول افعال مرکب، حرف ربط مرکب، فعل سببی، جمله‌های شرطی ممکن، جمله‌های شرطی ناممکن، بیان آرزو (فراز پیشرفته ۱ و ۲). ۴۰ پرسش سه گزینه‌ای شامل ۱۵ پرسش از سطح پایه، ۱۳ پرسش از سطح میانی و ۱۲ پرسش از سطح پیشرفته. - پرسش‌ها در قالب جمله‌هایی با حداکثر طول ۱۵ واژه ارائه می‌شود.

سازه	اجزا، فرمت، تعداد
واژه	واژه‌های مربوط به: • اعضای خانواده، لباس‌ها، فصل‌ها، کارهای روزانه، وسایل نقلیه، کارهای خانه، وسایل خانه، شغل‌ها، مکان‌های مختلف محله، آب‌وهوا، فعل‌های حرکتی، داروها، ورزش‌ها، برگزاری جشن (فراز پایه ۱ و ۲). • آمادگی برای سفر، ویژگی‌های رفتاری، توصیف حالت‌ها، صفات منفی، حوادث اورژانسی، حوادث طبیعی، مزه‌ها، برنامه‌های آینده، رفتارهای مؤدبانه، عادت‌های خرج کردن، راهکارهایی برای یادگیری زبان، کارهای بانکی، روش‌های آماده‌سازی و پختن غذا (فراز میانی ۱ و ۲). • حوادث در خبرها، آیین‌ها و مراسم، بیماری و درمان، فضاهای شهری، حراجی، انواع سفر، مشکلات جهانی، هنرها، مکان‌های عمومی، گردش در اینترنت، اهالی موسیقی، گونه‌های فیلم (فراز پیشرفته ۱ و ۲). • ۴۰ پرسش سه‌گزینه‌ای شامل ۱۵ پرسش از سطح پایه، ۱۳ پرسش از سطح میانی و ۱۲ پرسش از سطح پیشرفته. • پرسش‌ها در قالب جمله‌هایی با حداکثر طول ۲۰ واژه ارائه می‌شود.

۳-۴. اجرای آزمایشی آزمون

پس از تهیه محتوای آزمون براساس ویژگی‌های گام‌های دوم و سوم (۴-۲-۲)، آزمون در محیط آزمون‌ساز آنلاین دیجی‌فرم بارگذاری شد و لینک آزمون^۱ در اختیار ۱۷۸ نفر (زن و مرد) از دانشجویان دانشگاه شهید بهشتی متقاضی دوره فارسی آموزی با میانگین سنی ۱۸ تا ۳۰ سال، قرار داده شد. آزمون شامل ۱۰۰ پرسش سه‌گزینه‌ای و مدت زمان پاسخگویی به آن ۷۵ دقیقه در نظر گرفته شده بود. مدل اجرای آزمون، مدل خطی یا ثابت است. همان‌طور که اشاره شد در این نوع آزمون، به هر زبان‌آموز مجموعه‌ی مشابهی از پرسش‌ها داده می‌شود که ممکن است فقط ترتیب آن‌ها متفاوت باشد. این مدل مانند آزمون‌های رایج تهیه و نمره‌دهی می‌شود و نمرات با جمع کردن تعداد پاسخ‌های درست محاسبه می‌شود.

در ابتدا فارسی‌آموزان باید نام و نام خانوادگی خود را وارد می‌کردند. براساس تنظیمات انجام شده، فارسی‌آموزان فقط یک بار امکان ورود به آزمون را داشتند و ورود دوباره، به‌منزلهٔ تقلب در آزمون بود. برای جلوگیری از تقلب و

1. <http://d.digi-form.ir/w251258ae>

بالا بردن امنیت آزمون، ترتیب نمایش پرسش‌ها و ترتیب نمایش گزینه‌ها تصادفی انتخاب شد. تاریخ و زمان شروع و پایان نیز مشخص شد تا آزمون‌دهندگان فقط در آن بازه زمانی امکان حضور در آزمون و پاسخگویی به پرسش‌ها را داشته باشند. آزمون در سال ۱۴۰۲ برگزار شد. در جدول (۴)، ساختار آزمون به تفکیک مهارت‌های فرعی مشخص شده است:

