

طراحی و روایی سنجی آزمون بسندگی زبان فارسی

بهزاد قنسولی*

دانشیار دانشگاه فردوسی مشهد، ایران

(تاریخ دریافت: ۸۸/۶/۱۰، تاریخ تصویب: ۸۹/۲/۱۲)

چکیده

توسعه دوره‌های آموزش فارسی به غیر فارسی‌زبانان و تدوین کتاب‌های مختلف آموزشی مرتبط در این زمینه، تحولی فرهنگی است که با وجود آثار مثبت آن کاستی‌هایی را نیز به همراه دارد. مهم‌ترین این نقص‌ها گستره آزمون‌سازی زبان و به ویژه نبود آزمون بسندگی زبان فارسی است. این مقاله به تعریف و اهمیت آزمون‌سازی زبان و ارتباط آن با دوره‌های آموزش زبان فارسی پرداخته است. در این پژوهش ضمن ساخت یک نمونه آزمون بسندگی زبان فارسی با تأکید بر آزمون‌های شنیدن، خواندن و دستور، روایی و پایایی آزمون سنجیده شد. برای روایی سنجی آزمون، از فرایندی دو مرحله‌ای، شامل تحلیل داده‌های بازنگری سه دانشجوی خارجی دوره دکتری زبان و ادبیات فارسی و بهره‌گیری از نظریه بحث در سنجش روایی بهره گرفته شد. پایایی آزمون نیز به شیوه آلفا کرونباخ محاسبه شد. نویسنده همچنین به دلیل ابعاد مهم و تأثیرگذار آزمون در قلمرو سیاسی و فرهنگی کشور، به تولید و روایی سنجی آزمون ملی بسندگی زبان فارسی توجه کرده است.

واژه‌های کلیدی: آزمون‌سازی، بسندگی، زبان فارسی، پایایی، روایی.

۱- مقدمه

در نیم قرن اخیر یکی از قلمروهای مهم در زمینه زبان‌آموزی که روز به روز به میزان اثربخشی و اهمیت آن در گستره‌های مرتبط با زبان توجه شده است، آزمون‌سازی است. این نگاه معطوف به دلایلی است. از یک سو اهمیت داده‌هایی، که موتور آن آزمون است، موجب تغییرات اساسی در مبانی نظری و شالوده‌های عملی رشته‌های مرتبط با زبان می‌شود. روشن است که اعتبار و راز ماندگاری یک نظریه، همواره در گرو پیوستگی نتایج برگرفته از به کارگیری آن نظریه در عمل است. آزمون‌سازی، همان حلقه واسطه میان نظریه و تغییرات میدانی است. یک چنین دیدگاه عمل‌مدار (Practice-oriented) در زبان‌شناسی کاربردی (Applied Linguistics) که در آن نتایج کمی و کیفی تحقیقات در خدمت تغییرات، اصلاح و در مواردی طرح نظریات جدید می‌گردد، نامأنوس نیست (دیویس ۲۰۰۱). زیرا این نتایج، شاخصی برای تعیین حقیقت محسوب می‌شود. از سوی دیگر، نقش آزمون در تحولات اجتماعی و تأثیر آن بر زندگی فردی حائز اهمیت است. آزمون‌های ملی که چارچوب‌های آموزش و یادگیری را به بخش‌های آموزش کشور تحمیل می‌کند، دارای تأثیرات اجتماعی است. این تأثیر در ادبیات رایج آزمون‌سازی، بازشویی (Wash back) نامیده می‌شود که اگر به صورت هدفمند برای اصلاح و بهبود بخش آموزش به کار گرفته شود، تأثیری مثبت بر روند آموزش و یادگیری خواهند داشت. علاوه بر این، آزمون‌ها در شکل‌گیری مبادلات اقتصادی و مالی یک جامعه نیز اثرگذارند. حجم سرمایه‌گذاری‌های خانواده‌ها و نقل و انتقال سرمایه در بخش‌های خصوصی مرتبط با دوره‌های یاد شده، سالانه افزون بر میلیاردها تومان می‌شود که متأسفانه آمار دقیقی از آن در دست نیست.

بی‌گمان یکی از ابعاد مهم و تأثیرگذار آزمون‌ها، قلمرو سیاسی- فرهنگی و شکل‌دهی هویت ملی هر کشور است. به عنوان مثال، آزمون بسندگی زبان انگلیسی (تافل) نماد قدرت زبانی کشورهای انگلوساکسون است که سالانه میلیون‌ها تن در سراسر جهان تنها پس از دریافت نمره لازم در این آزمون، مجوز عبور از دروازه‌های دانشگاه‌ها را به دست می‌آورند. استفاده ابزاری از آزمون بسندگی زبان انگلیسی که در ظاهر تنها کارکردی روان‌سنجی (Psychometric) دارد و هدف اصلی آن بر مدار تعیین میزان آمادگی داوطلبان دانشگاهی برای تحصیل در دانشگاه‌های انگلیسی زبان است، در لایه‌های ناپیدای آن خاصیتی تشویقی- ترغیبی دارد (براون ۲۰۰۴). زبان فارسی چندی است که با سیاست‌های تشویقی دولت، مراحل آغازین حرکتی نو را کلید زده است و اکنون در کشورهای مختلف جهان کرسی‌های زبان فارسی

