

## اعتبارسنجی آزمون تعیین سطح مرکز آرفا در دانشگاه فردوسی مشهد

سعیده سرافراز یزدی<sup>۱</sup>  
عطیه کامیابی گل<sup>۲</sup>  
بهزاد قنسولی<sup>۳</sup>

### چکیده

هدف از تحقیق حاضر اعتبارسنجی و روایی سنجی آزمون تعیین سطح مرکز آموزش زبان فارسی به غیر فارسی زبانان دانشگاه فردوسی مشهد است. روش تحقیق حاضر توصیفی-تحلیلی بوده و جمعیت آماری حاضر ۵۱ نفر از فارسی آموزان آن مرکز در سال ۹۸ است. برای رسیدن به هدف تحقیق و تعیین پایایی آزمون، از آلفای کرونباخ استفاده شده است که نتیجه ۰/۸۲ برای این آزمون تلقی می‌شود و نشان از ثبات درونی آزمون است. برای به دست آوردن روایی آزمون از تحلیل گویه شامل درجه تشخیص و درجه سختی استفاده می‌شود و همچنین در یک مصاحبه کتبی نظرات و پاسخ‌های مدرسان و فارسی آموزان در سطح آزمون، سختی آزمون و کافی بودن زمان آزمون مورد بررسی قرار گرفته است. روایی محتوایی آزمون توسط اساتید مرکز «خوب» ارزیابی شده است. نتیجه ارزیابی روایی نشان داد معادل ۳۸ درصد سوالات مطلوب و معادل ۶۲ درصد، تشخیص بین زبان آموز قوی از ضعیف قائل نشده است پس ضریب تشخیص در آزمون ضعیف ارزیابی شده است که پس از آسیب شناسی، نتیجه را می‌توان تا حدودی متاثر از مدت یادگیری زبان فارسی بعضی از آنها، به وقفه افتادن بین دوره‌های آموزشی و کم آوردن مدت زمان برای پاسخگویی به آزمون دانست.

### کلید واژه‌ها:

تحلیل سوال، آزمون تعیین سطح، ضریب دشواری، ضریب تشخیص، آموزش زبان فارسی به غیر فارسی زبانان.

۱. کارشناس ارشد آرفا، دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد

۲. استادیار گروه زبان و ادبیات فارسی دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد. نویسنده مسئول: (kamyabigol@um.ac.ir)

۳. استاد گروه انگلیسی دانشکده ادبیات و علوم انسانی دانشگاه فردوسی مشهد

تاریخ دریافت: ۹۹/۰۵/۰۸ تاریخ پذیرش: ۹۹/۱۱/۰۱

بدون تردید آزمون سازی و ارزشیابی دو بخش جدایی ناپذیر آموزش هستند. از منظر آموزشی، آزمون سازی در نقطه مرکزی آموزش هر زبان قرار دارد. آزمون سازی علاوه بر این که بر اهداف آموزشی زبان تاثیر می گذارد، بر دستیابی موفق به اهداف از پیش تعیین شده آموزشی نیز موثر است (دیویس، ۱۹۹۰).

آموزش و آزمون سازی رابطه‌ای تنگاتنگ و دو طرفه دارند؛ به طوری که چستین (۱۹۸۸) معتقد است که آموزش و یادگیری بدون در نظر گرفتن فاکتور ارزشیابی به نظر آرمان-گرایانه، غیرکاربردی و بی ثمر خواهد بود. آزمون سازی زبان مسئله پیچیده‌ای است؛ این نوع آزمون سازی از علوم روان شناسی، زبان شناسی و آموزش زبان بهره می گیرد (کلاین بریلی، ۱۹۸۱).

یکی از مهمترین وظایفی که مدرسین با آن مواجه هستند ارزیابی زبان آموزان است. این وظیفه مهم نیاز به مهارت‌های قابل توجهی دارد؛ زیرا این آزمون‌ها به عنوان یک شاخص کلیدی و حساس برای ارتقا و ارزشیابی زبان آموزان به کار می روند. مسلماً کار آزمون با نوشتن سوالات به پایان نمی رسد، بلکه پس از آن باید به دنبال گردآوری شواهدی برای اطمینان از کیفیت مناسب سوال بود. برای دستیابی به چنین هدفی، طراحی آزمون باید یک فرایند نظام مند سنجیده جهت جمع آوری و تحلیل اطلاعات صحیح باشد. هدف از چنین قضاوت‌هایی صرفاً شناسایی عوامل خطا در اندازه گیری است. بدین جهت پس از اجرای یک آزمون و تصحیح برگه‌های آزمون، تجزیه و تحلیل سوالات، روش مناسبی برای ارزشیابی نهایی در مورد کیفیت آزمون است و البته در آزمون‌های چند گزینه‌ای با توجه به عینی تر بودن فرایند تصحیح و نمره گذاری، بهتر می توان در مورد کیفیت آزمون قضاوت کرد. در تجزیه و تحلیل سوال‌های یک آزمون، اصولاً دو جنبه اندازه گیری که شامل کمی و کیفی است، استفاده می شود. جنبه‌های کمی شامل قدرت تمیز و درجه سختی سوال‌ها می شود و جنبه کیفی مواردی چون شکل و مراحل طراحی سوال‌ها و بررسی پاسخ‌های صحیح و غلط را در بر می گیرد.

تجزیه و تحلیل سوال‌های آزمون مزیت‌های زیادی در ارزیابی موثر فراگیران دارد. مقاله‌ای با عنوان «تجزیه و تحلیل سوالات چهارگزینه‌ای» که توسط کهو

نوشته شده است، نشان می‌دهد که بررسی علمی آزمون‌ها، متداولترین ابزار برای ارزیابی یادگیری و پیشرفت تحصیلی است (برایمن، ۲۰۱۲). پس متداولترین ابزار اندازه‌گیری که در بعضی از بافت‌های آموزشی به کار می‌رود، آزمون‌های چهارگزینه‌ای است. در دنیای امروزی ارزیابی را به عنوان فرایندی منظم برای توصیف و جهت‌بخشی به کیفیت یادگیری و پیشرفت تحصیلی می‌دانند (بازرگان، ۱۳۸۱). جهت ارزیابی مناسب و مطابق با استانداردهای علمی، لازم است طراحان آزمون با اهداف آموزشی آشنا باشند و با استفاده از بودجه‌بندی مباحث و مواد آزمون و با بهره‌گیری از حوزه‌های شناختی به طراحی سوال‌های مناسب دست بزنند (شریفی، ۱۳۹۳؛ شریفی و گنجی، ۱۳۸۳؛ گنجی، ۱۳۷۱). هدف نهایی تحلیل سوال‌های آزمون، بررسی سوال‌ها در جهت تعیین نقاط قوت و ضعف آنهاست.

#### ۱.۱. بیان مسئله

همانطور که می‌دانیم یادگرفتن و آموزش زبان دوم/خارجی یکی از شاخه‌های فرعی زبان‌شناسی کاربردی است (اشمیت، ۲۰۱۲) و ارزشیابی و سنجش زبان آموز جزء جدایی‌ناپذیر و مکمل آموزش زبان است. آموزش زبان فارسی نیز از این قاعده مستثنی نیست؛ زبان فارسی یکی از قدیمی‌ترین زبان‌های دنیاست که به واسطه قدمت در حوزه‌های تاریخی، ادبی و علمی غنای بسیار داشته و امروزه این زبان، زبان رایج سه کشور ایران، افغانستان و تاجیکستان می‌باشد. در حال حاضر نیز بسیاری از فارسی‌آموزان با انگیزه‌آشنایی با زبان و ادبیات فارسی، فرصت‌های شغلی و تحصیل در ایران در حال یادگیری زبان فارسی هستند. بنابراین، یکی از راه‌های موثر معرفی و گسترش زبان فارسی، همین فارسی‌آموزان می‌باشند که هر کدام پس از اتمام دوره تحصیلی می‌توانند به معرفی و توسعه زبان و فرهنگ ایرانی بپردازند (گلپور، ۱۳۹۴). اما یکی از مشکلات جدی در آموزش زبان فارسی به غیر فارسی‌زبانان کمبود آزمون‌های علمی معتبر است. در بیشتر موارد، بدون سنجش صحیح فارسی‌آموزان خارجی بر اساس ملاک و معیار علمی، آنان را به دانشگاه‌های ایران وارد می‌کنیم. مرکز آرفای دانشگاه فردوسی مشهد با برگزاری یک آزمون تعیین سطح، میزان توانایی زبان آموز را برای ورود به کلاس‌های این مرکز می‌سنجد. این پژوهش بر

آن است تا با بررسی این آزمون، میزان اعتبار آزمون و انطباق با معیارهای استاندارد را مورد سنجش قرار دهد.