جدول ۴. ساختار آزمون تعیین سطح

نمره	تعداد پرسش	زیر مهارت‌ها
۴۰	۴۰	دستور
۴۰	۴۰	واژه
۲۰	۲۰	کارکرد
۱۰۰	مجموع	

۴-۴. تحلیل داده‌های پژوهش

«هدف از تحلیل پرسش‌های یک آزمون بررسی تک‌تک پرسش‌ها و تعیین میزان دقت و نارسایی آن‌هاست. در تحلیل پرسش‌های آزمون، نقاط ضعف و قوت یک آزمون و کیفیت پرسش‌های آن تعیین می‌شود؛ بنابراین لازم است که آزمون‌گر پس از اجرای هر آزمون، پرسش‌های آن را تحلیل کند و با استفاده از نتایج حاصل از آن به تجدیدنظر در آزمون و بهبود کیفیت پرسش‌ها برای استفاده‌های بعدی اقدام کند» (سیف، ۱۳۹۰، ۴۲۱). در فرایند تحلیل پرسش‌های این آزمون، میزان ضریب دشواری، همبستگی پرسش با کل آزمون و پایایی و روایی آن با استفاده از نرم‌افزار SPSS 26 سنجیده شد.

۴-۴-۱. محاسبه ضریب دشواری^۱

پرسش‌ها تأثیر بسیار مهمی در سختی آزمون دارند و ما می‌بایست میزان سختی آن‌ها را به دقت محاسبه نماییم (رودمعجنی و همکاران، ۱۳۹۸: ۴۲). بنا به تعریف، درصد کل آزمون‌دهندگان که به یک پرسش جواب درست

1. Item difficulty

می‌دهند، ضریب دشواری آن پرسش است که با حرف p نشان داده می‌شود و از طریق فرمول زیر محاسبه می‌شود:

$$P = 100(R/T)$$

تعداد افرادی که به پرسش جواب درست داده‌اند

تعداد کل آزمون‌شوندگان

هراندازه ضریب دشواری یک پرسش بزرگ‌تر (به ۱۰۰ نزدیک‌تر) باشد، آن پرسش آسان‌تر است و هراندازه که این ضریب کوچک‌تر (به صفر نزدیک‌تر) باشد، پرسش دشوارتر است. ضریب دشواری هر یک از پرسش‌ها در جدول (۵) قابل مشاهده است.