فعال‌تر از گذشته شده‌اند. علاقه‌مندان به یادگیری فرهنگ، زبان و ادبیات فارسی برای تعیین سطح دانش زبان فارسی، نیازمند به آزمونی مانند تافل‌اند تا در سطوح مناسب آموزش قرار گیرند. متأسفانه جای این آزمون خالی است. در این مقاله نویسنده با بررسی جایگاه رشته آزمون‌سازی در علوم زبان‌شناختی، به اهمیت آزمون بسندگی فارسی پرداخته به مراحل نخستین طراحی، تدوین و روایی سنجی یک نمونه آزمون بسندگی فارسی توجه کرده است.

۲- مفهوم بسندگی کلی زبان چیست؟

در ادبیات آزمون‌سازی، تعاریف متعددی برای بسندگی کلی زبان طرح شده است که هر یک برگرفته از چارچوب نظری حاکم بر دوره‌های خاص مکتب زبان‌شناسی کاربردی است. برای نمونه در دوره حاکمیت ساختارگرایی (Structuralism) در زبان‌شناسی که متأثر از مکتب روانشناسی رفتارگرایی است، زبان و دانش کلی زبان تعریفی برگرفته از نوعی سازوکار شرطی شدن محض دارد. در این دیدگاه، توانایی کلی زبان را می‌توان در سابقه و شرایط یادگیری پیشین آن جستجو کرد. از این رو بسندگی کلی زبان با توجه به سطح و میزان آشنایی قبلی زبان‌آموز، تعیین می‌شود. در این دیدگاه، دانش زبان در حد مجموعه‌ای از مهارت‌ها محدود شده و ماهیت پیچیده زبان تنزل پیدا می‌کند (استرن ۱۹۸۴). راه حل دیگری که اکنون مقبولیت بیشتری در میان اهل نظر یافته است بر طراحی مدلی از دانش زبانی تأکید می‌کند که بر اساس دانش گویشور بومی زبان استوار است. سؤال اصلی در این مدل آن است که دانستن یک زبان و ویژگی‌های آن چیست. استرن چهار ویژگی در پاسخ به این سؤال مطرح می‌کند که عبارتند از:

- ۱- چیرگی بر صورت‌های زبان،
- ۲- چیرگی بر معانی شناختی، عاطفی، فرهنگی-اجتماعی،
- ۳- ظرفیت لازم در به کارگیری حداکثری از زبان برای ایجاد ارتباط،
- ۴- خلاقیت در استفاده از زبان.

با این وجود در بررسی تاریخچه تحولات بسندگی کلی زبان مباحث نظری، مقیاس‌های دسته‌بندی آزمون‌های استاندارد از مهم‌ترین رویکردهای تعریفی‌اند.

از دیدگاه نظری، عده‌ای مانند آلر (۱۹۷۹) بسندگی زبان را در چارچوب محتوای زبانی (Linguistic Context) آن تفسیر می‌کنند، بدان گونه که افزون بر مفاهیمی چون واژگان، دستور و آواشناسی، مفاهیم دیگری مانند معناشناسی، گفتمان و ویژگی‌های جامعه‌شناختی را

در آزمون بسندگی زبان دخیل می‌دانند. با این حال گروهی دیگر (استرن ۱۹۸۴) به نقش رفتاری زبان که در مهارت‌های زبانی تجلی پیدا می‌کند، بها داده و از این رو توانش زبانی را در چارچوب سنجش مهارت‌های خواندن، نوشتن، گفتن و شنیدن تعریف کرده‌اند. گروهی دیگر، بسندگی زبان را بر محور مقیاس‌های رتبه‌ای تعریف کرده‌اند. بر این اساس، رتبه‌ها تعاریف متفاوتی از بسندگی زبان را در بر می‌گیرند. به عنوان مثال مقیاس توانش زبانی FSI پنج طبقه‌بندی تعریفی برای بسندگی زبان ارائه داده است؛ مانند ۱- توانش محدود، ۲- توانش کاربردی محدود، ۳- توانش پایین، ۴- توانش خوب زبان و ۵- توانش دو زبانی یا بسیار خوب. معیار دسته‌بندی افراد بر اساس چارچوب استفاده کاربردی و ارتباطی زبان و نیز معیارهای زبان‌شناسی تعیین شده است.

بر اساس آنچه گفته شد، جدول زیر تصویری کلی از اجزای بسندگی زبان که نظریه‌پردازان به آن توجه دارند ارائه می‌دهد.