### ۲.۱. سوال‌های پژوهش

سوال‌هایی که پژوهش حاضر قصد پاسخ‌گویی به آن‌ها را دارد، به شرح زیر است:

- آزمون تعیین سطح مرکز آرفای دانشگاه فردوسی مشهد، از لحاظ پایایی و روایی دارای چه ویژگی‌هایی است؟
- میزان رضایت مدرسان از آزمون تعیین سطح مرکز آرفا دانشگاه فردوسی مشهد چقدر است؟
- میزان رضایت فارسی‌آموزان از آزمون تعیین سطح مرکز آرفا دانشگاه فردوسی مشهد چقدر است؟
- طبق نظر مدرسان مرکز آرفای دانشگاه فردوسی مشهد، چند درصد از فارسی‌آموزان به طور صحیح از طریق آزمون مورد نظر تعیین سطح شده‌اند؟

### ۳.۱. اهداف پژوهش

تجزیه و تحلیل سوال‌های یک آزمون، فرایندی است که شامل بررسی کیفیت و کمیت تک تک سوال‌ها و کل آزمون می‌شود. از اهداف اصلی این پژوهش، مطابقت دادن اجزای آزمون مورد تحقیق با تعریف آزمون استاندارد است و همچنین تحلیل آزمون تعیین سطح مرکز آموزش زبان فارسی به غیر فارسی‌زبانان دانشگاه فردوسی مشهد از لحاظ پایایی و روایی است. هدف این آزمون، سنجش توانایی کلی فرد در به‌کارگیری زبان مورد نظر است و تعیین توانایی زبان‌آموز برای ورود به کلاس‌های درس است. همچنین در ادامه با مصاحبه‌ای که از مدرسان و فارسی‌آموزان گرفته می‌شود، میزان رضایت آن‌ها از آزمون مورد بررسی قرار می‌گیرد و در انتها درصد فارسی‌آموزانی که طبق نظر مدرسان به طور صحیح تعیین سطح شده‌اند، بررسی می‌شود.

## ۲. چارچوب نظری

چارچوب نظری پژوهش حاضر به مبانی نظری مربوط به آزمون‌سازی می‌پردازد.

### ۱.۲. مبانی نظری آزمون‌سازی

قبل از پرداختن به موضوع آزمون‌سازی، بهتر است با چند اصطلاح کلیدی یعنی همان اندازه‌گیری، ارزشیابی و سنجش که گاهی به‌جای آزمون به‌کار می‌روند، آشنا شویم. ارزشیابی به فرایند نظام‌مند جمع‌آوری، تحلیل و تفسیر داده گفته می‌شود. این بخش از آزمون‌سازی به مدرس یا معلم قدرت تصمیم‌گیری در مورد میزان دستیابی موفق فراگیر به اهداف آموزشی از پیش تعیین شده را می‌دهد (سیف، ۱۳۸۲؛ ۱۳۹۰؛ فرهادی، جعفرپور، و بیرجندی، ۲۰۱۷؛ بکمن، ۲۰۰۸). آزمون‌ها زیر مجموعه ارزشیابی هستند (براون، ۲۰۰۱؛ ۲۰۰۳؛ ۲۰۰۸).

### ۱.۲.۲. اندازه‌گیری

در اندازه‌گیری، ویژگی‌ها یا صفات افراد مشخص می‌شود و میزان وجود ویژگی‌ها یا صفات به صورت کمی گزارش می‌شود. بنابراین، می‌توان گفت که اندازه‌گیری عبارت است از فرایندی که تعیین‌کننده میزان برخورداری فرد از یک ویژگی است (هیوز، ۱۹۹۸؛ فرهادی، جعفرپور، و بیرجندی، ۲۰۱۷؛ فالچر، ۲۰۰۷).

### ۲.۲.۲. آزمون

فرایند اندازه‌گیری نیاز به ابزار مناسب و دقیق دارد که به آن آزمون می‌گویند. آزمون‌ها طوری آماده می‌شوند که در آن آزمودنی، تمرین‌های خاصی را انجام می‌دهد. مانند نوشتن جواب تشریحی یا انتخاب جواب چهارگزینه‌ای و غیره.

### ۳.۲.۲. سنجش

سنجش به فرایندی گفته می‌شود که طی آن درک بهتری از دانش کنونی فراگیر به دست می‌آید. در این فرایند ویژگی‌های کیفی و کمی دخیل هستند. این بدان معنا است که سنجش می‌تواند به سادگی قضاوت شخصی مدرس بر اساس مشاهده رفتار زبان‌آموزان در کلاس باشد و یا می‌تواند به پیچیدگی یک آزمون استاندارد چند ساعته باشد (بیرجندی و مصلانژاد، ۱۳۸۹).

## ۲.۲.۴. روایی

طراحی و استاندارد کردن آزمون، عملی پیچیده است و مستلزم دانش و مهارت در زمینه‌های روان‌سنجی است (شریفی، ۱۳۹۳). در زیر، مراحل تدوین و استاندارد کردن آزمون به اختصار شرح داده می‌شود. طراحان آزمون ابتدا باید هدف یا هدف‌های آزمون را که قصد ساختن آن دارند، تعیین کنند. پس از تعیین اهداف و محتوای آزمون، سوال‌های آزمون که بر اساس اصول نظری علوم وابسته به آزمون است، نوشته می‌شوند (کرمی، ۱۳۸۲).

روایی آزمون، عبارت است از میزان کارایی آن برای اندازه‌گیری خصیصه‌ای که به منظور اندازه‌گیری آن خصیصه ساخته شده است (سیف، ۱۳۸۲؛ ۱۳۹۰). روایی انواع متفاوتی دارد که برخی از آنها شامل روایی صوری، روایی محتوایی، روایی ملاکی و روایی سازه می‌شود (درنیه، ۲۰۰۷؛ کرمی، ۱۳۸۲؛ هومن، ۱۳۷۴).

## ۲.۲.۵. پایایی و روش تعیین آن

ارزیابی عینی آزمون‌ها مستلزم بررسی درجه پایایی و روایی آن‌ها در موقعیت معین است (درنیه، ۲۰۰۷). پایایی، یک ابزار اندازه‌گیری است که به دقت آن اشاره می‌کند. یک آزمون پایا، آزمونی است که نتایج به‌دست آمده از آن قابل تکرار باشد و از این لحاظ معنایی نزدیک به تعمیم‌پذیری در روایی بیرونی نیز دارد (درنیه، ۲۰۰۷؛ سیف، ۱۳۷۵).

روش تعیین پایایی به دو صورت است: یکی تعیین پایایی نمره‌های آزمون که به روش‌هایی چون باز آزمایی، آزمون هم‌تا و غیره اجرا می‌شود؛ و صورت دوم پایایی که به هم-بستگی درونی سوال‌های آزمون بستگی دارد و از راه‌های متفاوتی چون دو نیم کردن، کودر ریچارسون، آلفای کرونباخ و پایایی بین مصححان قابل محاسبه است (دلاور، ۱۳۷۷؛ فرهادی، جعفرپور و پیرجندی، ۲۰۱۷). در پژوهش حاضر از روش آلفای کرونباخ برای محاسبه پایایی آزمون تعیین سطح استفاده شده است.

## ۳. پیشینه‌ی پژوهش

درباره‌ی ارزیابی آزمون‌های گوناگون تاکنون تحقیقات مختلفی صورت گرفته است. از جمله موسوی (۱۳۷۹) در پژوهشی با عنوان «بررسی راهکارهای نظری

آزمون بسندگی و تهیه یک آزمون بسندگی استاندارد زبان فارسی برای غیر فارسی زبان «، به طراحی آزمونی چهار بخشی با هدف بررسی توانایی های مهارت زبانی فارسی آموزان پرداخته است. البته با این - که آزمون شامل مهارت های شنیدن و خواندن و اجزای زبانی دستور زبان و واژگان می - شود، متأسفانه به مهارت های تولیدی مثل گفتار و نوشتار، نمی پردازد.

