

ضرورت رویکرد نظام‌مند به آزمون‌سازی در زبان فارسی: مقایسه‌ی برآوردهای مدرّسان و واقعیت آماری

سید اکبر جلیلی

عضو هیأت علمی دانشگاه بین‌المللی امام خمینی (ره)

مهدی خدادادیان

عضو هیأت علمی دانشگاه بین‌المللی امام خمینی (ره)

چکیده

در مرحله‌ی «پیش‌آزمون»، که یکی از مراحل مهم در آزمون‌سازی می‌باشد، پرسش‌های آزمون به صورت آزمایشی روی تعدادی زبان‌آموز که از لحاظ زبانی، هم‌سطح با آزمون‌دهندگان اصلی هستند، آزمایش می‌شوند و از رهگذر بررسی‌های آماری و تحلیل پاسخ‌های آزمون‌دهندگان آزمایشی، اطلاعات آماری همچون ضریب سهولت و ضریب تمیز پرسش‌ها استخراج می‌گردد. همچنین، پرسش‌های نامعتبر و دارای اشکال نیز مشخص گشته، بازبینی و اصلاح می‌شوند. پژوهش حاضر به پاسخ به این سوال می‌پردازد: آیا می‌توان در زمینه‌ی تشخیص میزان سهولت پرسش‌های آزمون، به معیاری به نام «برآوردهای مدرّسان» تکیه کرد؟ به عبارتی دیگر، برآوردهای مدرّسان زبان فارسی از سختی/آسانی پرسش‌ها (یعنی حدس‌ها، قضاوت‌ها یا پیش‌بینی‌ها) تا چه اندازه با ضرایب سهولت/دشواری برآمده از تحلیل آماری پاسخ‌های آزمون‌دهندگان (واقعیت آماری) انطباق و همخوانی دارند و می‌تواند ما را از طی مراحل پیش‌آزمون و تحلیل آماری بی‌نیاز سازند. در حقیقت، این پژوهش می‌خواهد ضرورت یا عدم ضرورت پیش‌آزمون را -که در رویکرد علمی-آماری و نظام‌مند نوین به آزمون‌سازی مطرح می‌شود- از طریق مقایسه‌ی «برآوردهای مدرّسان» و «واقعیت آماری آزمون» نشان دهد. به این منظور، ۱۵۱ پاسخنامه، مربوط به یکی از آزمون‌های پایان‌دوره‌ی خواندن و درک مطلب زبان فارسی سطح پایه (مقدماتی)، در مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) بررسی شد و پس از محاسبه‌ی ضریب سهولت هر یک از ۲۸ پرسش این آزمون، این ضرایب سهولت، با برآوردهای ۳۰ نفر از مدرّسان مقایسه شدند. نتایج نشان داد که صرفاً با اتکا به تجربه یا برآوردهای شمی مدرّسان نمی‌توان گفت که پرسش‌های یک آزمون، سخت، آسان یا ایده‌آل هستند و این نشان از ضرورت نهادینه‌سازی رویکرد علمی-آماری و نظام‌مند به آزمون‌سازی، در زمینه‌ی ارزشیابی و سنجش دانش زبانی فارسی‌آموزان دارد.

کلیدواژه‌ها: آزمون، برآوردهای مدرّسان، پیش‌آزمون، ضریب سهولت

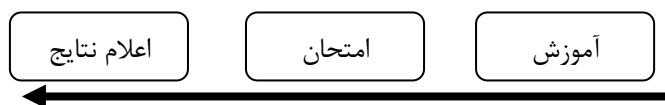
۱. مقدمه

در رویکرد سنتی و غیرعلمی به مقوله‌ی ارزیابی و ارزشیابی آموزشی، به‌ویژه در زمینه‌ی آموزش زبان، از «آزمون» صرفاً به‌عنوان ابزاری برای تعیین نمره‌ی یادگیرندگان یک زبان یا شرکت‌کنندگان در یک دوره‌ی آموزش زبان استفاده می‌شود. در این رویکرد، آنچه که از آن با نام «امتحان» یاد می‌شود، در مدت کوتاهی و بدون صرف دقت چندانی طراحی گردیده و در اختیار شرکت‌کنندگان قرار می‌گیرد. پس از امتحان نیز قضاوت‌هایی احساسی، همراه با معیارهایی احساسی صورت می‌گیرد که نمونه‌ای از تجلی این قضاوت‌ها را می‌توان در گزاره‌هایی مشاهده نمود که از سوی امتحان‌دهندگان یا دست‌اندرکاران مطرح می‌گردد: «سوال‌ها خیلی سخت/آسان بودند»، «معلوم نیست این سوال‌ها/واژه‌ها مربوط به کتابی باشند که ما خوانده‌ایم!»، «من اصلاً فلان سوال را نفهمیدم»، «جواب فلان سوال کاملاً در خود سوال مشخص بود و نیازی به خواندن ادامه‌ی سوال نبود» یا «این امتحان خیلی سخت/آسان بود و همه در آن رد/قبول شدند». همچنین، در رویکرد سنتی، امتحان‌گیرندگان نیز پاسخ‌هایی به اعتراضات امتحان‌دهندگان می‌دهند که این پاسخ‌ها خود نشان از عدم وجود معیارها و استانداردهایی برای طراحی امتحان هستند. در مجموع، در یک نگاه سنتی، امتحان صرفاً ابزاری برای تعیین نمره یا رتبه است و خاصیتی جز اعلام قبولی یا عدم قبولی افراد شرکت‌کننده ندارد، اما در رویکرد نظام‌مند و علمی، «آزمون» دیگر یک ابزار صرف برای قبولی یا عدم قبولی نیست و ویژگی‌های دیگری دارد. می‌توان تفاوت‌های «آزمون» در رویکرد علمی و نظام‌مند نوین، و «امتحان» در رویکرد سنتی را این‌گونه خلاصه نمود:

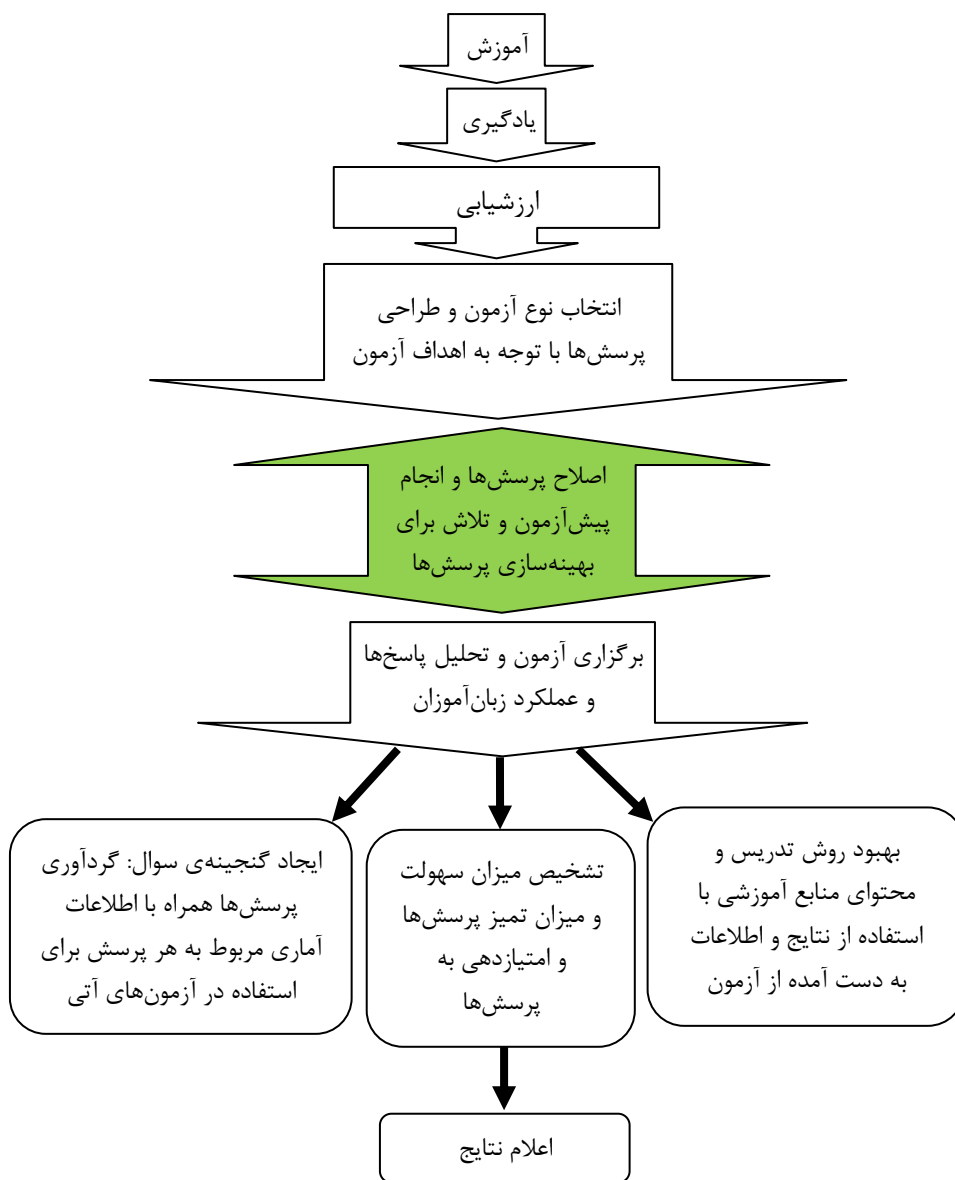
«آزمون» در رویکرد علمی و نظام‌مند	«امتحان» در رویکرد سنتی
«آزمون»، یکی از حلقه‌های زنجیره‌ی آموزش و یادگیری است.	«امتحان»، پایانی بر آموزش است.
آزمون معتبر، محتوا و چگونگی تدریس مدرّسان، و محتوا و چگونگی یادگیری یادگیرندگان را تحت تأثیر قرار می‌دهد (براون، ۲۰۱۰: ۳۸).	آزمون و نتایج آن تأثیری بر محتوا و روش تدریس یا یادگیری ندارد.
«پیش‌آزمون» برگزار می‌گردد تا پرسش‌ها پیش از برگزاری آزمون اصلی، استاندارد شوند. پرسش استاندارد، تعاریف و ویژگی‌های خاص خودش را دارد.	سوال‌ها بدون بررسی و بلافاصله پس از طراحی، در اختیار امتحان‌دهندگان قرار می‌گیرند و با قضاوت‌های شخصی و شمی طراحان ارزیابی می‌شوند.
به نوع آزمون و هدف آزمون توجه می‌شود.	نوع امتحان‌ها تفاوتی ندارد، و هدف امتحان صرفاً اعلام قبولی/ردی است.

«آزمون» در رویکرد علمی و نظام‌مند	«امتحان» در رویکرد سنتی
روى پرسش‌ها تحلیل آماری صورت می‌گیرد و پرسش‌های خیلی سخت، ایده‌آل، و آسان مشخص می‌شوند.	هیچ تحلیلی بر روی سوال‌ها انجام نمی‌شود و تنها قبول شدن یا نشدن امتحان‌دهندگان مشخص می‌شود.
پاسخ‌های زبان‌آموزان به هر پرسش، مورد تجزیه و تحلیل قرار می‌گیرد؛ میزان سهولت/دشواری و میزان تمیز هر پرسش مشخص می‌شود و براساس آن، برای هر پرسش امتیازی کمتر یا بیشتر در نظر گرفته می‌شود.	در صورتی که امتحان سخت بوده باشد و اعتراضات زیاد باشد، نمره‌ها اصطلاحاً «روی نمودار» می‌روند و در صورتی که سوال‌ها برای شرکت‌کنندگان در امتحان، آسان بوده باشند، تنها به منزله‌ی باهوش بودن امتحان‌دهندگان و رضایت از کار آموزش‌دهندگان تلقی می‌شود.
«آزمون‌سازی» فرایندی پیچیده، دقیق و هدفمند است. از تحلیل پرسش‌ها و نتایج، برای بهینه‌سازی آموزش استفاده می‌شود.	طراحی سوال، کاری ساده، بی‌زحمت و تک‌بُعدی است و از نتایج امتحان، به دلیل عدم تحلیل‌شان بهره‌ای برده نمی‌شود.
به ابعاد مختلف مانند ابعاد روانشناختی و اجتماعی آزمون و نتایج آن توجه می‌گردد (مک‌نامارا و زوور، ۲۰۰۶: ۲).	به ابعاد و تأثیرات مختلف آزمون و نتایج آن توجه نمی‌گردد.