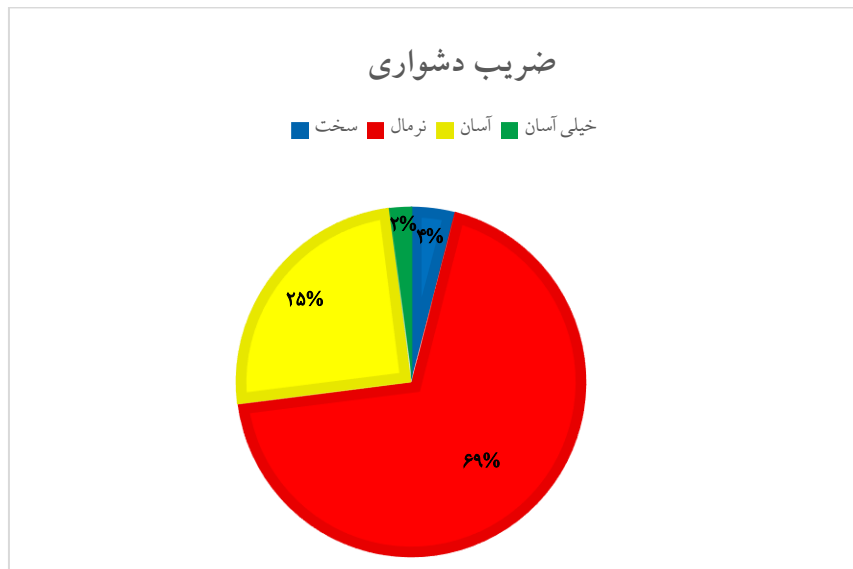
جدول ۵. ضریب دشواری پرسش‌های آزمون

ردیف	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری	ردیف	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری
۱	۸۸	۹۰	۰/۴۹	۵۱	۹۳	۸۵	۰/۵۲
۲	۹۴	۸۴	۰/۵۲	۵۲	۴۹	۱۲۹	۰/۲۷
۳	۱۰۱	۷۷	۰/۵۶	۵۳	۶۵	۱۱۳	۰/۳۶
۴	۷۹	۹۹	۰/۴۴	۵۴	۸۱	۹۷	۰/۴۵
۵	۹۹	۷۹	۰/۵۵	۵۵	۵۱	۱۲۷	۰/۲۸
۶	۷۱	۱۰۷	۰/۳۹	۵۶	۹۶	۸۲	۰/۵۳
۷	۶۱	۱۱۷	۰/۳۴	۵۷	۱۰۱	۷۷	۰/۵۶
۸	۱۰۰	۷۸	۰/۵۶	۵۸	۸۲	۹۶	۰/۴۶
۹	۸۴	۹۴	۰/۴۷	۵۹	۸۰	۹۸	۰/۴۴
۱۰	۵۵	۱۲۳	۰/۳۰	۶۰	۱۰۴	۷۴	۰/۵۸
۱۱	۶۹	۱۰۹	۰/۳۸	۶۱	۱۰۱	۷۷	۰/۵۶
۱۲	۵۱	۱۲۷	۰/۲۸	۶۲	۷۲	۱۰۶	۰/۴۰
۱۳	۹۷	۸۱	۰/۵۴	۶۳	۱۰۳	۷۵	۰/۵۷

رتبه	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری	رتبه	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری
۱۴	۸۴	۹۴	۰/۴۷	۶۴	۱۰۹	۶۹	۰/۶۱
۱۵	۷۸	۱۰۰	۰/۴۳	۶۵	۸۶	۹۲	۰/۴۸
۱۶	۶۶	۱۱۲	۰/۳۷	۶۶	۱۰۵	۷۳	۰/۵۸
۱۷	۵۸	۱۲۰	۰/۳۲	۶۷	۷۳	۱۰۵	۰/۴۱
۱۸	۷۰	۱۰۸	۰/۳۹	۶۸	۶۳	۱۱۵	۰/۳۵
۱۹	۷۵	۱۰۳	۰/۴۳	۶۹	۱۰۵	۷۳	۰/۵۸
۲۰	۵۸	۱۲۰	۰/۳۲	۷۰	۶۰	۱۱۸	۰/۳۳
۲۱	۴۹	۱۲۹	۰/۲۷	۷۱	۵۵	۱۲۳	۰/۳۰
۲۲	۷۴	۱۰۴	۰/۴۱	۷۲	۷۴	۱۰۴	۰/۴۱
۲۳	۸۳	۹۵	۰/۴۶	۷۳	۴۲	۱۳۶	۰/۲۳
۲۴	۹۱	۸۷	۰/۵۱	۷۴	۸۵	۹۳	۰/۴۷
۲۵	۳۳	۱۴۵	۰/۱۸	۷۵	۵۷	۱۲۱	۰/۳۲
۲۶	۵۴	۱۲۴	۰/۳۰	۷۶	۸۳	۹۵	۰/۴۶
۲۷	۹۰	۸۸	۰/۵۰	۷۷	۸۲	۹۶	۰/۴۶
۲۸	۸۵	۹۳	۰/۴۷	۷۸	۶۱	۱۱۷	۰/۳۴
۲۹	۵۹	۱۱۹	۰/۳۳	۷۹	۷۴	۱۰۴	۰/۴۱
۳۰	۹۷	۸۱	۰/۵۴	۸۰	۷۷	۱۰۱	۰/۴۳
۳۱	۵۲	۱۲۶	۰/۲۹	۸۱	۱۰۱	۷۷	۰/۵۶
۳۲	۸۷	۹۱	۰/۴۸	۸۲	۱۰۸	۷۰	۰/۶۰
۳۳	۴۹	۱۲۹	۰/۲۷	۸۳	۹۰	۸۸	۰/۵۰