شیخ الاسلامی (۱۳۷۸)، در پایان نامه ای با عنوان «بررسی روایی آزمون ورودی تحصیلات تکمیلی در ایران»، به بررسی روایی آزمون ورودی دوره های کارشناسی ارشد پرداخته است. شیخ الاسلامی به مقایسه نمرات آزمون ورودی تحصیلات تکمیلی دوره های کارشناسی ارشد ناپیوسته سال های ۱۳۷۶ و ۱۳۷۷ با نمرات آزمون تافل می پردازد. برای این مقایسه، وی میزان همبستگی بین دو آزمون را مورد بررسی قرار داده است.

قاسمی و رزانه (۱۳۸۴) در پایان نامه ای با عنوان «بررسی روایی پیشگویی قسمت زبان عمومی آزمون های ورود کارشناسی ارشد مجموعه زبان انگلیسی» به بررسی روایی پیشگویی بخش زبان عمومی آزمون های ورودی کارشناسی ارشد مجموعه زبان های انگلیسی سال های ۱۳۸۱ و ۱۳۸۲ پرداخته است. برای رسیدن به این هدف، تعداد ۶۷ نفر از دانشجویان مقطع کارشناسی ارشد در سه رشته آموزش، مترجمی و زبان و ادبیات انگلیسی که در دانشگاه علامه طباطبایی تحصیل کردند، انتخاب شدند. آزمون معیار انتخاب شده، آزمون آیلتس بود که به شکل فراگیر در سراسر جهان استفاده می شود. نتایج این پژوهش نشان داد که قسمت اول زبان عمومی ورودی مقطع کارشناسی ارشد مجموعه زبان انگلیسی و زیر مجموعه های آن دارای روایی پیشگویی مناسبی است. دوم این که قسمت مصاحبه آزمون زبان عمومی که در سال های ۱۳۸۱ و قبل از آن مورد استفاده قرار می گرفت، پیشگویی مناسبی از عملکرد آتی داوطلبان در آزمون مهارت زبان انگلیسی دارد.

در پایان نامه ای با عنوان «بررسی روایی پیشین آزمون جدید کنکور زبان انگلیسی» پیشقدم (۱۳۸۱) به بررسی روایی آزمون کنکور زبان انگلیسی پرداخته است. با توجه به عدم موفقیت برنامه چهار ساله زبان انگلیسی در رشته های زبان و ادبیات، دبیری و آموزش زبان در دانشگاه های کشور، سازمان سنجش

در سال ۱۳۸۱ تصمیم بر ایجاد تغییراتی در آزمون رشته‌های زبان انگلیسی گرفت که حاصل آن آزمون جدیدی تحت عنوان زبان تخصصی شد. پژوهش پیشقدم به بررسی روایی پیشین آزمون جدید زبان انگلیسی می‌پردازد. نتایج این پژوهش که با دو گروه از دانشجویان شبانه و روزانه انجام شد، مشخص کرد که دانشجویان روزانه هم در نمرات کنکور و هم در نمرات پایان ترم از دانشجویان شبانه بهتر عمل کردند که نشان‌گر عملکرد خوب آزمون جدید سازمان سنجش است.

در مقاله‌ای با عنوان «طراحی و روایی‌سنجی آزمون بسندگی زبان فارسی» قنسولی (۱۳۸۹) از میان مدل‌های مختلف اجزای بسندگی زبان، مدل چند بعدی زبان به کار گرفته است. در این مدل به دو مهارت اصلی زبان یعنی خواندن، شنیدن و دو جزء زبانی دستور و واژگان توجه شده است. انعکاس این مدل در مدل‌های استاندارد به کار رفته شده در تافل و آیلتس دیده می‌شود. در طراحی این آزمون تنها به دو مهارت اصلی ادراکی خواندن و شنیدن توجه شده است و مهارت‌های تولیدی صحبت کردن و نگارش به دلیل محدودیت‌های تحقیق بررسی نشده‌اند. در این پژوهش علاوه بر طراحی یک نمونه آزمون بسندگی زبان فارسی که شامل مهارت‌های شنیدن، خواندن و اجزای زبانی دستور و واژگان است، روایی و پایایی آزمون نیز محاسبه شده است. برای محاسبه روایی آزمون، از دو مرحله که شامل تحلیل بازنگری سه دانشجوی خارجی دوره دکتری زبان و ادبیات فارسی و استفاده از نظریه بحث در سنجش روایی است، استفاده شده است. پایایی آزمون هم از طریق محاسبه آلفای کرونباخ مورد بررسی قرار گرفته است. نتایج این آزمون نشان داد که دانشجویان خارجی در خواندن و درک مفاهیم عملکرد بهتری نسبت به مابقی مهارت‌های آزمون داشته‌اند و پایایی آزمون ۷۰٪ تعیین شده است.

در پژوهشی با عنوان «آزمون مهارتی فارسی بر پایه چهار مهارت اصلی زبانی» جلیلی (۱۳۹۰) تلاش کرده است یک آزمون بسندگی در زبان فارسی طراحی کند. تفاوت پژوهش موسوی با جلیلی این است که موسوی (۱۳۷۹) از ساختار آزمون تافل PBT استفاده کرده و آزمون وی شامل بخش‌های مجزایی برای دستور زبان و واژگان می‌شد، ولی به دو مهارت تولیدی زبان، یعنی گفتار و



نوشتار توجه نکرده است؛ اما جلیلی از ساختار کلی آزمون آیلتس استفاده کرده است و بخش‌های دستور و واژگان را به طور ضمنی در مهارت‌های دیگر می‌سنجد و بخش جداگانه‌ای برای آن در نظر نگرفته است و در پایان دو بخش مجزا برای ارزیابی توانایی‌ها فارسی‌آموز در مهارت‌های گفتار و نوشتار در نظر می‌گیرد. جلیلی نام آزمون خود را آمفا (آزمون مهارت زبان فارسی) می‌گذارد. شقاقی و متانی (۱۳۹۰) در مقاله‌ای با عنوان «هنجاریابی آزمون بین‌المللی زبان فارسی» به هنجاریابی آزمون استاندارد موسسه کانون زبان ایران پرداخته‌اند. این آزمون به ترتیب شامل بخش‌های شنیدن، واژگان، خواندن و دستور است و مهارت‌های گفتار و نوشتار جایگاهی در این آزمون ندارند.

در پژوهشی با عنوان «ارزیابی روایی و پایایی آزمون وزارت علوم و تحقیقات با هدف دستیابی به یک آزمون مهارت برای زبان فارسی» ممقانی (۱۳۹۰) به ارزیابی روایی و پایایی آزمون زبان انگلیسی وزارت علوم، تحقیقات و فناوری (MSRT) پرداخته است. هدف از این پژوهش، استخراج الگویی مناسب از MSRT برای طراحی آزمون بسندگی در زبان فارسی است. جهت دستیابی به این الگو از دو آزمون زبان فارسی دیگر یعنی، آمفا (آزمون مهارت زبان فارسی) و آزمون کانون زبان ایران نیز استفاده شد. روایی و پایایی این دو آزمون نیز محاسبه شده است. در نهایت نتایج این پژوهش نشان داده که آزمون MSRT نمی‌تواند الگوی مناسبی برای آزمون بسندگی زبان فارسی باشد؛ چرا که تافل پایه این آزمون بوده و به نظر ممقانی، آزمون آیلتس سازگاری بیشتری با هدف سنجش فارسی‌آموزان دارد. همچنین مشخص شد که آزمون آمفا نسبت به آزمون کانون، نمونه بهتری برای الگوبرداری است.

### ۳/۱. پژوهش‌های انجام شده در مرکز آذفا دانشگاه فردوسی مشهد

مرکز آموزش زبان فارسی به غیر فارسی‌آموزان دانشگاه فردوسی مشهد در مهر ماه سال ۱۳۹۲ مجوز رسمی وزارتی برای تأسیس دریافت کرد. این مرکز دارای ۲۵ کلاس است. مرکز آموزش زبان فارسی به غیر فارسی‌زبانان با ایجاد بستری مناسب برای ادامه تحصیل فارسی‌آموزان در ایران، فرصتی برای یادگیری زبان فارسی در بازه زمانی کوتاه ایجاد کرده است (وبگاه مرکز آذفا، ۱۳۹۸). پژوهش‌های انجام شده تا کنون در مرکز آذفا دانشگاه فردوسی مشهد در

حوزه‌های خواندن و نوشتن آموزش زبان فارسی به غیر فارسی زبان، عملکرد مدرسان، تهیه و تدوین متون برای غیر فارسی‌زبانان بوده است. با توجه به سابقه مرکز آزفا دانشگاه فردوسی و بررسی پژوهش‌های انجام شده، تا به حال آزمون‌های این مرکز از لحاظ روایی‌سنجی و اعتبارسازی مورد بررسی قرار نگرفته است که هدف نهایی این تحقیق در این راستا تنظیم شده است.