بنابراین، می‌توان رویکرد علمی و نظام‌مند به آزمون‌های زبان با رویکرد سنتی را، به شکل‌های زیر (شکل‌های ۱ و ۲) نشان داد:



شکل ۱. امتحان در نگاه خطی و تک‌بُعدیِ رویکرد سنتی



شکل ۲. آزمون در رویکرد علمی و نظام‌مند

در راستای همین تحولات در حوزه‌ی آزمون‌های زبان، کتاب‌های موجود در حوزه‌ی آزمون‌سازی^۱، مانند جعفرپور (۱۳۷۵)، فرهادی و همکاران (۲۰۰۷)، کومب و همکاران (۲۰۰۷) و براون (۲۰۱۰)، آزمون‌های زبان را براساس اهداف‌شان به چند نوع تقسیم می‌کنند و همچنین، ویژگی‌های هر آزمون را شرح می‌دهند. در این کتاب‌ها، سازوکارهایی برای طراحی انواع پرسش‌های آزمون‌های زبان مطرح، و نشان داده می‌شود که هر پرسش تا چه اندازه می‌تواند هوشمندانه و مفید طراحی گردد و پس از برگزاری هر آزمون، چه تحلیل‌هایی بر روی هر پرسش صورت می‌گیرد تا ارزیابی دقیقی از توانایی‌های زبانی آزمون‌دهندگان صورت پذیرد.

بیرجندی و مصلی‌نژاد (۲۰۱۰: ۱۱) از مقوله‌ای به نام «تأثیرات و پیامدهای آزمون»^۲ سخن می‌گویند و اعتقاد دارند که اطلاعات برآمده از نتایج آزمون را می‌توان در ارزیابی جنبه‌های گوناگونی از فرایند آموزش (تعیین میزان واقع‌گرایانه بودن اهداف آموزشی، مناسب بودن روش‌ها و منابع آموزش، کیفیت یادگیری و...) به کار گرفت؛ آنها (همان‌جا) عقیده دارند که

«نتایج آزمون نه تنها نقاط ضعف در یادگیری هر یک از یادگیرندگان را مشخص می‌سازند، بلکه در صورت مدّ نظر قرار دادن نتایج به صورت کلی، می‌توانند نقاط ضعف در آموزش را نیز روشن سازند؛ مثلاً وقتی که اکثر یادگیرندگان در مورد یک پرسش مشابه، ضعیف عمل می‌کنند، نمی‌توان گفت که این تقصیر آنها است، بلکه این بیشتر متوجه تدریس است؛ شاید مدرس در تلاش برای آموزش مواردی بوده که غیر قابل‌دستیابی توسط یادگیرندگان بوده‌اند، یا از روش‌هایی ناکارآمد برای اهداف آموزشی‌اش بهره برده است (البته در صورتی که آزمون مناسب بوده باشد). پاسخ‌های یادگیرندگان به پرسش‌های آزمون و مباحثه‌ی پس‌آزمونی^۳ در خصوص نتایج، می‌تواند سرنخ‌هایی را برای پی بردن به منبع و منشأ دشواری پرسش‌ها فراهم آورد».

گیتر (۱۹۹۵؛ نقل از همان: ۱۳) «پیامد آزمون» را به‌عنوان «تأثیر آزمون بر آموزش و یادگیری» تعریف می‌کند و کوهن (۱۹۹۴؛ نقل از همان: ۱۴) آن را «چگونگی تأثیر ابزارهای ارزشیابی بر رویه‌ها و باورهای آموزشی» می‌داند.

آزمودن، یک عمل ظریف و پیچیده است. ظریف است، چون قرار است تصمیماتی بر مبنای نتایج آزمون گرفته شوند. یعنی تصمیم‌گیری مناسب، نیاز به اطلاعات دقیق دارد و داده‌های دقیق، از طریق آزمودن و ارزشیابی دقیق به دست می‌آیند (بیرجندی و مصلی‌نژاد، ۲۰۱۰: ۱۹)؛ پیچیده است، چون شامل رشته‌های مطالعاتی متفاوتی می‌شود. اگر طراح آزمون از مکاتب روان‌شناختی، جامعه‌شناختی، روش‌شناختی، و فلسفی

۱. اصطلاحات «آزمون‌سازی» و «آزمودن» در این پژوهش معادل "testing" به کار رفته‌اند.

۲. testing backwash effects

۳. post-test discussion

آگاه نباشد، دستیابی به سنجش‌های دقیق، تقریباً غیرممکن خواهد بود. مثلاً در آزمون، یک طراح آزمون تلاش می‌کند که پیشرفت یادگیری آزمون‌دهنده را ارزشیابی کند؛ روان‌شناسی، مسئولیت علمی تعریف یادگیری و بررسی متغیرهای گوناگون مؤثر در یادگیری را بر عهده دارد. جامعه‌شناسی نیز علمی است که جامعه‌ی کلاس را تعریف می‌کند. همین‌طور روش‌شناسی، معین‌کننده‌ی چارچوب‌های آموزش و یادگیری است. گذشته از اینها، زبان‌شناسی نیز در آزمون [مهارت‌های] زبان نقش دارد؛ زیرا زبان‌شناسی، زبان را به عنوان موضوعی که قرار است آموخته شود، تعریف می‌کند. بنابراین، آزمون واقعاً ظریف و پیچیده است. آزمون باید به عنوان یک مرحله‌ی طبیعی در فرایندهای آموزش و یادگیری مد نظر قرار گیرد. کارشناسان اعتقاد دارند که آموزش و آزمون، دو روی یک سکه هستند. یعنی آنها بسیار درهم تنیده‌اند (همان: ۲۰ و ۲۱).

۲. بیان مسأله و ضرورت پژوهش

آنچه که موضوع این پژوهش می‌باشد، مسأله‌ی «پیش‌آزمون»^۱ است؛ موسوی (۱۹۹۷: ۱۰۴) «پیش‌آزمون» را این‌گونه تعریف می‌کند: «آزمایش یک آزمون که به‌تازگی طراحی شده، اما هنوز به‌طور کامل شکل نگرفته است. آزمون تحت طراحی، می‌تواند براساس نتایج تجزیه و تحلیل پرسش‌ها که از نتایج پیش‌آزمون به دست آمده، بازبینی و اصلاح شود». همان‌گونه که در شکل (۲) نیز مشاهده شد، این مرحله به صورت یک مرحله‌ی دوسویه و رفت و برگشتی است و معنای این دوسویگی آن است که با هر بار پیش‌آزمون، آزمون‌ساز به دستاوردهایی ناآشنا می‌گردد؛ مثلاً: (۱) پرسش‌های استاندارد را بازشناسی می‌کند؛ (۲) پرسش‌های بسیار سخت یا بسیار آسان را، که نمی‌توانند تمایز میان زبان‌آموزان ضعیف و قوی را نشان دهند، شناسایی می‌نماید؛ (۳) با توجه به تعداد انتخاب هر گزینه در یک پرسش چندگزینه‌ای یا با توجه به محتوای پاسخ‌ها برای پرسش‌های نوشتاری/تولیدی، دلایل احتمالی را که موجب سختی یا آسانی یک پرسش شده‌اند، کشف می‌نماید و (۴) مجدداً پس از رفع موارد یاد شده، پیش‌آزمونی دیگر را برگزار می‌نماید تا دوباره به استانداردهای پرسش‌ها نزدیک‌تر گردد. در واقع، مهم‌ترین تفاوت آزمون در رویکرد علمی نوین با رویکرد سنتی را می‌توان همین مرحله‌ی «پیش‌آزمون» دانست.