جدول شماره ۱- خلاصه نظریات مختلف در تفسیر بسندگی زبان برگرفته از استرن (۱۹۸۴)

نسبتاً انتزاعی					نسبتاً واقعی
مفهوم واحد	مفهوم دو بعدی	مفهوم سه بعدی	مفهوم چهار بعدی	مفهوم چند بعدی	
نظریه آلر یا توانایی انطباق‌گفتمان با بافت برون زبانی (Pragmatic expectancy Grammar)	نظریه کامینز مبنی بر توانش زبانی و توانش ارتباطی	نظریه کانالی و اسوین مبنی بر توانش دستوری، توانش زبانی - اجتماعی و توانش راهبردی	نظریه استرن مبنی بر چهار مهارت اصلی شنیدن، گفتن، خواندن و نوشتن	توانایی کاربرد عملکردی زبان توانایی آوایی، خطی، واژگانی و دستوری در ارتباط با مهارت‌های شنیدن، گفتن، خواندن و نوشتن کارگیری از زبان	

همان‌گونه که از جدول فوق پیداست، به هر اندازه از طیف انتزاعی به سوی مفاهیم چند بعدی حرکت کنیم، بسندگی زبان، مفهومی سنجش‌پذیرتر و واقعی‌تر به خود می‌گیرد. نکته

حائز اهمیت در این جدول، نگاه چند بعدی به سازه بسندگی زبان است. در آزمون‌های معتبر بسندگی زبان انگلیسی، نظیر آزمون‌های تافل و آیلتس، به مفاهیم چهار بعدی و چند بعدی توجه بیشتری شده است. در آزمون آیلتس از مقیاس رتبه‌بندی نه بخشی (Stanine) نیز استفاده شده است. با توجه به سیر تحولات آزمون‌های بسندگی با این وجود بحث‌های فراوانی در زمینه اجماع کلی برای تعریفی واحد از سازه بسندگی وجود دارد، اما گذشته از تفاوت نظرات، اینک در سطح جهانی شاهد اجرای آزمون‌های بسندگی معتبر برای تعیین توانایی‌های زبانی و میزان موفقیت آینده داوطلبان دانشگاهی در کشورهای انگلیسی زبان‌ایم. در نگاهی کلی هدف اصلی یک آزمون بسندگی زبان، دریافت اطمینان از کاربرد یک زبان برای اهداف علمی، به ویژه موفقیت داوطلبان خارجی در درک شنیداری، گفتاری، نوشتن و خواندن در زبان دوم است.

۳- اهمیت جایگاه آزمون بسندگی زبان در توسعه و تعمیق زبان فارسی

در سال‌های اخیر تلاش‌هایی در راستای توسعه زبان فارسی، به ویژه در سطح جهانی از راه ایجاد کرسی‌های زبان فارسی صورت گرفته است، اما این دگرگونی‌ها بیشتر تغییراتی آموزشی‌اند (نگاه کنید به انتشار رو به رشد کتاب‌های متعدد در آموزش زبان فارسی، به غیر فارسی زبانان مانند ضرغامیان، ۱۳۸۴؛ ضیاء حسینی، ۱۳۸۵؛ صفار مقدم، ۱۳۸۲). این موضوع با نگاهی انتقادی به یکی از کتاب‌های مطرح در این زمینه بررسی شده است، (قنسولی ۱۳۸۳). اما آنچه اکنون بیشتر نگران‌کننده است، غفلت غیر قابل قبول ما در درک نقش و اهمیت آزمون‌ها، به ویژه آزمون بسندگی زبان فارسی در فرایند توسعه زبان فارسی است. با این حال باید به این نکته توجه داشت که همه آزمون‌ها از یک درجه اهمیت و کاربرد واحد برخوردار نیستند. بعضی از آزمون‌ها دامنه‌ای محدود داشته و کاربردی به اصطلاح وابسته به سرفصل درس دارند؛ مانند آزمون‌های پیشرفت تحصیلی (Achievement Test) نتایج این آزمون‌ها تعمیم‌پذیر نبوده و بیشتر کاربردی موقعیتی و محلی (Local) دارند. در نقطه مقابل، آزمون بسندگی زبان، قرار دارد با کاربردی وسیع و پیشگو (Predictive). در بخش قبلی اجزای بسندگی مورد توجه قرار گرفت.

یکی از عوامل مهم در ایجاد آزمون‌های بسندگی زبان، جایگاه بین‌المللی زبان‌هاست. توسعه دانش و علوم مختلف در یک کشور، پیوندی نزدیک با تحکیم جایگاه بین‌المللی زبان آن کشور پدید می‌آورد. بی‌گمان اقبال قابل توجه فراگیرندگان زبان انگلیسی دستیابی به دانشی

است که در کشورهای انگلوساکسون انباشته شده است. این جایگاه بستر توسعه آزمون بسندگی تافل را ایجاد کرده است. در کشور ما پیشرفت علمی در بخش‌های پزشکی و علوم هسته‌ای و نیز جاذبه‌های معنوی و فرهنگی و ادبی ذخیره‌های علمی کشور محسوب می‌شوند و می‌توانند بستر لازم برای توسعه آزمون بسندگی فارسی را فراهم آورند. آزمون بسندگی حاوی سازوکارهای تأثیرگذار در بخش‌های آموزش و یادگیری زبان‌اند که از آن به عنوان تأثیر بازشویی یاد می‌شود. چنانچه آزمون بسندگی در راستای اهداف آموزش و یادگیری مؤثر زبان طراحی شود، آثار مثبتی در توسعه صحیح و علمی زبان فارسی بر جای خواهد گذاشت. به عنوان نمونه، آزمون اختصاصی زبان انگلیسی کنکور سراسری کشور تأثیری مثبت بر روند یادگیری داوطلبان نداشته است، زیرا بخش‌های محدودی از بسندگی زبان، یعنی خواندن، دستور و واژگان اهداف این آزمون‌اند و از این رو داوطلبان ورود به دانشگاه‌ها، تنها خود را برای این آزمون‌ها آماده می‌کنند و از این رو بخش‌های مهم‌تر یادگیری زبان که مهارت‌های نوشتن، شنیدن و گفتن را شامل می‌شود، از دایره زبان‌آموزی و یادگیری خارج می‌شوند. از این روست که در طراحی آزمون بسندگی زبان فارسی، می‌باید دقت لازم در گنجاندن اصولی مهارت‌های چهارگانه زبان بر اساس مفاهیم چهار بعدی و چند بعدی بسندگی که در جدول شماره ۱ به آن اشاره شد، صورت گیرد.