#### ۴. روش شناسی پژوهش

در بخش روش‌شناسی به زیربخش‌هایی چون طرح پژوهش، آزمودنی‌ها، ابزار پژوهش، جمع‌آوری داده و شیوه تحلیل داده پرداخته می‌شود.

##### ۴.۱. طرح پژوهش

شیوه انجام پژوهش حاضر به منظور تجزیه و تحلیل داده‌ها، «توصیفی-تحلیلی» است. داده‌های پژوهش، از طریق انجام آزمونی در قالب آزمون تعیین سطح مرکز آزفای دانشگاه فردوسی مشهد جمع‌آوری شده است. به منظور بررسی و تعیین اعتبار آزمون پیشنهاد شده، از روش میدانی استفاده شده است.

##### ۴.۲. آزمودنی‌ها

جامعه آماری این پژوهش به دو بخش تقسیم می‌گردد. اولین بخش شامل تمامی متقاضیان فارسی‌آموزان مرکز آموزش زبان فارسی به غیر فارسی‌زبانان دانشگاه فردوسی مشهد است. تعداد کل این جامعه آماری ۵۱ نفر است و بازه سنی ۲۵ تا ۳۸ می‌باشد. لازم به ذکر است در این جامعه آماری ۳۶ آزمودنی مرد و ۱۵ آزمودنی زن بودند که انحراف معیار  $1/77$  می‌باشد. دومین بخش شامل ۱۵ مدرس مرکز آزفای دانشگاه فردوسی مشهد که ۵۱ نفر بعد از تعیین سطح در کلاس‌های آن‌ها قرار گرفتند. از این مدرسان در مورد ویژگی‌های آزمون و درستی تعیین سطح آزمودنی‌ها مصاحبه به عمل آمد. این مدرسان ۲۹-۳۳ ساله بودند. ۲ نفر مرد و ۱۳ نفر زن و سابقه کار آن‌ها بین ۲-۴ سال در مرکز آزفای دانشگاه فردوسی مشهد بود. ۹ نفر دارای مدرک کارشناسی ارشد و ۶ نفر هم دانشجوی دکترا بودند.

#### ۳.۴. ابزار پژوهش

- ابزارهایی که در این پژوهش به کار گرفته شده اند شامل:
۱. اولین و اصلی ترین ابزار به کار رفته در پژوهش حاضر، آزمون تعیین سطح مرکز آرفای دانشگاه فردوسی مشهد است.
  ۲. به غیر از آزمون، از نرم افزار SPSS نسخه ۲۳ برای محاسبه ضریب آلفای کرونباخ جهت تعیین پایایی درونی آزمون استفاده شد.
  ۳. در مرحله بعدی از درجه دشواری و تشخیص (تمییز) برای اندازه گیری روایی آزمون استفاده شد.
  ۴. در مرحله بعد فرم نظرسنجی که در اختیار مدرسان برای ارزیابی آزمون قرار می گیرد.
  ۵. در آخرین مرحله مصاحبه به صورت کتبی از فارسی آموزان و مدرسان صورت گرفته که در آن نظر آن‌ها در مورد آزمون مورد سنجش قرار گرفت.

#### ۱.۳.۴. محتوا و ساختار آزمون

آزمون تعیین سطح مرکز آرفای دانشگاه فردوسی مشهد، از نوع چند گزینه‌ای است که در محتوای آن از واژگان فارسی که در کتاب‌های فارسی این مرکز موجود است، استفاده شده است و از اصطلاحات و ترکیباتی استفاده شده که درک و فهم آن نیاز به تخصص علمی خاصی ندارد. در این آزمون دانش دستور و واژگانی سنجیده می‌شود (جدول ۱) که از مبنای کتاب‌های موجود در مرکز آرفای دانشگاه فردوسی مشهد است. اگر فارسی آموز در سطح بالایی از دانش واژگان و دستور قرار داشته باشد، آزمون شفاهی‌ای بر مبنای کتاب‌های موجود در مرکز گرفته می‌شود. تعداد سوالات آزمون ۵۰ عدد می‌باشد که به هر سوال یک امتیاز اختصاص داده شده است که به ترتیب از سوال ۱ تا ۱۰ مربوط به کتاب اول آموزش نوین زبان فارسی، ۱۱ تا ۲۰ مربوط به کتاب دوم آموزش نوین زبان فارسی، ۲۱ تا ۳۰ مربوط به کتاب سوم آموزش نوین زبان فارسی، ۳۱ تا ۴۰ مربوط به کتاب چهارم آموزش نوین زبان فارسی و ۴۱ تا ۵۰ مربوط به کتاب پنجم آموزش نوین زبان فارسی است و با توجه به نمره کسب شده سطح زبان آموز تعیین می‌شود. مدت زمان آزمون ۳۰ دقیقه در نظر گرفته شده است و آزمون شامل بخش‌های درک واژه و درک متن می‌باشد. در این آزمون نمره منفی در نظر گرفته نمی‌شود.

### جدول ۱: ساختار آزمون تعیین سطح

بخش بندی	جزئیات بخش
واژگان	درک واژه شامل مقوله‌های واژگانی (اسم، فعل، حرف اضافه، فعل مرکب، صفت)
متن به هم ریخته	به منظور ارزیابی توانایی تشخیص پیوستگی و انسجام درون متنی
درک متن	درک دگرگفت و ارجاع درون متن

#### ۴.۴. برگزاری آزمون

زمانی که برای این آزمون در نظر گرفته شد برطبق نظر طراحان ۳۰ دقیقه برای پاسخ‌گویی به ۵۰ سوال چهارگزینه‌ای بود. پس از جمع‌آوری آزمون‌ها پژوهش‌گران به نمره‌گذاری آن‌ها پرداختند. نمره منفی در این آزمون مد نظر قرار نگرفت و آزمون در مرکز آف‌زای دانشگاه فردوسی مشهد و در نیمسال اول سال ۹۸ برگزار شد.

#### ۴.۵. شیوه تحلیل داده‌ها

تجزیه و تحلیل داده‌ها طی مراتب ذیل صورت پذیرفت:

ابتدا آمار توصیفی در مورد آزمون مورد بررسی قرار گرفت. سپس به منظور تعیین پایایی درونی آزمون چهار گزینه‌ای تعیین سطح زبان فارسی به غیر فارسی‌زبانان، فرمول آلفای کرونباخ مورد استفاده قرار گرفت. علاوه بر پایایی، روایی درونی آزمون‌ها نیز از طریق محاسبه شاخص درجه دشواری و درجه تشخیص (تمییز) سوال‌های آزمون با استفاده از ارزش پی به دست آمد.

ارزش پی به عنوان شاخص درجه دشواری با استفاده از فرمول و بر اساس تقسیم تعداد کل پاسخ‌های صحیح گروه قوی و ضعیف به تعداد کل آزمودنی‌ها، محاسبه می‌شود (دانشگاه وسکانسن، ۲۰۲۰؛ دانشگاه واشنگتون، ۲۰۲۰)

#### ۵. ارائه، واکاوی و تفسیر داده‌ها

با توجه به گسترش آموزش زبان فارسی به غیر فارسی‌زبانان به عنوان زبان دوم / خارجی، آزمون سازی نقشی حیاتی در بحث آموزش ایفا می‌کند. در همین

راستا در این بخش از پژوهش به تجزیه و تحلیل داده‌ها و ارائه یافته‌های حاصل از ابزارهای تحقیق توسط روش‌های آماری پرداخته می‌شود. لذا برای جمع آوری اطلاعات در راستای اهداف تحقیق، آزمونی بین ۵۱ نفر از فارسی آموزان مرکز آموزش زبان فارسی به غیر فارسی زبانان دانشگاه فردوسی مشهد که بازه سنی بین ۲۵ تا ۳۸ دارند توزیع و جمع آوری گردید. پس از برگزاری آزمون و جمع آوری برگه‌های آزمون به نمره گذاری آن‌ها پرداخته شد. به منظور تجزیه و تحلیل داده‌های آزمون از نرم افزار اس پی اس اس نسخه ۲۳ استفاده شد.