در همین خصوص، آدرسون و همکاران (۱۹۹۵؛ نقل از آدرسون، ۲۰۰۵: ۱۶۸) انگاره‌ای چندمرحله‌ای را برای آزمون‌سازی معرفی می‌کنند که پنجمین مرحله‌اش «پیش‌آزمون، تحلیل نتایج، و اصلاح آزمون» نام دارد. حتی کسانی که صحبتی از پیش‌آزمون نکرده‌اند (مانند بکمن و پالمر، ۱۹۹۶) نیز، هرچند سه بخش

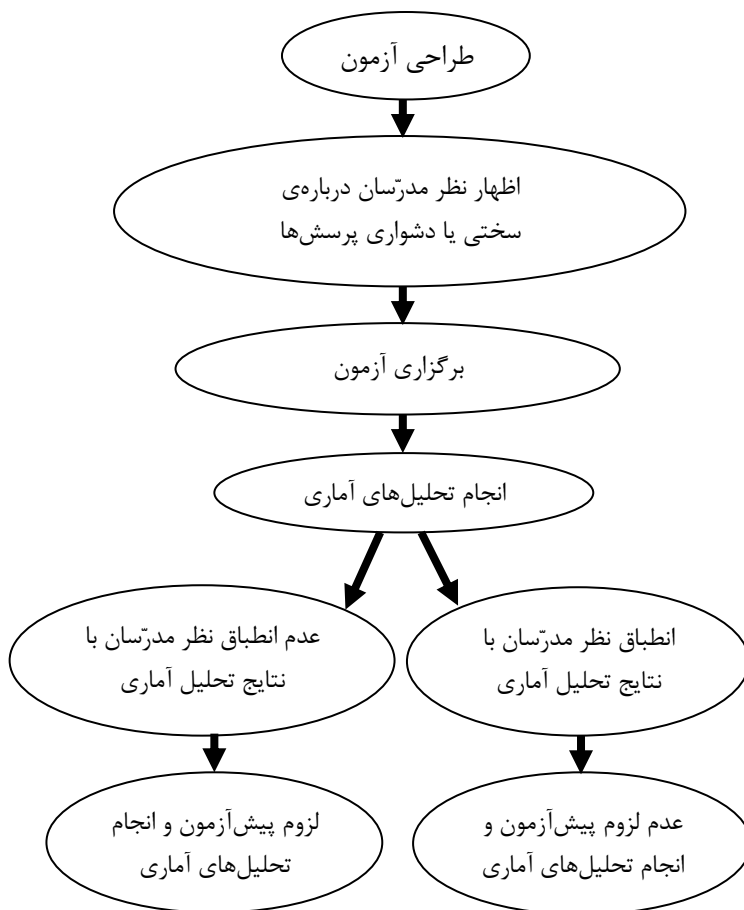
^۱. pre-test

اصلی، شامل «طراحی»، «عملیاتی‌سازی»^۱ و «برگزاری و اجرا»^۲، برای آزمون‌سازی برمی‌شمارند، اعتقاد دارند که بخش سوم، شامل مواردی همچون گردآوری بازخوردها، تحلیل آزمون و... است (نقل از آلدرسون، ۲۰۰۵: ۱۶۸). هاگس (۲۰۰۳: ۵۸) نیز مراحل پنجم و ششم طراحی و برگزاری آزمون را به «انجام پیش‌آزمون بر روی افراد مشابه با آزمون‌دهندگان اصلی و تجزیه و تحلیل نتایج آن و انجام تغییرات ضروری در آزمون» اختصاص می‌دهد. در حقیقت، تحلیل آزمون‌ها چه پیش از آزمون اصلی و چه پس از آن، اهمیت دارد.

براون (۲۰۰۳: ۶۷) در بحث درباره‌ی «آزمون استاندارد شده»^۳ اذعان می‌دارد که یک آزمون استاندارد شده‌ی خوب، محصول فرایند کاملی از تحقیق و توسعه‌ی تجربی است که رویه‌ها و روندهای استاندارد را برای برگزاری و نمره‌دهی تعیین می‌نماید. او در ادامه (ص: ۷۶) در خصوص مراحل طراحی آزمون تافل^۴ به کارهای مختلفی اشاره می‌کند که بر روی پرسش‌های مختلف از لحاظ علمی و آماری انجام می‌گیرند تا محتوای نهایی آزمون مورد نظر شکل بگیرد.

اما متأسفانه آنچه در حوزه‌ی آزمون‌سازی در آموزش زبان فارسی به غیر فارسی‌زبانان مشاهده می‌شود، برخلاف روند مطرح شده در رویکرد علمی نوین به آزمون‌سازی است؛ یعنی عدم لزوم پیش‌آزمون و استخراج اطلاعات آماری، و اکتفای صرف به نظرات مدرّسان و طراحان آزمون. بنابراین، مسأله‌ی پژوهش حاضر این است که آیا می‌توان صرفاً به نظرات بازبینی‌کنندگان آزمون، که معمولاً مدرّسان هستند، اکتفا کرد و لزوم مقوله‌ای به نام «پیش‌آزمون» یا تحلیل نتایج پس از آزمون را رد کرد؟ به زبانی ساده‌تر، اگر ما مجموعه‌ای از پرسش‌ها را طراحی کنیم، پس از یک یا چند بار بازبینی پرسش‌ها توسط مدرّسان، آیا می‌توانیم قاطعانه بگوییم که هر پرسش، ایده‌آل است یا نیست؟ اگر پاسخ این سوال مثبت باشد و مدرّسان بتوانند دقیقاً همان چیزی را بگویند که تحلیل‌های آماری پیش‌آزمون یا آزمون اصلی می‌گویند، پس می‌توان گفت که نیازی به تحلیل آماری یا برگزاری پیش‌آزمون وجود ندارد و می‌توانیم نظرات مدرّسان را، معیاری برای تعیین میزان سهولت یا دشواری پرسش‌های آزمون بدانیم و براساس همین پیش‌بینی‌ها یا حدس‌های مدرّسان، می‌توانیم به هر پرسش، قبل یا پس از آزمون امتیاز (وزن) بدهیم و نتایج عملکرد زبان‌آموزان را اعلام نماییم. مطالب فوق را می‌توان به شکل نمودار زیر نیز نشان داد:

1. operationalisation
2. administration
3. standardized test
4. TOEFL



شکل ۳. نمودار بیان مسأله‌ی پژوهش

به همین منظور، نگارندگان این پژوهش سعی خواهند نمود تا با انجام یک مقایسه میان «نتایج حاصل از تحلیل آماری پرسش‌های یک دوره آزمون زبان فارسی در سطح پایه» با «برآوردها (قضاوت‌ها، پیش‌بینی‌ها یا حدس‌ها)ی مدرّسان درباره‌ی سختی یا آسانی پرسش‌ها»، ضرورت یا عدم ضرورت وجود مرحله‌ای به نام پیش‌آزمون و نیز ضرورت یا عدم ضرورت رویکرد علمی-آماري نوین و نظام‌مند به آزمون‌سازی را نشان دهند؛ در نتیجه، سوال و فرضیه‌ی پژوهش این‌گونه مطرح می‌شود:

سوال: آیا برآوردها (پیش‌بینی و حدس) مدرّسان در خصوص میزان دشواری، آسانی، یا ایده‌آل بودن پرسش‌های یک آزمون، با واقعیت آن آزمون، که برآمده از تحلیل‌های آماری بر روی پاسخ‌های زبان‌آموزان است، انطباق دارد؟

فرضیه: پیش‌بینی و حدس مدرّسان در خصوص میزان دشواری، آسانی، یا ایده‌آل بودن پرسش‌های یک آزمون، با واقعیت آن آزمون، که برآمده از تحلیل‌های آماری بر روی پاسخ‌های زبان‌آموزان است، انطباق ندارد.

۳. تنگنای پژوهش

محدودیت و تنگنای پژوهش این است که متغیرهای موجود در پژوهش از یک سنخ نیستند؛ متغیر اول (ضرایب سهولت که با توجه به پاسخ‌های فارسی‌آموزان و با استفاده از فرمول ضریب سهولت استخراج شده)، طیفی از اعداد یا درصدهای بین صفر تا یک است، اما متغیر دوم (پیش‌بینی یا حدس مدرّسان)، طیفی از واژگان است که به هر پرسش اختصاص داده شده است. به عبارتی دیگر، وقتی یک مدرّس، گزینه‌ی سخت یا آسان را برای یک پرسش انتخاب کرده است، نمی‌توان گفت که منظور او از این توصیف، دقیقاً همان درصد یا عددی است که نتایج آماری به آن اشاره دارد. مثلاً وقتی یک مدرّس می‌گوید که فلان پرسش «خیلی سخت» است، نمی‌توانیم بگوییم که منظور آن مدرّس از «خیلی سخت»، همان ضریب سهولت میان ۱۸٪ تا ۳۸٪ است که در پیوستار موجود در شکل (۴) (در بخش ۱،۲،۵) طبق عددهای به دست آمده از آمار وجود دارد. در واقع، مدرّس، یک توصیف ذهنی و انتزاعی را ارائه می‌دهد که می‌تواند معیاری غیر از اعداد و ارقام واقعی داشته باشد؛ معیارهایی مانند در نظر گرفتن سطح زبانی زبان‌آموزانی که خودش با آنها کلاس داشته یا روش تدریس یا محتوای منابعی که خودش تجربه‌ی تدریس آنها را داشته است و ...

بنابراین، در این پژوهش با دو طیف توصیفی سر و کار داریم؛ یکی، طیفی از توصیف‌هایی که براساس آمارها و فرمول ضریب سهولت بیان شده در منابع آزمون‌سازی است و ما در پیوستار موجود در شکل (۴) آنها را به مجموعه‌ای از اعداد اختصاص داده‌ایم و طیف دیگر، توصیف‌هایی که کاملاً شمی، ذهنی و انتزاعی هستند که هرچند ممکن است با واقعیت‌های آماری نیز تصادفاً انطباق داشته باشند، اما به هر حال، نمی‌توان گفت که این دو مقوله از یک سنخ هستند و بین مثلاً مفهوم «سخت» از نظر مدرّسان و مفهوم «سخت» از نگاه آمار، انطباق وجود دارد و این، مانعی برای بهره‌گیری از فرمول‌های آمار استنباطی در این پژوهش می‌شود.

شاید خالی از لطف نباشد که این مسأله را با بیان نکاتی میان‌رشته‌ای روشن‌تر سازیم. نخست آن که شاید این پژوهش، به‌نوعی، از نوع پژوهش‌های اثبات‌گرایی و تجربه‌گرایی محسوب شود که به دنبال کشف

حقیقت از راه آزمایش و تجربه و مشاهده است^۱ و برای تعیین سطح دشواری یا آسانی آزمون‌ها به توصیف‌هایی انتزاعی و بدون اساس علمی-آماری، مانند «سخت» یا «آسان» اکتفا نمی‌کند.

دوم آن که شاید بتوان گفت -بدون قصد تأیید یا رد مباحث فلسفی در حوزه‌ی دانش فلسفه- توصیف‌هایی که مدرسان در این پژوهش به پرسش‌های آزمون اختصاص داده‌اند، صرفاً واژه‌هایی هستند که از دید هر شخصی می‌توانند معانی متفاوتی داشته باشند. مثلاً «سخت» در نظر یک مدرس، یک تعریف دارد و در نظر مدرس دیگر تعریفی دیگر. بنابراین، برای رسیدن به تعاریف واحد، نیاز به پژوهش‌های میدانی و مشاهدات عینی با معیارهای مشخص داریم.^۲