۴- مراحل تحقیق

۴-۱- طراحی و تدوین چارچوب کلی آزمون

برای شروع طراحی آزمون لازم است تا با بررسی مدل‌های اجزای بسندگی زبان یک مدل زبانی انتخاب گردد. بر مبنای جدول شماره ۱ مدل چند بعدی زبان که با واقعیت شرایط آزمون زبان تناسب دارد، انتخاب شد. در این مدل به ۲ مهارت اصلی زبان، مانند خواندن، شنیدن و نیز مهارت‌های دستور و واژگان توجه شده است. این مدل به مدل‌های استاندارد کنونی، مانند تافل و آیلتس به لحاظ شکلی شباهت دارد. به عنوان اقدامی جدید طراحی آزمون به ۲ مهارت اصلی خواندن و شنیدن محدود گردیده و مهارت‌های گفتن و نوشتن به دلیل محدودیت‌های تحقیق لحاظ نگردید. تعداد سؤالات آزمون به ترتیب ۴۰ سؤال چهار گزینه‌ای برای دستور، ۴۰ سؤال چهار گزینه‌ای برای واژگان، ۶ متن خواندن هر کدام با ۵ سؤال چهار گزینه‌ای و ۳۰ سؤال چهار گزینه‌ای برای مهارت شنیدن که به ۳ بخش گفتگوهای کوتاه،

متوسط و بلند تقسیم شده بودند.

۱-۱-۴- سؤالات خواندن

در انتخاب متن‌ها به طور نسبی به ویژگی‌های تناسب متن‌ها با سطح دانش و آگاهی آزمودنی‌های خارجی زبان توجه شد. بدین منظور متن‌ها از واژه‌های ادبی، اجتماعی، تاریخی و علمی تشکیل شده‌اند. در تعیین سطح دشواری متن‌ها، گرچه از فرمول‌های رایج سطح دشواری خواندن مانند سنجش فاگ (Fog reliabilities index) استفاده نشد، با این وجود طرح نظرات یک استاد زبان و ادبیات فارسی و یک استاد زبان‌شناسی به پالایش متن‌ها کمک شایانی کرد. میانگین تعداد واژگان ۳۵۰ واژه تعیین شد.

۱-۱-۲- سؤالات شنیدن

یکی از دشوارترین مراحل طراحی آزمون بسندگی زبان فارسی، طراحی موقعیت‌های بیانی متفاوت و متنوع برای آزمون بود. این موقعیت‌ها به ۳ دسته مکالمات کوتاه شامل مکالمه ۲ تن در حد بیان یک موضوع و پاسخ به آن، مکالمه متوسط که شامل تبادل و بیان اطلاعات در حد چند جمله توسط شخصیت‌های آن موقعیت و مکالمه بلند که منحصرأ متشکل از سخنرانی‌های علمی در حد ۳ دقیقه بود، تقسیم شدند.

۱-۱-۳- سؤالات ساختار

سؤالات ساختار به ۲ دسته ۲۰ سؤالی یعنی ۲۰ سؤال چهار گزینه‌ای که تنها حاوی ۱ گزینه درست و ۳ گزینه نادرست و ۲۰ سؤال چهار گزینه‌ای تشخیص خطا (Error identification) تقسیم شدند. هر بخش دارای راهکار خاص خود بود. سؤالات پیرامون واژگان دستوری (Function words) طراحی شدند. از این رو از مورد به پرسش گذاردن واژگان قاموسی (Lexical words) در تنه سؤالات پرهیز شد. سؤالات، قبل از انجام آزمون توسط دو استاد زبان و ادبیات فارسی نقد و بررسی شد و اصلاحات لازم صورت گرفت.

۵- نتایج آزمون

در این بخش به ۳ موضوع میانگین کل نمرات آزمودنی‌ها، میانگین نمرات آنها در بخش‌های مختلف آزمون، سنجش پایایی آزمون، سنجش روایی آزمون به دو شیوه تحلیل

کیفی مبتنی بر پروتکل‌های بازنگری و تحلیل بر مبنای نظریه بحث خواهیم پرداخت.

۵-۱- محاسبه میانگین‌ها

میانگین نمرات کسب شده دانش پژوهان خارجی در این آزمون در بخش‌های مختلف و نیز میانگین کل نمران آنها در جدول زیر محاسبه شد.