#### ۱.۵. تحلیل سوالات آزمون

هدف از تحلیل سوالات آزمون بررسی تک تک سوالات و تعیین میزان دقت و نارسایی آنهاست. در تحلیل سوال‌های آزمون، نقاط ضعف و قدرت یک آزمون و کیفیت سوالات آن تعیین می‌شود. پس لازم است که آزمون‌گر بعد از اجرای هر آزمون، سوالات آن را مورد تحلیل قرار دهد و با استفاده از نتایج آن به تجدید نظر در آزمون و بهبود کیفیت سوال‌ها برای استفاده بعدی اقدام نماید (سیف، ۱۳۹۰؛ سلیمی زاده، ۱۳۷۷). در فرایند تحلیل سوالات آزمون پژوهش حاضر، تمرکز بر ضریب دشواری، ضریب تشخیص و پایایی است که در ادامه به این مباحث می‌پردازیم.

طبق روش تحلیل سوال، جهت محاسبه شاخص‌های آماری باید دو گروه قوی و ضعیف آزمودنی‌ها تعیین شوند و نحوه پاسخ دهی آنان به گزینه‌های مختلف سوال به دست آیند. نمره‌های آزمون از بالاترین نمره به پایین‌ترین نمره تنظیم می‌شوند و بر اساس تعداد اوراق درصدی از بالا و پایین به عنوان گروه قوی و ضعیف مشخص می‌گردند و چگونگی جواب دادن سوال در جدول‌های تحلیل سوال درج می‌گردند. درصد کل آزمودنی‌هایی که به یک سوال جواب مثبت می‌دهند، ضریب دشواری آن سوال را مشخص می‌کنند.

#### ۱.۱.۵ محاسبه ضریب دشواری

بنا به تعریف، درصد کل آزمودنی‌هایی که به یک سوال جواب درست می‌دهند، ضریب دشواری آن سوال است که از طریق فرمول زیر محاسبه می‌شود:

$$\text{ضریب دشواری} = \frac{\text{تعداد پاسخ صحیح گروه ضعیف} + \text{تعداد پاسخ صحیح گروه قوی}}{100 \times (\text{افراد گروه تعداد کل})}$$

اگر درجه سختی سوال بیش از ۸۰٪ باشد سوال آسان است و اگر درجه سختی کمتر از ۳۰٪ باشد سوال آزمون بسیار سخت است و اگر در بازه ۳۰٪ تا ۸۰٪ قرار بگیرد ضریب سختی سوال مطلوب ارزیابی می‌شود (دانشگاه وسکانسن، ۲۰۲۰؛ دانشگاه واشنگتون، ۲۰۲۰؛ کرگهد و نمرف، ۲۰۰۴).

### ۲.۱.۵ تفسیر ضریب تشخیص

از فعالیت‌های دیگر که برای تحلیل سوالات چند گزینه‌ای انجام می‌شود محاسبه‌ی ضریب تشخیص یا تمییز می‌باشد. ضریب تشخیص که با (Discrimination) نشان داده می‌شود قدرت سوال را در تشخیص بین گروه قوی و گروه ضعیف آزمودنی‌ها تعیین می‌کند. هر قدر ضریب تشخیص بزرگتر باشد، قدرت تشخیص آن سوال بیشتر و هر قدر کوچکتر باشد قدرت تشخیص آن کمتر خواهد بود. اگر ضریب سوالی منفی باشد نشان می‌دهد که در آن سوال گروه قوی بدتر از گروه ضعیف عمل کرده‌اند. بنابراین تمام ضریب‌های تشخیص سوال آزمون باید مثبت باشند. مقدار ۱۰٪ به بالا ضریب تشخیص مطلوب به شمار می‌رود و اگر ضریب سوالی کمتر از این مقدار باشد سوالاتی هستند که به تجدید نظر نیاز دارند. برای محاسبه ضریب تشخیص یک سوال از فرمول زیر استفاده می‌شود (دانشگاه وسکانسن، ۲۰۲۰؛ دانشگاه واشنگتون، ۲۰۲۰؛ کرگهد و نمرف، ۲۰۰۴).

$$\text{ضریب تشخیص} = \frac{\text{تعداد پاسخ صحیح گروه ضعیف} + \text{تعداد پاسخ صحیح گروه قوی}}{100 \times (\text{افراد گروه تعداد کل})}$$

### ۳.۱.۵ نتایج ضریب دشواری و ضریب تشخیص آزمون

نتایج ضریب دشواری و ضریب تشخیص آزمون به شرح زیر است:  
جدول ۲ توزیع ضریب دشواری و تشخیص برای سطح کیفی سوالات آزمون را نشان می‌دهد.

جدول ۲: نتایج درجه تشخیص و درجه سختی مربوط به هر گویه

سوال	درجه سختی	درجه تشخیص
۱	٪۶۸	۰
۲	٪۳۶	٪۱۲
۳	٪۵۲	٪۸
۴	٪۴۴	٪۲
۵	٪۴۶	٪۶
۶	٪۲۸	٪۸
۷	٪۱۶	۰
۸	٪۶۰	٪۸
۹	٪۶۲	٪۶
۱۰	٪۳۸	٪۱۰
۱۱	٪۳۶	٪۱۲
۱۲	٪۳۸	٪۱۴
۱۳	٪۱۶	٪۴
۱۴	٪۳۲	٪۴
۱۵	٪۲۸	٪۸
۱۶	٪۴۰	٪۸
۱۷	٪۶۲	٪۶
۱۸	٪۶۰	٪۴
۱۹	٪۴۸	۰
۲۰	٪۵۰	٪۶

۲۱	%۳۴	%۱۰
۲۲	%۲۲	%۱۴
۲۳	%۳۲	.
۲۴	%۳۶	%۴
۲۵	%۳۲	%۱۶
۲۶	%۲۸	.
۲۷	%۶۰	.
۲۸	%۴۰	%۴
۲۹	%۲۶	%۱۸
۳۰	%۳۲	%۱۲
۳۱	%۳۴	%۲۲
۳۲	%۳۴	%۱۸
۳۳	%۲۸	%۱۲
۳۴	%۳۴	%۱۰
۳۵	%۳۲	%۱۶
۳۶	%۴۸	%۴
۳۷	%۵۸	%۲
۳۸	%۶	%۲
۳۹	%۵۴	%۱۰
۴۰	%۴۶	%۱۰
۴۱	%۳۰	%۶
۴۲	%۴	.



۴۳	%۴۴	%۴
۴۴	%۲۸	.
۴۵	%۲۲	%۲
۴۶	%۲	%۴
۴۷	%۴۲	%۱۰
۴۸	%۳۴	%۱۴
۴۹	%۵۲	%۱۲
۵۰	%۲۶	%۱۸

بر اساس جدول ۲ توزیع فراوانی ضریب دشواری برای سطح کیفی سوالات آزمون به صورت زیر می‌باشد: از مجموع ۵۰ سوال، ۱۵ سوال معادل ۳۰ درصد سخت و ۳۵ سوال معادل ۷۰ درصد نرمال ارزیابی شده است. شماره سوال‌ها بر اساس درجه دشواری (سخت، نرمال، آسان) در جدول ۳ ذکر شده است و با توجه به آمار به دست آمده هیچ کدام از سوال‌ها آسان ارزیابی نشده‌اند.

جدول ۳: نتایج درجه دشواری هر گویه

سخت (کمتر از ۳۰)	آسان (بیشتر از ۸۰)	نرمال (۳۰ تا ۸۰)
۶، ۷، ۱۳، ۱۵، ۲۲، ۲۶، ۲۹		۳۷، ۳۹، ۴۰، ۱، ۲، ۳، ۴، ۵، ۸، ۹،
۳۱، ۳۳، ۴۱، ۳۸، ۴۴، ۴۵		۱۰، ۱۱، ۱۲، ۳۵، ۳۶، ۱۸، ۱
۴۶، ۵۰		۱۴، ۱۶، ۱۷، ۲۰، ۲۱، ۳۸، ۳۴
		۲۳، ۲۴، ۲۵، ۴۰، ۲۷، ۲۸، ۳۰
		۴۲، ۴۳، ۴۷، ۴۸، ۴۹

و بر اساس جدول ۴ توزیع ضریب تشخیص آزمون به صورت زیر می‌باشد: از مجموع ۵۰ سوال، ۱۹ سوال معادل ۳۸ درصد مطلوب و ۲۹ سوال معادل ۶۲ درصد ضعیف ارزیابی شده است.