مثالی از دنیای پزشکی ارائه می‌کنیم که شاید مسأله را روشن‌تر سازد. شاید پزشکان براساس آزمایش و تجربه دریافته باشند که اگر گلبول‌های سفید انسان که سیستم ایمنی بدن را شکل می‌دهند، در تعدادهای معینی باشند، طیفی از حالات جسمی خیلی خوب، خوب، بد، خیلی بد، و وخیم را نشان می‌دهند. مثلاً برای توصیف کسی که گلبول‌های سفید در هر میکرولیتر از خونش به ۵۰۰ عدد رسیده، شاید بگویند اوضاعش خیلی وخیم است و به فلان بیماری دچار شده، و در توصیف کسی که تعداد گلبول‌هایش مثلاً ۲۰۰۰ عدد باشد، می‌گویند حالش خوب نیست. در این حالت، اگر یک نفر بالای سر یک بیمار برود، و با دیدن وضعیت او بگوید حالش خیلی بد است یا خیلی وخیم است، آیا می‌توان حرف او را دال بر این دانست که دقیقاً تعداد گلبول‌های سفید آن بیمار فلان مقدار معین است؟

با توجه به مطالب فوق، می‌توان این‌گونه نتیجه گرفت که نمی‌شود یک متغیر اسمی با ۳۰ نظر از سوی مدرسان را، که هر کدام، یک توصیف شمی و انتزاعی هستند، با متغیر دیگر که اعدادی عینی و قابل اندازه‌گیری هستند، در قالب فرمول‌های موجود در آمار استنباطی مقایسه نمود. بنابراین، به دلیل این محدودیت، در پژوهش حاضر، امکان بهره‌گیری از آمار استنباطی وجود ندارد.

۱. چپمن (۱۳۸۴: ۱۵۰) در وصف اثبات‌گرایان این‌گونه می‌گوید: «هدف اثبات‌گرایان منطقی این بود که زبانی مناسب برای بیان مشاهدات علمی ایجاد کنند و در پی شناسایی گزاره‌های معنادار علمی و تمایز آنها از گزاره‌های بی‌معنا بودند».

۲. همان‌طور که «ویتگنشتاین و پیروانش در انگلستان و آمریکا [نیز] فقر فلسفه‌ی زبان را آشکارا نشان داده‌اند و اعتقاد دارند که واژه‌ها یا کلمات همانند افزارهایی هستند (همانند افزارهای موجود در جعبه‌ی ابزار، مثل چکش، اره، پیچ‌گوشی، خط‌کش، میخ، و پیچ‌ها)؛ حال همان‌گونه که کاربرد این چیزها مختلف است، به همان‌سان، کاربرد واژه‌ها نیز گوناگون است و...» (بوخنسکی، ۱۳۷۹). ویتگنشتاین بر نظریه‌ی کاربردی یا ابزاری زبان تأکید دارد. وی با بیان نظریه‌ی «بازی‌های زبانی» در پژوهش‌های فلسفی چنین عنوان می‌کرد که معنای کلمات را از استعمال آنها در موقعیت‌های مختلف می‌توان فهمید و کلمات، بیرون از هر متن و موقعیتی فاقد معنا هستند و حد مشترکی ندارند (آلوانی مگرد، ۱۳۸۵).

۴. پیشینه‌ی پژوهش

در زمینه‌ی میزان سهولت پرسش‌های آزمون‌های زبان می‌توان به پژوهش نوروززاده (۱۳۹۰) اشاره کرد که به کاربرد آزمون‌سازی انطباقی در تعیین سطح زبان انگلیسی دانشجویان پرداخته است؛ وی با استفاده از ضریب سهولت پرسش‌ها به طراحی یک آزمون انطباقی رایانه‌ای پرداخته است که می‌تواند سطح زبانی دانشجویان زبان انگلیسی را تشخیص دهد.

در حوزه‌ی ارزشیابی و آزمون‌سازی در زبان فارسی، با رویکرد آموزش زبان فارسی به غیر فارسی‌زبانان نیز چند پژوهش انجام شده است که به‌عنوان نمونه می‌توان از این پژوهش‌ها یاد کرد: موسوی (۱۳۷۹) در زمینه‌ی طراحی آزمون بسندگی فارسی؛ جلیلی (۱۳۹۰) در زمینه‌ی طراحی آزمون مهارتی فارسی (آمفا)؛ ممقانی (۱۳۹۰) در زمینه‌ی دستیابی به یک آزمون مهارت برای زبان فارسی. اما از آن‌جا که موضوع این پژوهش‌ها ارتباط مستقیمی با موضوع پژوهش حاضر ندارد، از توضیح مبسوط‌شان خودداری می‌شود.

۵. مبانی نظری

در این بخش، در خصوص مراحل تهیه‌ی آزمون، و واکاوی ماده‌ها از طریق سه شاخص سهولت/دشواری پرسش‌ها، میزان فرق‌نمایی یک پرسش (به دلیل ارتباط مستقیم این شاخص با سختی/آسانی یک پرسش) و چگونگی توزیع گزینه‌های مربوط به یک پرسش چهارگزینه‌ای (از این لحاظ که تأثیر مستقیمی بر انتخاب گزینه‌ها توسط زبان‌آموزان دارد و می‌تواند از عوامل سختی/آسانی یک پرسش باشد) سخن گفته خواهد شد. لازم به ذکر است که از بیان مباحثی همچون «روایی» و «پایایی» که مربوط به ارزیابی درونی و محتوایی یک آزمون هستند و در انجام این پژوهش نیز نقشی ندارند، اجتناب می‌گردد.

۱.۵. مراحل تهیه‌ی آزمون

جعفرپور (۱۳۷۵: ۳۷-۴۴) برای تهیه‌ی آزمون، چهار مرحله متصور است. این چهار مرحله به شرح ذیل می‌باشند:

۱) طرح آزمون^۱: آماده کردن طرح آزمون (توصیف هدف آزمون، آماده‌سازی رئوس مطالب آزمون، گزینش نوع ماده‌ها (سوال‌ها)، تعیین درجه‌ی دشواری آزمون و تهیه‌ی دستورالعمل‌های بخش‌های مختلف آزمون)

^۱. test plan

۲) نوشتن ماده‌ها: طراحی سوالات آزمون با رعایت مواردی همچون: «طراحی سوال به تعداد کافی»، «هر ماده یک نکته‌ی مشخص را اندازه‌گیری کند و نه بیشتر»، «وظیفه‌ی آزمون‌دهنده در برابر هر ماده کاملاً مشخص شود» و...