رتبه	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری	رتبه	تعداد پاسخ درست	تعداد پاسخ نادرست	ضریب دشواری
۳۴	۸۹	۸۹	۰/۵۰	۸۴	۷۱	۱۰۷	۰/۳۹
۳۵	۷۵	۱۰۳	۰/۴۲	۸۵	۴۳	۱۳۵	۰/۲۴
۳۶	۶۱	۱۱۷	۰/۳۴	۸۶	۶۹	۱۰۹	۰/۳۸
۳۷	۶۱	۱۱۷	۰/۳۴	۸۷	۶۱	۱۱۷	۰/۳۴
۳۸	۹۰	۸۸	۰/۵۰	۸۸	۸۰	۹۸	۰/۴۴
۳۹	۳۲	۱۴۶	۰/۱۷	۸۹	۵۹	۱۱۹	۰/۳۳
۴۰	۴۷	۱۳۱	۰/۲۶	۹۰	۶۴	۱۱۴	۰/۳۵
۴۱	۹۱	۸۷	۰/۵۱	۹۱	۵۹	۱۱۹	۰/۳۳
۴۲	۱۰۶	۷۲	۰/۵۹	۹۲	۶۷	۱۱۱	۰/۳۷
۴۳	۶۵	۱۱۳	۰/۳۶	۹۳	۵۲	۱۲۶	۰/۲۹
۴۴	۱۲۶	۵۲	۰/۷۰	۹۴	۸۱	۹۷	۰/۴۵
۴۵	۱۱۹	۵۹	۰/۶۶	۹۵	۶۸	۱۱۰	۰/۳۸
۴۶	۹۲	۸۶	۰/۵۱	۹۶	۳۲	۱۴۶	۰/۱۷
۴۷	۱۱۰	۶۸	۰/۶۱	۹۷	۶۹	۱۰۹	۰/۳۸
۴۸	۱۱۶	۶۲	۰/۶۵	۹۸	۵۵	۱۲۳	۰/۳۰
۴۹	۱۰۹	۶۹	۰/۶۱	۹۹	۹۰	۸۸	۰/۵۰
۵۰	۱۰۳	۷۵	۰/۵۷	۱۰۰	۲۹	۱۴۹	۰/۱۶

به منظور نشان دادن توزیع فراوانی ضریب دشواری برای سطح کیفی پرسش‌های آزمون از نمودار (۱) استفاده کردیم.



نمودار ۱. توزیع فراوانی ضریب دشواری پرسش‌های آزمون

براساس این نمودار نتایج نشان می‌دهد که از مجموع ۱۰۰ پرسش، ۴ پرسش سخت، ۶۹ پرسش نرمال، ۲۵ پرسش آسان و ۲ پرسش خیلی آسان ارزیابی شده است. قابل توجه است که هیچ‌یک از پرسش‌ها از لحاظ ضریب دشواری، خیلی سخت نبوده و فراوانی این سطح صفر است. بیشترین درصد و فراوانی مربوط به سطح نرمال است.

۴-۲. محاسبه پایایی آزمون

پایایی^۱، یکی از مهم‌ترین ویژگی‌های آزمون‌ها به‌طور عام و آزمون‌های زبان به‌طور خاص است. پایایی به‌دقت، اعتمادپذیری، ثبات یا تکرارپذیری آزمون اشاره دارد ([کاپلان و ساکوزا](#)^۲). از آنجاکه امکان برگزاری و اجرای این آزمون یک بار وجود داشت، از روش کودر-ریچاردسون برای بررسی همسانی درونی آزمون و تعیین پایایی استفاده شده است. در این روش محاسبه پایایی، اگر نتیجه به دست آمده بالای ۷۰ درصد باشد، به این معناست که آزمون از پایایی خوبی برخوردار است. برای محاسبه پایایی به روش کودر-ریچاردسون به اندازه‌گیری جمع حاصل ضرب نسبت پاسخ‌های درست و نادرست ($\sum pq$)، میانگین، انحراف معیار و

واریانس نیاز است که این ارقام را می‌توان در جدول (۶) مشاهده کرد.

$$r_1 = \frac{n}{n-1} \left(1 - \frac{\sum pq}{S^2} \right)$$

1. Reliability

2. Kaplan & Saccuzzo

جدول ۶. شاخص‌های توصیفی نمرات

واریانس	انحراف معیار	میانگین	Σpq
۱۵۷۰	۳۹,۶۲	۴۳	۱۰۴

بررسی نتایج نشان داد که این آزمون از پایایی ۰/۹۳ برخوردار است. می‌توان بالا بودن ضریب پایایی را ناشی از بالا بودن تعداد پرسش‌های آزمون، یکسان بودن تقریبی ضریب دشواری پرسش‌ها و بالا بودن تعداد آزمون‌دهندگان دانست.