جدول ۴: درجه تشخیص هر گویه

کمتر از ۱۰٪ ضعیف	۱۰٪ به بالا مطلوب
۳۶، ۳۷، ۳۸، ۴۱، ۴۲، ۴۳	۱۱، ۱۰، ۲، ۲۱، ۲۲
۴۴، ۴۵، ۴۶، ۲۸، ۲۷، ۲۶	۲۵، ۲۹، ۴۹، ۵۰، ۴۸
۲۴، ۲۳، ۲۰، ۱۹، ۱۸، ۱۷	۴۷، ۴۰، ۳۹، ۳۵، ۳۴
۱۶، ۱۵، ۱۲، ۸، ۷	۳۳، ۳۲، ۳۱، ۳۰
۹، ۶، ۵، ۴، ۳، ۱	

شماره سوال‌ها براساس مطلوب بودن و ضعیف بودن سوال در جدول ۴ شده است.

#### ۲.۵. بررسی آمار توصیفی داده‌ها

پس از برگزاری، جمع آوری و نمره گذاری آزمون‌ها، نمرات آزمودنی‌ها با استفاده از نرم افزار آماری اس پی اس نسخه ۲۳ پردازش شد. جدول ۵ آمار توصیفی مربوط به دانش شرکت کنندگان را بر اساس ابزار اندازه گیری به کار رفته در این پژوهش نشان می دهد.

جدول ۵: آمارتوصیفی مربوط به دانش شرکت کنندگان در آزمون

انحراف- معیار	میانگین	حداکثر نمره	حداقل نمره	نمره کل	تعداد سوال	تعداد افراد	آزمون
۱۹/۷۶	۹۲/۸۷	۹۴	۶۸	۱۳۴	۵۰	۵۱	کتاب ۱
۱۶/۹۳	۸۰/۵۴	۷۶	۳۳	۱۵۴	۵۰	۵۱	کتاب ۲
۱۶/۴۰	۷۶/۹۲	۸۸	۴۲	۱۵۴	۵۰	۵۱	کتاب ۳
۱۸/۸۷	۶۷/۷۴	۱۳۴	۷۹	۱۵۴	۵۰	۵۱	کتاب ۴
۱۶/۵۳	۹۵/۹۷	۱۲۹	۷۲	۱۵۴	۵۰	۵۱	کتاب ۵

همان‌طور که در جدول ۵ مشاهده می شود میانگین نمره دانش زبانی آزمودنی‌ها در کتاب پنجم ۹۵/۹۷ است که در مقایسه با کتاب های دیگر حاکی از دشوار بودن سوالات مربوط به این بخش است. انحراف معیار کتاب ۱ در مقایسه با

دیگر کتاب‌ها بیشتر است که این مطلب حاکی از آن است که این سوالات توانسته در بین آزمودنی‌ها بهتر تمایز قائل شود. حداقل نمره مربوط به کتاب ۲ می‌باشد که بیشتر اشتباهات مربوط به به کارگیری افعال در زمان‌های درست بود که این می‌تواند به دلیل تداخل زبان اول و دوم باشد. با توجه به اطلاعات به دست آمده پژوهش‌گران در مورد پیشینه فارسی آموزان شرکت کننده در آزمون، اکثر آزمودنی‌های ما افرادی بودند که در یک دوره کوتاه مدت در مرکز مشغول به تحصیل بودند و پس از یک وقفه چند ماهه، مجدد به مرکز بازگشتند و با توجه به سیاست‌های مرکز مجدداً برای تعیین سطح مورد آزمون قرار گرفتند این مطلب باعث شده بود اکثر از انتهای آزمون شروع به پاسخ دادن بکنند این مطلب باعث شده که در تنظیم زمان دچار مشکل شده و همین دلیل باعث ایجاد تغییراتی در نمره کل شود.

با توجه به نمرات کسب شده فارسی آموزان و بررسی پاسخ‌های آزمون، سخت‌ترین بخش‌های آزمون مربوط به انتخاب فعل، فعل مرکب و درک متن می‌باشد. از میان حجم نمونه فقط یک آزمودنی بود که به سطح عالی رسیده است و او سال‌هاست که در ایران زندگی می‌کند ۶۰٪ در سطح میانه و ۳۳٪ در سطح مبتدی قرار گرفتند.

### ۳.۵. سنجش پایایی آزمون

از آنجایی که این آزمون دارای گویه‌های چند گزینه‌ای است، برای محاسبه پایایی این آزمون، از فرمول آلفای کرونباخ استفاده شده است. این فرمول به سنجش ثبات درونی (internal consistency) می‌پردازد. پایایی ثبات درونی مقیاسی برای همگن بودن یا سازگار بودن بخش‌های مختلف موجود در یک آزمون با یکدیگر است. در روش آلفای کرونباخ هر سوال آزمون با تک تک سوال‌های دیگر مقایسه می‌شود. در محاسبه آلفای کرونباخ گفته شده است که اگر ضریب آلفا بیش از ۷۰٪ باشد آزمون از پایایی قابل قبولی برخوردار است (مومنی و قیومی، ۱۳۸۹). پایایی این آزمون ۸۲٪ تعیین شده است که نتیجه بسیار مطلوبی برای این آزمون تلقی می‌شود (درنیه، ۲۰۰۷).

### ۴.۵. سنجش روایی آزمون

روایی، ویژگی دیگری از یک آزمون استاندارد است. پایایی، یک شاخص آماری است (جعفرپور، ۱۳۷۷). سوال روایی این است که آیا آنچه طراح

آزمون ادعا کرده است فقط در آزمون قرار گرفته است و یا ممکن است عوامل دیگری نیز در آزمون وارد شده باشند. البته روایی یک امر نسبی است.

#### ۱.۴.۵. سنجش روایی محتوایی و صوری

برای محاسبه روایی محتوایی و روایی (صوری) یک آزمون، باید گفت که در واقع هیچ روش آماری برای تعیین آن وجود ندارد. به جای روش آماری، از نظر متخصصان آن رشته استفاده می‌شود. در کل از متخصصان خواسته می‌شود که نظرشان در مورد میزان مطابقت سوال‌های آزمون با هدف آموزشی آن مجموعه را بیان کنند. در این آزمون برای دستیابی به این مهم، فرم نظرسنجی‌ای در اختیار ۱۵ مدرس مرکز آذفای دانشگاه فردوسی مشهد قرار گرفت. این مدرسان ۲۹-۳۳ ساله بودند. ۲ نفر مرد و ۱۳ نفر زن بودند و سابقه کار آن‌ها بین ۲-۴ سال در مرکز آذفای دانشگاه فردوسی مشهد بود. ۹ نفر دارای مدرک کارشناسی ارشد و ۶ نفر هم دانشجوی دکترا بودند. معیار نظرسنجی بر اساس شاخص‌های، عالی، خوب، متوسط و بد در جدول ۶ ارائه شده است که نتایج آن به شرح زیر است:

جدول ۶: نتایج نظر سنجی مدرسان درباره‌ی آزمون

گویه‌های نظر سنجی	عالی	خوب	متوسط	بد
محتوای آزمون	-	۹ نفر	۶ نفر	-
تناسب سوالات با سطح آزمودنی	۳ نفر	۱۰ نفر	۲ نفر	-
زمان پاسخ دهی به سوالات	۵ نفر	۸ نفر	۲ نفر	-
ترتیب سوالات آزمون	۹ نفر	۳ نفر	۳ نفر	-
تعداد سوالات و ساختار آزمون	-	۱۵ نفر	-	-

همانطور که در جدول ۶ قابل مشاهده است، در پاسخ به سوال اول که در مورد محتوای آزمون بود ۶۰٪ مدرسان مرکز، محتوای آزمون را مطابق با کتاب‌های مرکز آزمون اعلام کردند و در پاسخ به سوال دوم ۶۶٪ تناسب سوالات با سطح آزمون را خوب ارزیابی کردند و در پاسخ به زمان پاسخ دهی به سوالات ۵۵٪ زمان را متناسب اعلام کردند و در پاسخ به سوال چهارم بیشتر از ۶۰٪ مدرسان ترتیب سوالات را بر اساس محتوای کتاب‌ها مرکز آزمون مشاهده خوب ارزیابی کردند و در پاسخ به سوال آخر تمامی مدرسان تعداد سوالات و ساختار آزمون را مطابق با معیارهای آزمون دانستند و در نظر سنجی انجام شده هیچ یک از مدرسان گزینه «بد» را انتخاب نکرده بودند.