۳) بازبینی ماده‌ها: پس از نوشتن ماده‌های آزمون، آنها باید مورد بازبینی دقیق قرار گیرند و بهتر است این کار یا با فاصله‌ی زمانی انجام شود یا توسط شخص دیگری، تا خطاها و کاستی‌های احتمالی یافت و اصلاح شوند. پس از این کار، آزمونگر باید از میان ماده‌های تأیید شده، بهترین‌ها را برگزیند و آنها را برحسب ملاحظات مورد نظر خود، مرتب کند. آزمون تهیه شده در این مرحله را آزمون آزمایشی می‌گویند. یعنی این آزمون باید «پیش‌آزمایی» شود. به همین دلیل، این آزمون را «پیش‌آزمون» نیز می‌گویند.

۴) پیش‌آزمایی: پس از این که آزمون آزمایشی مهیا شد، لازم است آن را با تعدادی آزمون‌دهنده که مشخصاتی همانند آزمون‌دهنده‌های اصلی دارند، آزمایش کرد. هدف پیش‌آزمایی، واکاوی کمی‌کارکرد ماده‌های آزمون با این گروه نمونه است. در واکاوی ماده‌ها^۱، پاسخ‌های داده شده به هر ماده (پرسش) مورد بررسی قرار می‌گیرد. واکاوی ماده‌ها را می‌شود پس از استفاده از آزمون در دفعات بعد نیز به‌کرات انجام داد تا از این راه کیفیت ماده‌ها و در نتیجه، کیفیت آزمون را بهبود بخشید.

در واقع، با توجه به مطالب فوق که از جعفرپور (۱۳۷۵) نقل شد، می‌توان چنین برداشت کرد که اگر پیش‌آزمون و واکاوی ماده‌ها قبل از آزمون اصلی انجام نشود، خود آزمون اصلی، حکم همان پیش‌آزمون را خواهد داشت و این از لحاظ اصول آزمون‌سازی چندان مطلوب به نظر نمی‌رسد. جعفرپور (۱۳۷۵: ۴۴) واکاوی ماده‌ها را روشی تشخیصی برای دستیابی به چهار هدف عنوان می‌کند: (۱) شناسایی ماده‌های ضعیف کاستی‌مند، (۲) شناسایی گزینه‌هایی که توجه هیچ آزمون‌دهنده‌ای را به خود جلب نکرده‌اند، (۳) تعیین درجه‌ی آسانی یا دشواری ماده‌ها و تعیین این که ماده‌ها تا چه حد بین آزمون‌دهندگان دارای دانسته‌های کمتر و بیشتر، تمیز می‌دهد و (۴) شناسایی کاستی‌های دستورالعمل‌ها و تعیین زمان مناسب برای برگزاری آزمون.

۲.۵. واکاوی ماده‌ها

واکاوی پاسخ‌های داده شده به هر پرسش، چه در مرحله‌ی پیش‌آزمون صورت گیرد و چه پس از آزمون، براساس سه شاخص انجام می‌شود. پر واضح است که منظور ما از این سخن، شیوه‌ی طراحی پرسش‌ها و اصول طراحی آنها نیست، بلکه منظور ما مجموعه مواردی است که از آنها اصطلاحاً با عنوان «تحلیل

^۱. item analysis

پرسش» یا «تحلیل آیتم» یاد می‌شود. در حقیقت، پس از برگزاری آزمون، با استفاده از پاسخ‌های آزمون‌دهندگان، می‌توان مشخص نمود که خود پرسش‌ها، و نه پاسخ‌ها، چه ویژگی‌هایی دارند. سیف (۱۳۸۷: ۳۰۷) اعتقاد دارد که «هدف از تحلیل سوال‌های آزمون، واریسی تک‌تک سوال‌ها و تعیین میزان دقت و نارسایی‌های آنها است. در تحلیل سوال‌های آزمون، نقاط قوت و ضعف یک آزمون و کیفیت همه‌ی سوال‌های آن تعیین می‌شود».

این تجزیه و تحلیل، همان‌گونه که در شکل (۲) نیز نشان داده شده است، می‌تواند اطلاعات و دستاوردهای بسیار مفیدی در زمینه‌های مختلف آموزش، یادگیری، و تألیف منابع درسی داشته باشد. سیف (همان‌جا) می‌گوید:

«لازم است معلمان پس از اجرای هر آزمون، سوال‌های آن را تحلیل کنند و با استفاده از نتایج حاصل، به تجدید نظر در آزمون و بهبود کیفیت سوال‌ها برای استفاده‌های بعدی اقدام نمایند... برای تهیه‌ی یک آزمون خوب، تحلیل پاسخ‌های آزمون‌شوندگان و بررسی تجربی آنها ضروری است. تحلیل پاسخ‌های آزمون‌شوندگان همچنین، اطلاعات تشخیصی لازم را برای بررسی کیفیت یادگیری دانش‌آموزان و مشکلات آموزشی معلمان فراهم می‌آورد».

بنابراین، پس از آن که یک آزمون برگزار می‌شود و تصحیح می‌گردد، مطلوب است که کارایی و تأثیر پرسش‌های آن، مورد ارزیابی قرار گیرد؛ این کار با بررسی پاسخ‌های زبان‌آموزان به هر پرسش صورت می‌پذیرد. (بیرجندی و مصلی‌نژاد، ۲۰۱۰: ۱۴۱).

کتاب‌های مختلفی در حوزه‌ی آزمون‌سازی از نویسندگانی مانند جعفرپور (۱۳۷۵)، سیف (۱۳۸۷)، فرهادی و همکاران (۲۰۰۷)، و بیرجندی و مصلی‌نژاد (۲۰۱۰) سه شاخص را برای هر پرسش معرفی می‌کنند و برای هر شاخص، تعریفی ارائه می‌دهند و شرایط آن را توصیف می‌نمایند. این سه شاخص عبارت هستند از «میزان سهولت»، «میزان تمیز (افتراق‌نمایی)»، و «توزیع گزینه‌ها». در ادامه، به اختصار به شرح این سه شاخص و نحوه‌ی محاسبه‌ی آنها از منابع مذکور و همچنین براون (۲۰۰۸) می‌پردازیم.