۳-۴-۴. محاسبه روایی آزمون

ابزارهایی که برای گردآوری داده‌ها به کار می‌روند، باید از روایی^۱ برخوردار باشند. روایی به این معناست که روش یا ابزار به‌کاررفته تا چه اندازه می‌تواند خصوصیت موردنظر در آزمون را درست اندازه‌گیری کند. روایی به هدفی که آزمون برای تحقق بخشیدن به آن درست شده است، اشاره می‌کند؛ به عبارت دیگر، آزمونی دارای روایی است که برای اندازه‌گیری آنچه موردنظر است، کافی و مناسب باشد. روایی بیشتر به خصوصیات آزمون بستگی دارد. رابطه بین روایی و پایایی از این قرار است که یک آزمون باید پایا باشد تا بتواند روایی داشته باشد.

پس برای اینکه یک آزمون روا باشد، نخست باید پایا باشد؛ یعنی پایایی پیش شرط روایی است و روایی آزمون کامل‌کننده پایایی است. در پژوهش حاضر به منظور اندازه‌گیری روایی آزمون از روایی محتوایی استفاده شده است.

سنجش روایی محتوایی^۲: باید گفت که درواقع، هیچ‌گونه روش آماری برای محاسبه روایی محتوایی یک آزمون وجود ندارد، بلکه برای تعیین روایی محتوایی یک آزمون، قضاوت متخصصان درباره اینکه پرسش‌های آزمون تا چه اندازه با هدف‌های برنامه مطابقت دارد، ملاک قرار می‌گیرد (بیرجندی و مصلا نژاد، ۲۰۱۰، ۱۶۶). از آنجاکه آزمون تهیه شده براساس محتوای کتاب‌های درسی فراز بوده است، به نظر می‌رسد آزمون از روایی محتوایی برخوردار باشد؛ اما نمی‌تواند روایی آزمون را تضمین کند. برای دستیابی به روایی آزمون طراحی شده باید به نظر کارشناسان این حوزه

1. Validity

2. Content validity

استناد کرد.

در این راستا پرسش نامه‌ای با ۱۰ گویه طراحی و همراه با ویژگی‌ها و مشخصات آزمون (جدول ۲) در اختیار ۲۰ نفر از مدرسان مجموعه فراز فارسی قرار داده شد تا براساس پاسخ آن‌ها بتوان روایی محتوایی آزمون را سنجید. معیار نظرسنجی براساس شاخص‌های خیلی خوب، خوب، متوسط و ضعیف طراحی شده است. پرسش نامه و نتایج آن در جدول (۷) آمده است.

جدول ۷. نظرات کارشناسان آموزش زبان فارسی به غیرفارسی‌زبانان درباره آزمون

گویه‌ها	ضعیف	متوسط	خوب	خیلی خوب
۱. آیا آزمون با سطوح توانمندی مشخص شده در چارچوب مرجع فارسی مطابقت دارد؟			۲۰٪	۸۰٪
۲. آیا مطالب آزمون به‌طور صحیح نمایانگر توانایی‌های زبانی موردنیاز برای هر سطح توانمندی (پایه، میانی، پیشرفته) هستند؟		۱۰٪	۲۰٪	۷۰٪
۳. آیا محتوای آزمون توسط متخصصان زبان یا مربیان آشنا با سطوح توانمندی مورد ارزیابی تأیید شده است؟				۱۰۰٪
۴. آیا موارد آزمون به‌گونه‌ای طراحی شده‌اند تا سطوح توانمندی را به‌طور مؤثر متمایز کنند و به‌طور دقیق فرد را در سطوح مناسب قرار دهند؟		۳۰٪	۵۰٪	۲۰٪
۵. آیا آزمون مواردی را شامل می‌شود که شبیه به استفاده و ارتباط واقعی زبانی در مواقع واقعی باشند؟			۳۰٪	۷۰٪
۶. آیا آزمون از نظر فرهنگی و بافتی مناسب آزمون‌دهندگان است؟		۳۰٪		۷۰٪
۷. آیا معیارهای امتیازدهی برای هر سطح توانمندی به‌طور واضح تعریف شده و با استانداردهای زبانی معتبر همخوانی دارند؟				۱۰۰٪
۸. آیا آزمون به پیش‌آزمایش گذاشته شده تا محتوای موارد آزمون و تصمیم‌های سطح‌بندی را ارزیابی کند؟	۸۰٪	۲۰٪		
۹. آیا یک فرایند مداوم برای ارزیابی و بازبینی آزمون براساس بازخورد از داوطلبان، مربیان و نتایج سطح‌بندی وجود دارد؟		۳۰٪	۵۰٪	۲۰٪
۱۰. آیا آزمون متناسب با محتوای آموزشی موردنظر طراحی شده است؟			۵۰٪	۵۰٪

همان‌طور که در جدول (۷) ملاحظه می‌شود، در پاسخ به پرسش اول ۸۰ درصد از کارشناسان معتقدند که این آزمون با سطوح توانمندی مشخص‌شده در چارچوب مرجع زبان فارسی مطابقت دارد.