جدول شماره ۲. میانگین نمرات دانش پژوهان خارجی در بخش‌های مختلف آزمون

خواندن	دستور	شنیدن	
۳۰	۴۰	۳۰	تعداد سؤالات
۱۷	۱۹	۱۶	میانگین
۵۱			میانگین کل

میانگین نمرات دانش پژوهان به خوبی نشان می‌دهد که دانش پژوهان حدوداً به نیمی از سؤالات با موفقیت پاسخ داده‌اند. بهترین کارکرد در بخش خواندن و درک مفاهیم است. بخش دستور، دشوارترین بخش به نظر می‌رسد. میانگین کل نمرات، پایین‌تر از انتظار است. در بخش سنجش روایی آزمون، به دلایل این کاستی اشاره خواهد شد.

۵-۲- سنجش پایایی آزمون بسندگی

برای سنجش پایایی (Reliability) آزمون از فرمول آلفا کرونباخ استفاده شد و پایایی آزمون ۷۰٪ معین شد که نتیجه‌ای قابل قبول برای این آزمون تلقی می‌شود. به همین گونه آزمون فریدمن که مکمل سنجش پایایی این آزمون است، سطح معناداری آن را تأیید کرد.

۵-۳- سنجش روایی آزمون بسندگی

در نگاهی کلی، روایی یک آزمون را در شواهد و اسنادی می‌توان یافت که نشان دهنده ارتباط میان سؤالات آزمون و سازه مورد سنجش در یک زمینه خاص باشد. در این نگاه هر آزمون می‌بایست ویژگی‌های روایی مخصوص به خود را داشته باشد که ضرورتاً قابل تعمیم به آزمون‌های دیگر نیست (کین ۲۰۰۱). در نظریه سنجش روایی بر مبنای اصل بحث، کین

(۲۰۰۱) به نکات زیر در فرایند سنجش روایی توجه می‌کند:
تمام استنباط‌هایی که منجر به نتیجه‌گیری از نمره‌ها می‌شود باید به طور کامل و آشکار بیان شوند.

سنجش روایی مبتنی بر اصل بحث زمانی موجب تلفیق و همگونی تمامی داده‌های موجود و مرتبط برای استنباط و نتیجه‌گیری با یکدیگر می‌شود.
باید مفاهیم مشکل ساز این فرایند، با استفاده از شواهد موجود و تحقیقات جدید ارزیابی شوند.

اصل بحث با استفاده از شکلی قابل قبول تر مورد تجدید نظر قرار گرفته و یا به کلی مردود اعلام شود.

بر اساس آنچه که گفته شد برای سنجش روایی آزمون بسندگی فارسی از مدل فولچر و دیویدسون (۲۰۰۷) استفاده کردیم. این مدل که از چارچوبی مفهومی (Conceptual Frame) بر مبنای نظریه بحث (Argument) برخوردار است، اساساً روایی را فرایند بحث درباره معنای نمرات و کاربرد نمرات می‌داند. در این فرایند، باید اطلاعات بایسته را درباره معنای نمرات و کاربرد آن فراهم آورد. چارچوب مفهومی فولچر و دیویدسون در زیر چنین آمده است:



چارچوب اساسی فرایند بحث در سنجش روایی بر اساس مدل فولچر و دیویدسون (۲۰۰۷)

برای فهم بهتر این چارچوب، باید مفاهیم فوق تعریف شوند. اجزای مدل مذکور شامل داده‌های موجود، طرح ادعا، توجیه، داده‌های پشتیبانی و رد ادعاست. داده‌های موجود

(Grounds)، شامل حقایق، شواهد و داده‌های مرتبط است. طرح ادعا (Claim) نتیجه‌ی بحثی است که ما در جستجوی آنیم. توجیه (Warrant) پلی میان داده‌های مرتبط و ادعاست و توجیه‌کننده ادعای مبتنی بر داده‌های مرتبط است. داده‌های پشتیبانی (Backing) به هر داده‌ای که برای توجیه لازم است گفته می‌شود. داده‌ها شامل مدل‌های نظری، تحقیقات قبلی و داده‌های پشتیبانی است. رد ادعا (Rebutal) به داده‌هایی گفته می‌شود که در بخش توجیه نمی‌توان نبود آن را توجیه کرد.