#### ۲.۴.۵ تحلیل رضایت فارسی آموزان

به منظور تحلیل کیفی روایی آزمون، میزان رضایت‌مندی فارسی آموزان از آزمون تعیین سطح مرکز آزمون دانشگاه فردوسی مشهد مورد پرسش قرار گرفت. در پایان آزمون، طی مصاحبه کتبی، از آزمودنی‌ها خواسته شد تا نظر خود را در مورد سطح آزمون، زمان پاسخ-گویی به آزمون و ساختار آزمون بازگو نمایند که جدول ۷ نشان دهنده نتایج آن می‌باشد:

جدول ۷: نظر سنجی از آزمودنی‌ها

سوالات	عالی	خوب	بد
سطح آزمون	۹ نفر	۲۳ نفر	۱۹ نفر
زمان آزمون	۵ نفر	۱۰ نفر	۳۶ نفر
ساختار آزمون	۲ نفر	۳۰ نفر	۱۹ نفر

#### ۶. نتیجه‌گیری و پیشنهادهای آموزشی و پژوهشی

در این بخش علاوه بر پاسخ به سوالات تحقیق و بحث در مورد یافته‌های آن، به ارائه پیشنهادهایی در خصوص پژوهش‌های بعدی پرداخته می‌شود.

#### ۱.۶ پاسخ به سوالات تحقیق

در این بخش قصد داریم از طریق تعیین درستی یا نادرستی فرضیه‌ها به

سوالات تحقیق پاسخ دهیم.

در خصوص سوال اول تحقیق و همچنین فرضیه اول که در مورد میزان پایایی و روایی آزمون تعیین سطح مرکز آzfای دانشگاه فردوسی مشهد مطرح شده بود، بر اساس واکاوی داده‌ها و تجزیه و تحلیل آن‌ها به طور خلاصه می‌توان گفت که آزمون چهارگزینه‌ای مرکز آzfای دانشگاه فردوسی مشهد از روایی صوری خوب (طبق نظر مدرسان) و روایی درونی ضعیفی برخوردار است. با توجه به تحلیل گویه‌ای انجام شده از ۵۰ سوال آزمون، حدود ۶۲/۰٪ از سوالات، از ضریب تشخیص ضعیف برخوردار بوده است که این نشان از روایی قابل قبولی نمی‌باشد. در خصوص میزان پایایی آزمون تعیین سطح مرکز آzfای دانشگاه فردوسی مشهد که برای محاسبه آن از فرمول آلفای کرونباخ استفاده شد، آزمون از شاخص پایایی ۸۰٪ برخوردار بود که نشان از ثبات درونی آزمون است.

به منظور پاسخ گویی به سوال دوم که در خصوص میزان رضایت مدرسان از آزمون تعیین سطح مرکز آzfای دانشگاه فردوسی مشهد قرار گرفته بود که سوالات درباره محتوای آزمون، تناسب سوالات با سطح آزمودنی، زمان پاسخ دهی به سوالات، ترتیب سوالات، تعداد سوالات و ساختار آزمون مورد سوال قرار گرفت که با توجه به نظر متخصصان از آنجا که پیکره و چارچوب دقیقی برای تعیین سطح آزمودنی در زبان فارسی وجود ندارد و شیوه تعیین سطح در این پژوهش با استناد به معیارهای خارجی است، به منظور اعتبار بخشی بیشتر به آزمون طراحی شده، برگه‌های آزمون در اختیار مدرسان آزمودنی‌ها قرار گرفت تا آنان با مقایسه نمره آزمودنی و سطح آزمون، نظر خود را درباره نتایج آزمون و مناسب بودن آن بازگو نمایند. از دیدگاه مدرسان، متناسب بودن نمره آزمون با سطح آزمون حدود ۷۸٪ اعلام شد.

در پاسخ به سوال سوم تحقیق که میزان رضایت مندی فارسی آموزان از آزمون تعیین سطح مرکز آzfای دانشگاه فردوسی مشهد را مورد پرسش قرار داده است در پایان آزمون، طی مصاحبه کتبی، از آزمودنی‌ها خواسته شد تا نظر خود را در مورد سطح آزمون، زمان آزمون و ساختار آزمون بازگو نمایند که از ۵۱ نفر آزمون دهنده ۲۳ نفر سطح آزمون را خوب ارزیابی کرده، ۳۶ نفر از کم بودن

زمان آزمون ناراضی بوده و ۳۰ نفر ساختار آزمون را خوب ارزیابی کرده بودند. به علت آشنایی نیمی از فارسی آموزان با زبان فارسی و اینکه اکثر فارسی آموزان دوره، سه تا چهار ماه در ایران بوده‌اند و در کلاس‌های مرکز شرکت داشته‌اند و هم‌چنین به دلیل داشتن یک وقفه بین گذراندن دوره‌ها، سیاست مرکز بر این مبنا صورت گرفته که باید مجدد مورد آزمون قرار گیرند..

و در پاسخ به سوال چهارم در مورد این که چند درصد از فارسی آموزان به طور صحیح تعیین سطح شده‌اند باید گفت طبق نمره کسب شده و نظر مدرسان حدود ۷۸٪ فارسی آموزان به طور صحیح تعیین سطح شده‌اند و این به این معنا است که حدود ۲۲٪ از فارسی آموزان درست تعیین سطح نشده و پس از آغاز کلاس، مدرسان مجبور به درخواست تعیین سطح مجدد آن دسته بودند. با وجود آن که پایایی و روایی آزمون دو خصوصیت مهم یک آزمون می‌باشند، ولی قابلیت اجرا و کاربردی بودن آزمون نیز از اهمیت بسیار بالایی برخوردار است. قابلیت اجرا و کاربردی بودن به معنای سهولت در برگزاری، امتیازدهی و تفسیر نتایج است (فرهادی، جعفرپور، و بیرجندی، ۲۰۱۷). آزمون طراحی شده هم، از لحاظ برگزاری، سهولت در نمره‌دهی و سهولت در تفسیر نتایج، توسط مدرسان مناسب ارزیابی شد.

از نظر بکمن و پالمر (۲۰۱۰) مهمترین چیزی که باید در طراحی یک آزمون زبان مطرح شود کاربردی بودن آن در موقعیت‌های مشابه زبانی است. پیش بینی می‌شود که این آزمون تعیین سطح تنها در مراکزی که از کتاب آموزش نوین زبان فارسی استفاده می‌شود، قابلیت استفاده داشته باشد و نمی‌تواند آزمونی کلی برای سنجش دانش زبان فارسی باشد؛ چراکه به دلیل منبع - محور بودن آزمون دچار محدودیت بزرگی است.

### ۲.۶. نتیجه‌گیری

نتایج حاصل از تحلیل داده‌ها که از بررسی آزمون تعیین سطح ۵۱ نفر فارسی‌آموز مرکز آذفای دانشگاه فردوسی مشهد به دست آمده است، نشان می‌دهد که در بخش تجزیه و تحلیل آزمون، می‌توان چنین ادعا کرد که در آزمون طراحی شده، میزان ضریب آلفای کرونباخ آزمون ۸۰٪ که نشانه پایا بودن آزمون است پایایی یک شرط لازم برای روایی یا اعتبار است به عبارت دیگر یک آزمون معتبر باید پایا هم باشد ولی یک آزمون پایا همیشه معتبر نیست (بکمن، ۲۰۰۸). گفته شد آزمونی روایی دارد که آنچه را که ادعا کردیم بسنجد. در این آزمون، روایی محتوایی آن نیز، بر اساس نظر سنجی از مدرسان زبان فارسی، خوب ارزیابی شده است و در روایی آزمون بر اساس تحلیل سوال انجام شده ضریب تشخیص در آزمون ضعیف ارزیابی شده است.