از آنجایی که همراه با پرسش‌نامه، ویژگی‌ها و مشخصات آزمون (جدول ۲) نیز در اختیار مدرسان قرار گرفته بود، ۷۰ درصد از پاسخ‌دهندگان به گویه ۲، محتوای آزمون را بازتاب‌دهنده توانایی‌های موردنظر هر سطح دانستند. در پاسخ به پرسش سوم، همه مدرسان پس از بررسی آزمون اتفاق‌نظر داشتند که پرسش‌ها توسط متخصصان این حوزه طراحی شده است. نتایج گویه چهارم نشان داد ۳۰ درصد آن‌ها معتقدند که آزمون نمی‌تواند فارسی‌آموزان را در سطوح درست قرار دهد که می‌توان این موضوع را ناشی از طراحی آزمون براساس مهارت‌های فرعی دانست. با توجه به اینکه در پرسش‌های آزمون، ۲۰ پرسش به کارکرد (موضوعاتی براساس روابط و گفت‌وگوهای افراد در دنیای واقعی و در بسیاری از موارد منعکس‌کننده ویژگی‌های فرهنگی) اختصاص یافته بود، ۷۰ درصد مدرسان در پاسخ به گویه‌های پنج و شش، آزمون را اصیل^۱ دانستند.

همان‌طور که قبلاً نیز به آن اشاره شد آزمون قبل از برگزاری در اختیار ۱۰ نفر گویشور بومی قرار گفت تا آن‌ها براساس شم زبانی خود درستی پرسش‌ها و پاسخ‌ها را تأیید کنند؛ اما پیش‌آزمونی برگزار نشد و به همین دلیل، ۸۰ درصد مدرسان در گویه هشتم آزمون را از این منظر ضعیف دانستند. در گویه نهم ۵۰ درصد مدرسان در بخش پیشنهادهای پرسش‌نامه به این موضوع اشاره کرده‌اند که بهتر است آزمون از فرایند بازبینی و اصلاح برخوردار باشد. نتایج گویه دهم نشان داد چون مدرسان منتخب برای این پژوهش، کسانی بودند که مجموعه «فراز» را تدریس می‌کنند و تسلط کافی بر محتوای منبع آموزشی دارند، آن‌ها معتقد بودند آزمون منطبق بر محتوای آموزشی موردنظر است.

۵. بحث و نتیجه‌گیری

پژوهش حاضر به‌منظور تدوین، طراحی، اجرا و اعتبارسنجی یک آزمون تعیین سطح آنلاین به‌منظور استفاده از نتایج آن در وب-اپلیکیشنی که برای آموزش زبان فارسی به غیرفارسی‌زبانان در دست تهیه است و نیز مراکزی که از این منابع درسی در آموزش خود استفاده می‌نمایند، انجام شده است. آزمون تعیین سطح براساس محتوای مجموعه «فراز فارسی» برای سه زیرمهارت واژه، دستور و کارکرد (در مجموع ۱۰۰ پرسش) طراحی شد. به‌منظور اجرای این آزمون به‌صورت آنلاین پس از بررسی آزمون‌های مبتنی بر رایانه و انتخاب مدل خطی یا ثابت برای طراحی آزمون آنلاین،

1. Authentic

نقاط قوت و ضعف آزمون‌سازهای آنالین موجود را مورد مطالعه قرار دادیم و در پایان آزمون‌ساز آنالین دیجی‌فرم را برای اجرای آزمون تعیین سطح انتخاب کردیم که نمایانگر پاسخ پرسش اول پژوهش است.