اگر بخواهیم از این چارچوب مفهومی برای سنجش روایی آزمون بسندگی استفاده کنیم، باید گفت که داده‌های مرتبط در این آزمون به اطلاعات برگرفته از مهارت‌های مورد سنجش اشاره دارد. باید اذعان کرد از آنجایی که این تحقیق یکی از نخستین گام‌های برداشته شده برای تولید یک آزمون بسندگی زبان فارسی است، از این رو عاری از کاستی نیست. داده‌های موجود تنها به سنجش مهارت‌های شنیدن، خواندن و دستور محدود شده است و مهارت‌های گفتن و نوشتن در این آزمون گنجانیده نشده است. ادعای اصلی این آزمون، تولید محتوا برای سنجش دانش زبان فارسی در چارچوب مهارت‌های یاد شده در بالا است. از این رو، این آزمون، ادعایی برای سنجش مهارت‌های گفتن و نوشتن در زبان فارسی ارائه نمی‌دهد. توجیه اصلی این آزمون مربوط به داده‌هایی است که با ادعای سنجش مهارت‌های شنیدن، خواندن و دستور مرتبط است. در طراحی این آزمون، از مدل‌های موجود آزمون بسندگی به ویژه آزمون بسندگی زبان انگلیسی (TOEFL) برای داده‌های پشتیبانی آزمون بسندگی زبان فارسی استفاده شد. توجیه اصلی برای نبود مهارت‌های گفتن و نوشتن (rebuttal) در این آزمون، مربوط به دشواری سنجش عینی (objective) این مهارت‌هاست که در این تحقیق میسر نگردید. در سنجش روایی این آزمون به ۲ اقدام تحلیل کیفی آزمون و تحلیل مفهومی آزمون بر مبنای نظریه بحث اشاره شد. نگارنده به خوبی آگاه است که برای سنجش روایی سازه‌ای (Construct Validity) این آزمون، می‌بایست از تحلیل عامل (Factor Analysis) نیز استفاده می‌شد. علت اصلی نبود تحلیل عامل در سنجش روایی، تعداد کم آزمودنی‌های (۱۶ تن) شرکت‌کننده در این تحقیق است.

۱-۳-۵- تحلیل کیفی روایی آزمون بسندگی

به منظور تحلیل کیفی روایی (Validity) آزمون و تعیین ارتباط سؤالات آن با سازه‌های دانش زبان، ناگزیر از به کارگیری روش کیفی آن شدیم. دلیل عمده تعیین روایی به روش

کیفی، دسترسی محدود به دانش پژوهان خارجی زبان است که در ایران به تحصیل مشغول می‌باشند. اگر چه تعداد بیشتری از این دانش پژوهان را نمی‌توان در مؤسسات آموزش عالی و غیر دولتی بخش‌های مختلف کشور یافت، و از آنجایی که تنها تعداد محدودی در گسترهٔ تدوین این تحقیق یعنی شهر مشهد تحصیل می‌نمودند، از این رو تنها ۱۳ دانش‌پژوه دختر با ملیت‌های عربی و پاکسانی که در مدارس علوم اسلامی این شهر به تحصیل اشتغال داشتند و همچنین ۳ دانشجوی دکتری کره‌ای و عراقی دانشگاه فردوسی مشهد مورد پرسش قرار گرفتند. آنان ابتدا به سؤالات آزمون پاسخ دادند و پس از آزمون به روش بازنگری (retrospection) نظر آنها در زمینهٔ سطح دشواری مادهٔ آزمون‌ها، تناسب مادهٔ آزمون‌ها با دانش زمینهٔ آنان، وجود ابهام در سؤالات آزمون‌ها، ارتباط سؤالات آزمون با سطح دانش آنان و هر آنچه که آنان به عنوان نقطهٔ ضعف و کاستی سؤالات یافته بودند، مورد سؤال قرار گرفت. اصلاحات به طور عمده در بخش‌های شنیدن، خواندن و ساختار صورت گرفت.

بخشی از پروتکل بازنگری رویکرد دانشجوی اهل کرهٔ جنوبی و دانشجوی دورهٔ دکتری زبان و ادبیات فارسی دانشگاه فردوسی مشهد به شرح زیر است:

«این امتحان برای من سخت بود. ولی این امتحان در سطح پیشرفته بود. برای شنیدن، من تقریباً صحبت کردن را فهمیدم. اما موقع جواب دادن کمی به مشکل برخورددم. همه جمله را نخوانده بودم. چون وقت خیلی کم بود. من معمولاً یک یا دو جمله را خواندم. آن موقع سؤال دیگر شروع شد. من نمی‌توانستم جواب درست را انتخاب کنم..... دوم امتحان، برای من خیلی سخت بود که حرف اضافه، زمان، دستور زبان چه فرقی دارند. سوم امتحان، متینها (متن‌ها) جلب (جالب) بودند، اما سؤال‌ها برای من سخت است. مشکل این بود که تقریباً کلمه‌ها را نمی‌فهمیدم، کلمه‌های جدید خیلی زیاد بودند».

بخشی از پروتکل بازنگری رویکرد دانشجوی اهل کرهٔ جنوبی و دانشجوی دکتری زبان و ادبیات فارسی دانشگاه فردوسی مشهد در این بخش ارایه شده است.

فصله‌های (فصل‌های) موضوع و موضوع بعد کم بود. نمی‌توانستم آن چهار تا جمله‌ها را بخوانم. و قط (وقت) کم بود. من هنوز جمله‌ها را داشتم می‌خواندم ولی موضوع دیگر شروع می‌شد.

کشورمان از سوی چپ تا راست می‌نویسیم. من از این شیو (شیوه) عادت کرده‌ام. ولی ترتیب پاسخ الف ب پ ها با ترتیب ما مخالفت (مخالف) بود. پس گاهی این را اشتباه (اشتباه) می‌کردم، پاک می‌کردم و دوباره تغییر می‌دادم. به نظر من برای مثل ما پاسخ نامه دیگر هم درست بکنید.