### ۳.۶. پیشنهادها

- با توجه به مسائلی که در این پژوهش مورد بررسی قرار گرفت، می‌توان موضوعات زیر را به عنوان پیشنهادهایی برای تحقیقات آینده مطرح نمود:
- مقایسه آزمون تعیین سطح در مراکز دیگر با آزمون تعیین سطح مرکز آذفای دانشگاه فردوسی مشهد؛
  - تاثیر زبان مادری در موفقیت آزمون دهندگان در آزمون تعیین سطح؛
  - بررسی تاثیر محیط دوم در میزان موفقیت آزمون دهندگان در آزمون تعیین سطح؛
  - تهیه یک آزمون بر اساس پیکره سطح بندی شده زبان فارسی و نیازهای غیر عمومی فارسی زبانان.



## فهرست منابع

۱. بازرگان، عباس. (۱۳۸۱). ارزشیابی آموزشی. تهران: سمت.
۲. بیرجندی و مصلا نژاد. (۱۳۸۹). تاثیرات و پیامدهای آزمون. اصفهان: سپاهان.
۳. پیشقدم، رضا. (۱۳۸۱). بررسی روایی پیشین آزمون جدید کنکور زبان انگلیسی. (رساله-ی دکتری منتشر نشده). دانشگاه علامه طباطبایی.
۴. جلیلی، اکبر. (۱۳۹۰). آزمون مهارتی فارسی بر پایه چهار مهارت اصلی زبانی. (پایان نامه کارشناسی ارشد منتشر نشده). دانشگاه علامه طباطبایی.
۵. جعفرپور، عبدالجواد. (۱۳۷۷). آزمون شناسی زبان. شیراز: انتشارات دانشگاه شیراز.
۶. دلاور، علی. (۱۳۷۷). مقدمه‌ای بر تحلیل عاملی. تهران: رشد.
۷. سیف، علی اکبر. (۱۳۸۲). روشهای ارزشیابی و اندازه‌گیری آموزشی. تهران: دوران.
۸. سیف، علی اکبر. (۱۳۷۵). مقدمه‌ای بر نظریه یادگیری. تهران: دانا.
۹. سیف، علی اکبر. (۱۳۹۰). سنجش فرآیند و فراورده‌های یادگیری. تهران: دوران.
۱۰. سلیمی‌زاده، محمد کاظم. (۱۳۷۷). آشنایی با روش‌های تجزیه و تحلیل سوالات و آزمون‌ها. تهران: سازمان سنجش کل کشور.
۱۱. شریفی، پاشا. (۱۳۹۳). اصول روان سنجی و روان آزمایی. تهران: رشد.
۱۲. شریفی، کاظم، و گنجی، اکبر. (۱۳۸۳). معرفی روش طراحی آزمون‌های هدفدار. مجله دانشکده ادبیات و علوم انسانی دانشگاه تهران، ۱۶، ۱۳۳-۱۱۸.
۱۳. شقاقی، فرهاد، و متانی، ندا. (۱۳۹۰). هنجاریابی آزمون بین‌المللی زبان فارسی. (پایان-نامه کارشناسی ارشد منتشر نشده). دانشگاه علامه طباطبایی.
۱۴. شیخ‌الاسلامی، عباس. (۱۳۷۸). بررسی روایی آزمون ورودی تحصیلات تکمیلی در ایران. (رساله‌ی دکتری منتشر نشده). دانشگاه علامه طباطبایی.
۱۵. قنسولی، بهزاد. (۱۳۸۹). طراحی و روایی سنجی آزمون بسندگی زبان فارسی. پژوهش زبان‌های خارجی، ۵۷، ۱۱۵-۱۲۹.
۱۶. قاسمی‌ورزنده، سجاد. (۱۳۸۴). بررسی روایی پیشگویی قسمت زبان عمومی آزمون-های ورود کارشناسی ارشد مجموعه زبان انگلیسی. (رساله دکتری منتشر نشده). دانشگاه علامه طباطبایی.
۱۷. کرمی، ابوالفضل. (۱۳۸۲). کاربرد آزمون‌های روانی. تهران: روان سنجی.
۱۸. گلپور، لیلا. (۱۳۹۴). طراحی و اعتبار بخشی آزمون استاندارد زبان فارسی بر پایه چهار مهارت زبانی. (پایان‌نامه کارشناسی ارشد منتشر نشده). دانشگاه پیام نور.
۱۹. گنجی، حمزه. (۱۳۷۱). روان شناس عمومی. تهران: پیام نور.

۲۰. موسوی، ثریا. (۱۳۷۹). بررسی راهکارهای نظری آزمون‌های بسندگی و تهیه یک آزمون بسندگی استانده زبان فارسی برای غیر فارسی زبانان. (پایان‌نامه کارشناسی ارشد منتشر نشده). دانشگاه شیراز.

۲۱. ممقانی، هانیه. (۱۳۹۰). ارزیابی روایی و پایایی آزمون وزارت علوم و تحقیقات با هدف دستیابی به یک آزمون مهارت برای زبان فارسی. (پایان‌نامه کارشناسی ارشد منتشر نشده). دانشگاه علامه طباطبائی.

۲۲. مومنی، منصور و قیومی، علی. (۱۳۸۹). تحلیل آماری با استفاده از SPSS. تهران: انتشارات مؤلف.

۲۳. هومن، حیدر علی. (۱۳۷۴). مبانی پژوهش در علوم رفتاری. تهران: پارسا.

۲۴. وبگاه مرکز زبان فارسی به غیر فارسی زبانان دانشگاه فردوسی مشهد. (۱۳۹۸). معرفی مرکز، بازیابی شده در ۱۹ اردیبهشت ۹۷ از:

۲۵. <https://ctpl.umac.ir>

۲۶. Bachman, F., & Palmer S. (۲۰۱۰). Language Testing in Practice. Oxford: Oxford

۲۷. University Press.

۲۸. Bachman, T. (۲۰۰۸). Current concepts in validity and reliability for psychometric instrument: Theory and application. The American j Med , (۲) ۱۱۹ ۱۶-۷.

۲۹. Brown, H. D. (۲۰۰۱). Language assessment principle and classroom practice. New York: Pearson Education Inc.

۳۰. Brown, H. D. (۲۰۰۳). Language assessment principle and classroom practice. White Plains, New York: Pearson Education Inc.

۳۱. Brown, J. D. (۲۰۰۸). Testing in Language Programs. New York: McGraw-Hill.

۳۲. Bryman, A. (۲۰۱۲). Social research methods. Oxford: Oxford University Press.

۳۳. Chastain, K. (۱۹۸۸). Developing second language skills. New York: Harcourt Brace Jovanovich Publishers.

۳۴. Craighead, W. E. & Nemeroff, C. B. (۲۰۰۴). The concise Corsini encyclopedia of psychology and behavioral science. Hoboken, New Jersey: Wiley and Sons

Publishers.

۳۵. Davis, A. (۱۹۹۰). Principles of language Testing. Oxford: Oxford University Press.

۳۶. Do rnyei, Z. (۲۰۰۷). Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies. Oxford: Oxford University Press.

۳۷. Farhady, H., Jafarpur, A., & Birajandi, P. (۲۰۱۷). Testing language skills: From theory to Practice. Tehran: SAMT.

۳۸. Fulcher, J. (۲۰۰۷). Language testing and assessment. New York: Taylor and Francis Group.

۳۹. Hughes, A. (۱۹۹۸). Testing for language teacher. London: Cambridge University Press.

۴۰. Klein- Braley, C. (۱۹۸۱). Practice and problems in language testing. Proceedings of the First International Language Testing Symposium of the IUS. Frankfurt: Peter Lang.

۴۱. Schmitt, N. (۲۰۱۲). An introduction to applied linguistics. England: Hodder Education.

۴۲. The University of Wesconsin Website, Retrieved on July ۲۰۲۰, ۲۸ from:  
<https://www.uwosh.edu/testing/faculty-information/test-scoring/score-report-interpretation/item-analysis/-/item-difficulty>

۴۳. The University of Washington Website, Retrieved on July ۲۰۲۰, ۲۸ from:  
<https://www.washington.edu/assessment/scanning-scoring/scoring/reports/item-analysis/>



پژوهشگاه علوم انسانی و مطالعات فرهنگی  
پرتال جامع علوم انسانی