در پاسخ به پرسش دوم تحقیق، پس از اجرای آزمون، نتایج به دست آمده را از نظر ضریب دشواری هر پرسش آزمون، همبستگی میان پرسش‌های آزمون و روایی محتوایی ارزیابی کردیم. محاسبه ضریب دشواری هر پرسش نشان داد که از مجموع ۱۰۰ پرسش، ۴ پرسش سخت، ۶۹ پرسش نرمال، ۲۵ پرسش آسان و ۲ پرسش خیلی آسان بودند. قابل توجه است که هیچ‌یک از پرسش‌ها از لحاظ ضریب دشواری، خیلی سخت نبوده و فراوانی این سطح صفر بوده است. بیشترین درصد و فراوانی نیز مربوط به سطح نرمال است. برای محاسبه پایایی آزمون از روش کودر-ریچاردسون استفاده کردیم که بررسی نتایج نشان داد آزمون از پایایی 0.93 و بسیار بالا برخوردار است. برای سنجش روایی محتوایی آزمون، از نتایج پرسش‌نامه‌ای که در اختیار مدرسان باتجربه قرار دادیم، استفاده کردیم. شواهد نشان داد که آزمون از روایی محتوایی مطلوبی برخوردار است.

با توجه به نتایج به دست آمده از این پژوهش می‌توان آزمون تعیین سطح آنالین را برای تعیین سطح فارسی‌آموزان در مراکزی که منبع آموزشی آن‌ها «فراز فارسی» است، پیشنهاد داد. همچنین این آزمون از اعتبار لازم برای استفاده در وب-اپلیکیشن در دست تهیه نیز برخوردار است.^۱

۱. نتایج میزان عملکرد آزمون مزبور در تعیین سطح دانشجویان دانشگاه علوم پزشکی شهید بهشتی، که متقاضی شرکت در دوره‌های فارسی‌آموزی مرکز آرفای دانشگاه شهید بهشتی بودند، در پژوهش آتی ارائه می‌شود.

کتابنامه

- استاجی، معصومه؛ اعظم رستمی. (۱۴۰۱). طراحی و بررسی ساختار، پایایی و روایی آزمون بسندگی سنجش زبان پارسی (آسپا). فصلنامه علمی زبان‌پژوهی دانشگاه الزهرا (س)، پاییز ۱۴۰۱، سال چهاردهم، شماره ۴۴، صص ۹-۳۴.
- حیدرزاده، آبتین؛ پنجه‌علی، سمانه؛ درخشان، پگاه؛ نوایی، مجید؛ تورانی، حسن. (۱۳۹۸). مقایسه تأثیر دو نوع شیوه ارزیابی کاغذی و مبتنی بر رایانه بر نتایج و عملکرد داوطلبان آزمون زبان وزارت بهداشت، درمان و آموزش پزشکی، فصلنامه علمی-پژوهشی طب و تزکیه، زمستان، دوره ۲۸، شماره ۴، صص ۲-۱۳.
- رودم‌عجنی، محسن؛ قنسولی، بهزاد؛ قبول، احسان. (۱۳۹۸). سنجش مهارت گوش کردن در محیط دانشگاه: بررسی موردی آزمون بسندگی فارسی در دانشگاه فردوسی مشهد، مجله زبانشناسی و گویش‌های خراسان، پاییز و زمستان، سال یازدهم، شماره ۲، صص ۲۵-۵۴.
- سمائی، زهرا؛ کامیابی گل، عطیه؛ مهرآوران، محمود. (۱۴۰۲). طراحی و اعتبارسنجی اولین آزمون تعیین سطح زبان فارسی براساس منابع آموزشی جامعه المصطفی العالمیه، پژوهشنامه آموزش زبان فارسی به غیرفارسی‌زبانان، بهار و تابستان، سال دوازدهم، شماره اول، صص ۴۵-۷۲.
- سیف، علی‌اکبر. (۱۳۹۰). اندازه‌گیری، سنجش و ارزشیابی آموزشی (ویرایش پنجم). تهران: دوران.
- صحرائی، رضامراد؛ رضاپور، غلامحسین. (۱۳۹۸). ارزیابی روایی و پایایی آزمون تعیین سطح زبان فارسی برای غیرفارسی‌زبانان بنیاد سعدی، فصلنامه اندازه‌گیری تربیتی، دانشگاه علامه طباطبائی، سال نهم، شماره ۷۳، صص ۹۹-۱۱۱.
- کامیابی گل، عطیه؛ سرافراز یزدی، سعیده؛ قنسولی، بهزاد. (۱۳۹۸). اعتبارسنجی آزمون تعیین سطح مرکز آفا در دانشگاه فردوسی مشهد، دو فصلنامه تخصصی مطالعات آموزش زبان فارسی. سال چهارم، شماره ۸، صص ۸۵-۱۱۲.
- میردهقان فراشاه، مهین‌ناز؛ وکیلی فرد، امیررضا؛ منتظری، زینب؛ باقری، فرشته. (۱۳۹۵). چارچوب مرجع آموزش زبان فارسی به غیرفارسی‌زبانان: دستور، واژه و کارکرد برای سطوح پایه، میانی و پیشرفته. تهران: مرکز همکاری‌های علمی بین‌المللی وزارت علوم، تحقیقات و فناوری، تهران: نشر خاموش.
- میردهقان فراشاه، مهین‌ناز؛ باقری، فرشته؛ زهرا حاجی‌باقری و همکاران (۱۴۰۲). مجموعه آموزشی دوازده‌جلدی فراز فارسی. تهران: انتشارات دانشگاه شهید بهشتی.