در مورد بخش دوم (دستور) برای من بیش از همه کلمه‌ها سخت بود. نمی‌توانستم حل کنم.

در مورد بخش سوم این هم به من خیلی سخت بود. بعضی از کلمه‌ها اولین بار دیدم.

بخشی از پروتکل بازننگری رویکرد دانشجوی اهل عراق و دانشجوی دوره دکتری زبان و ادبیات فارسی دانشگاه فردوسی مشهد به شرح زیر می‌باشد.

«آن آزمایش برای ما آوردند. فعلاً خیلی ساخت (سخت) نیست ولی چم (کم) مشکل داشت. آن مشکل بود: ۱- زمان آزمایش شنیدن و نوشتن خیلی کم بود. ۲- آن صحبت ما شنیدیم بسیار سرعت زیاد بود. ۳- به نظر (نظر) من بهترین راه الان آن آزمایش حتماً تکرار با سابقه اطلاع و زمان زیاد می‌کنند.»

در بررسی پروتکل‌های بازننگری ۳ متن، دانشجویان فوق به ۲ نکته اساسی توجه داشته است. الف: زمان کم برای پاسخ به سؤالات آزمون. اولین دانشجوی اهل کره جنوبی، علت اصلی مشکل بودن سؤالات را در نداشتن وقت کافی برای پاسخ به سؤالات ذکر می‌کند. دومین دانشجوی اهل کره جنوبی نیز به صراحت به کمبود وقت اشاره می‌کند. دانشجوی اهل عراق نیز سرعت بالای شنیدن و کمبود زمان کافی را مشکل اصلی آزمون می‌داند. ب: دشواری واژگان. سخت بودن واژگان در بخش‌های مختلف آزمون به ویژه در بخش درک نوشتاری و متون خواندن یکی از مهم‌ترین دشواری‌های آزمودنی‌های فوق قلمداد شده است. برای اولین دانشجوی اهل کره جنوبی، بخش دستور و خواندن دارای واژگان دشوار است. دومین دانشجوی اهل کره جنوبی نیز دشواری پاسخ دادن به سؤالات خواندن را در دشواری واژگان این متن‌ها و سؤالات آن می‌داند. از موارد دیگری که به عنوان عامل دشواری به آن اشاره شده است، ضعف دانش دستوری (پروتکل اولین دانشجوی اهل کره جنوبی) و تفاوت شیوه نگارش میان زبان کره‌ای و فارسی (پروتکل دومین دانشجوی اهل کره جنوبی) است. هیچ یک

از دانشجویان خارجی به عدم تناسب ماده آزمون‌ها با دانش زمینه و وجود ابهام در سؤالات آزمون که بیشتر به عنوان موارد تحلیل آزمون گوشزد شده بود، ایرادی نگرفتند. اولین دانشجوی اهل کره جنوبی به جالب بودن متن‌ها اشاره می‌کند. بر اساس پروتکل‌های بازنگری دانشجویان فوق، تغییرات لازم در ماده آزمون‌ها صورت گرفت. بدین ترتیب که بسیاری از واژگان دشوار به واژگان نسبتاً ساده تغییر یافتند. زمان بیشتری به خواندن متون داده شد تا دانشجویان بتوانند با سهولت بیشتری به ماده آزمون‌ها پاسخ دهند.

نتایج بخش کمی این پژوهش تفاوت‌هایی با نتایج به دست آمده در مطالعه مقدماتی موسوی (۱۳۷۹) دارد. در حالی که نتایج این تحقیق نشان می‌دهد دانشجویان خارجی در بخش خواندن و درک مفاهیم بهترین عملکرد را دارند و بخش دستور دشوارترین بخش برای آنان است، نتایج پژوهش موسوی نشان می‌دهد که آزمودنی‌ها در بخش خواندن از کمترین مهارت و در بخش دستور از بالاترین مهارت برخوردارند. یکی از دلایل عمده دشواری آزمودنی‌ها در بخش دستور به میزان همبستگی پایین آن با کل آزمون اصلی است که در مطالعه وی با صراحت به آن اشاره شده است. سابقه یادگیری آزمودنی‌ها و نیز دوره‌های آموزشی قبلی آنان که در داخل و خارج کشور طراحی شده است، می‌تواند در این رابطه تأثیرگذار باشد.

در سال‌های اخیر از داده‌های بیانی به نحو بارزی در فرایند روایی سنجی آزمون‌ها استفاده شده است (گرین ۱۹۹۸). به عنوان مثال، نوریس (گرین ۱۹۹۸) از روش برون‌فکنی اندیشه (thinking-aloud) برای سنجش روایی سؤال‌های چند گزینه‌ای بهره گرفت. نکته قابل توجه در کاربرد داده‌های بیانی برای سنجش روایی آزمون‌ها تبیین دقیق‌تر ویژگی‌های مهارت‌هایی است که آزمون‌گر آنها را در ساخت آزمون مورد نظر دارد. بازخوانی فرایندهای شناختی آزمودنی‌ها در انجام یک آزمون ما را در تعیین انطباق این فرایندها و آنچه که به عنوان فرایندهای واقعی پیش‌بینی می‌کنیم، یاری می‌دهد. به این ترتیب می‌توان تشخیص داد که آیا آزمون طراحی شده با اهداف واقعی آن سنخیت و ارتباط دارد یا نه و این اگر چه تمام تلاش ما برای سنجش روایی یک آزمون نیست، اما تصویری دیگر از فرایند روایی سنجی را ایجاد می‌کند که از طریق روش‌های کمی امکان پذیر نیست.