Akhidenor-Bamidele, Anthonia. (2019). The Roles of Online Placement Test in English Language Teaching. *University of Bahrain – English Language Center: The 2nd International Conference*.

Al-Adawi, S. et al. (2016). Investigating Content and Face Validity of English Language Placement Test Designed by Colleges of Applied Sciences. *English Language Teaching*, Vol. 9.

- Alkhadher, Othman. Neil Anderson & David Clarke. (1994). Computer- Based Testing: A Review of Research and Practice. *European work and organizational psychologist*, 4 (2), 169-187.
- Birjandi, P. P. Mosallanejad. (2010). *An Overview of Testing and Assessment*. Esfahan: Sepahan Publication.
- Brown, H.D. (2004). *Language assessment: Principles and classroom practices*. New York: Pearson Education.
- Brown, J. D. (1997). *Language Testing and Teaching*. 141-158.
- Ejim, Samson. (2018). *Overview of Computer- Based Tests*. Research Proposal.
- Fulcher, Glenn, Fred Davidson. (2007). *Language testing and assessment*. Routledge.
- Haug, Tobias. W. Mann, F. Holzknicht. (2023). The Use of Technology in Sign Language Testing: Results of a Pre-Pandemic Survey. *Sign Language Studies*. DOI:10.1353/sls.2023.0003
- Heaton, J.B. (1988). *Writing English Language Tests*. Harlow: Longman.
- Hosseini, Monirosadat. Mohammad Z.A Jafri. (2014). Comparability of Test Results of Computer based Tests (CBT) and Paper and Pencil Tests (PPT) among English Language Learners in Iran, *Procedia - Social and Behavioral Sciences*.
- Hughes, Arthur. (2012). *Testing for Language Teachers*. Cambridge University Press.
- Kaplan, Robert. Dennis Saccuzzo. (2017). *Psychological Testing: Principles, Applications, and Issues*. 9th Edition. Cengage Learning.
- Karadeniz, S. (2009). The impacts of paper, web and mobile based assessment on students' achievement and perceptions. *Scientific Research and Essay*, 4 (10), 984 – 991. Retrieved May 15, 2011 from <http://www.academicjournals.org/sre>.
- Lamb, M. (2017). The motivational dimension of language teaching. *Language Teaching*, 50 (3), 301-346.
- Naqvi, Samia. (2023). Establishing Reliability and Validity of an Online Placement
- Nikoopour, J. M Amini Farsani, M Nasiri. (2011). On the Relationship Between Critical Thinking and Language Learning Strategies Among Iranian EFL Learners. *Journal of Technology & Education*, 5 (3), 195-200.
- Pollitt, Alastair. (2009). The Oxford Online Placement Test: The Meaning of OOPT Scores. www.oxfordenglishtesting.com.
- Salimi, H. Rashidy, A., Salimi, A. H., & AminiFarsani, M. (2011). Digitized and non-Digitized Language Assessment: A Comparative Study of Iranian EFL Language Learners, *International Conference on Languages, Literature and Linguistics IPEDR*, vol. 26. Singapore: IACSIT Press.
- Test in an Omani Higher Education Institution, *Languages*, 8(1), 61.
<https://doi.org/10.3390/languages8010061> - 22 Feb 2023.