۶- نتیجه

در این تحقیق، گام نخستین برای تولید و روایی سنجی آزمون ملی بسندگی زبان فارسی

برداشته شد. به طور مسلم در مراحل بعدی لازم است که اصلاحات صورت گرفته در این آزمون با شرکت تعداد بیشتری از دانش پژوهان خارجی زبان مورد بررسی و واکاوی قرار گیرد. نباید این حقیقت را از خاطر زدود که آزمون بسندگی زبان علاوه بر نقش‌های علمی و نشانگرهای پیشگویی آن از امتیازی فرهنگی و حتی سیاسی برخوردار است. این آزمون به عنوان نماد قدرت ملی و بین‌المللی زبان فارسی تلقی می‌شود و در بیشتر موارد، انتقال دهنده ارزش‌های فرهنگی - زبانی یک ملت است. زبان تأثیرگذار و فرهنگ مدار فارسی با وجود تمهیدات چند ساله اخیر در بخش آموزش و گسترش محسوس آن در میان علاقه‌مندان به یادگیری آن، فاقد ابزار سنجش بسندگی زبان است. تعیین مناسب سطح دانش زبان فارسی و آگاهی از میزان موفقیت داوطلبان غیر فارسی زبان در تحصیلات دانشگاهی و حوزوی در ایران و خارج از آن نیازی ضروری است که به ارتقای جایگاه علمی، فرهنگی و سیاسی ایران کمک می‌کند. برگزاری آزمون بسندگی زبان فارسی در زمان‌های مختلف، در نمایندگی‌های جمهوری اسلامی ایران در کشورهای مختلف، موجب تحریکات بیشتر فرهنگی و سیاسی و در نتیجه تقویت جایگاه ایران اسلامی در عرصه بین‌المللی می‌شود، امری که در سند چشم‌انداز توسعه بیست ساله کشور مورد توجه قرار گرفته است. در طراحی آزمون بسندگی زبان فارسی می‌توان از مدل‌های کنونی در زبان‌های برجسته و جهانی دنیا الهام گرفت و آن را با ساختار ملی و فرهنگی و ادبی زبان فارسی انطباق داد. بدون شک در این مسیر به دلیل ماهیت بین‌رشته‌ای حاکم بر آزمون بسندگی زبان تنها استادان زبان فارسی نیستند که باید نقش‌آفرینی نمایند هر چند که در مراحل از طراحی از نظرات آنان نیز باید سود جست، بلکه باید زمینه اصلی فعالیت در قلمرو زبان‌شناسی کاربردی تعریف گردد.

Bibliography

- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practice*. New York: Longman.
- Davis, A. (2001). "The Logic of Testing Languages for Specific Purposes". *Language Testing*. 18(2), 133-147.
- Fulcher, G., and Davidson, F. (2007). *Language Testing and Assessment*. Routledge.
- Ghonsooly, B. (1383/2004). "Naghdi Bar Sathe Doshvarie Va Methodologie Ketabe Amuzeshe Farsi Be Gheire Farsi Zaban" (Analysis of Readability and Methodology of Teaching Farsi to Non-native Speakers of Farsi). *Journal of Faculty of Letters of Ferdowsi University of Mashhad*. 48, 93-103.

Green, A. (1998). *Verbal Protocol Analysis in Language Testing Research*. Cambridge: Cambridge University Press.

Kane, M. T. (2001). 'Current concerns in validity theory.' *Journal of Educational Measurement*. 38, 4, 319-341.

Mousavi, S. (1379/ 2000). Barresiye Rahkarhayeh Nazariye Azmoonhaye Basandeghi va Tahiyeyeh Yek Azmoone Basandeghye Estandeye Zabane Farsi Baraye Gheire Farsi Zaban (A Study of the Approaches to Proficiency Tests and Preparing a Standard Persian Proficiency Test for the Speakers of Other Languages). Unpublished MA dissertation, Shiraz University.

Oller, J. W. (1979). *Language Tests at School*. London: Longman.

Safar Moghaddam, A. (1382/2003). *Farsi Omumi 1: Sakhtarhaye Payeh* (General Farsi 1: Basic Structures). Tehran: Council for the Development of Persian Language and Literature.

Stern, H. H. (1984). *Fundamental Concepts of Language Teaching*. Oxford: Oxford university press.

Zarghamian, M. (1384/2005). *Doreye, Amuzeshe Zabane Farsi* (A Course in Teaching Persian Language). Tehran: Council for Persian Language and Literature.

Zia Hoseini, S. M. (1385/2006). *Raveshe Tadrise Zabane Farsi Be Farsi Amuzane Khareji* (Methodology of Teaching Persian to Foreign Language Learners of Persain). Tehran: Sokhan Publications.