۱.۲.۵. میزان سهولت^۱

یک پرسش می‌تواند میزانی از سختی یا آسانی را داشته باشد؛ این سختی یا آسانی را تنها می‌توان از عملکرد زبان‌آموزان در آزمون اصلی یا پیش‌آزمون دریافت. یعنی آنچه که برای طراحان آزمون می‌تواند روشن سازد که هر پرسش تا چه اندازه سخت، آسان، یا ایده‌آل بوده است، پاسخ‌هایی است که زبان‌آموزان به هر پرسش می‌دهند. مسلماً یک پرسش کاملاً آسان، همان اندازه بی‌ارزش است که یک پرسش خیلی

^۱. facility

سخت؛ بدین معنی که وجود پرسش خیلی آسانی که تقریباً همه‌ی زبان‌آموزان، حتی ضعیف‌ترین‌شان، به آن پاسخ صحیح داده‌اند، مانند این است که به همه‌ی زبان‌آموزان یک نمره داده شده است و وجود یک پرسش کاملاً سخت نیز مانند آن است که از همه‌ی آنها، بدون استثنا یک نمره کم شده است و حتی قوی‌ترین زبان‌آموز هم به آن پرسش، پاسخ صحیح داده است. برای محاسبه‌ی «ضریب سهولت»^۱ یک پرسش، از فرمول زیر (تعداد پاسخ‌های درست تقسیم بر تعداد کل پاسخ‌ها) استفاده می‌شود:

$$\text{ضریب سهولت} = \frac{\text{تعداد پاسخ‌های صحیح کل آزمون‌دهندگان به یک پرسش}}{\text{تعداد آزمون‌دهندگان}}$$

$$IF = \frac{\sum C}{N}$$

ایده‌آل‌ترین وضعیت برای یک پرسش آن است که ضریب سهولتش ۵۰٪ (۰/۵۰) باشد. در این حالت، نیمی از آزمون‌دهندگان به پرسش مورد نظر، پاسخ درست داده‌اند و نیمی پاسخ اشتباه، که این می‌تواند بهترین وضعیت برای یک پرسش از لحاظ ضریب تمیز که در ادامه توضیح خواهیم داد، باشد. گاهی نیز به جای ضریب سهولت از اصطلاح «ضریب دشواری»^۲ استفاده می‌شود که در این صورت از فرمول زیر بهره می‌بریم:

$$\text{ضریب دشواری} = 1 - \text{ضریب سهولت}^3$$

بیشتر منابع آزمون‌سازی مانند فرهادی و همکاران (۲۰۰۷) و بیرجندی و مصلی‌نژاد (۲۰۱۰) عنوان کرده‌اند که ضریب سهولت قابل قبول و مطلوب برای یک پرسش، بین ۳۷٪ (۰/۳۷) تا ۶۳٪ (۰/۶۳) خواهد بود و هر چه این ضریب به صفر (سختی مطلق) نزدیک‌تر باشد، پرسش سخت‌تر است و هر چه به یک (آسانی مطلق) نزدیک‌تر باشد، پرسش آسان‌تر خواهد بود.

بنابر مطالب فوق، نگارندگان، درجه‌بندی میزان سهولت یک پرسش را بر روی یک پیوستار (شکل ۴) نشان داده‌اند. در واقع، این پیوستار نشان می‌دهد که از لحاظ علم آزمون‌سازی، هر یک از واژه‌های «سخت»، «آسان»، «ایده‌آل» و... با چه ضریب سهولتی تطابق دارند.

1. facility index

2. difficulty index

3. difficulty index = 1 - IF (Item Facility)

واقعاً سخت!	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعاً آسان!
از ۰/۱۸ تا ۰/۱۰۰	از ۰/۱۹ تا ۰/۳۶	از ۰/۳۷ تا ۰/۴۴	از ۰/۴۵ تا ۰/۵۵	از ۰/۵۶ تا ۰/۶۳	از ۰/۶۴ تا ۰/۸۲	از ۰/۸۳ تا ۰/۱۰۰

شکل ۴. پیوستار انطباق ضرایب سهولت با توصیف‌های واژگانی و کیفی

ضریب سهولت یک پرسش، عددی است که از محاسبه‌ی پاسخ‌های زبان‌آموزان و به‌طور کلی، از عملکرد آنها در آزمون حاصل می‌شود. بنابراین، همان‌گونه که در شکل (۲) ترسیم شده است، یکی از فواید تحلیل‌های آماری این است که مشخصات هر پرسش از لحاظ شاخص‌های سهولت، تمیز، و توزیع گزینه‌ها استخراج می‌گردد و در بانک سوال بایگانی می‌شود تا در آزمون‌های بعدی نیز انتخاب پرسش به‌راحتی صورت پذیرد. در نتیجه، لازم است که مشخصات آزمون‌دهندگان از لحاظ ملیت، سن، جنس، ضریب هوشی و... نیز به مشخصات سوال افزوده شود؛ زیرا ممکن است یک پرسش سخت برای عده‌ای از افراد معین، برای عده‌ای دیگر با شرایطی متفاوت، خیلی آسان باشد. پس دشواری یا آسانی یک آزمون، چیزی است که با توجه به پاسخ‌های زبان‌آموزان در پیش‌آزمون‌ها یا آزمون اصلی استخراج می‌شود و هیچ پرسشی نمی‌تواند یک برجسته‌ی قطعی سخت یا آسان (برای هر فردی با هر شرایط و پیشینه‌ی زبانی) داشته باشد.

فرهادی و همکاران (۲۰۰۷: ۱۵۶) ساختار سؤالات^۱ و نحوه‌ی آرایش آنها و پاسخ‌های صحیح را نیز از عوامل مؤثر بر روایی^۲ آزمون می‌دانند و اعتقاد دارند که سؤال‌هایی که خوب تنظیم نشده‌اند و یا دارای ابهام هستند، موجب ناروایی آزمون می‌شوند، چون چنین سؤالاتی مانع از بروز استعداد واقعی آزمون‌دهنده می‌شوند؛ اگر یک آزمون‌دهنده سؤالی را که بد تنظیم شده از دست بدهد، معلوم نخواهد شد که واقعاً جواب آن را نمی‌دانسته یا به دلیل نقصی که در سؤال بوده نتوانسته آن را پاسخ دهد. همچنین، تنظیم سؤالات آزمون باید به‌صورت «از ساده به مشکل» باشد. در غیر این صورت، آزمون‌دهنده ممکن است در همان آغاز، دلسرد و ناامید شود و بدین ترتیب، نمرات ارزش و روایی نداشته باشند.

۲،۲،۵. میزان تمیز (فرق‌نمایی)^۳

معمولاً یک آزمون می‌خواهد تفاوت میان افرادِ بادانش و کم‌دانش را نشان دهد و کسانی را به مراحل بعدی بفرستد که نسبت به دیگران، از لحاظ دانش، برتری دارند. البته این ویژگی همه‌ی آزمون‌ها نیست و گاهی هدف آزمون ممکن است انتظار دیگری را به وجود آورد. مثلاً گاهی ما نمی‌خواهیم که افراد با یکدیگر رقابت

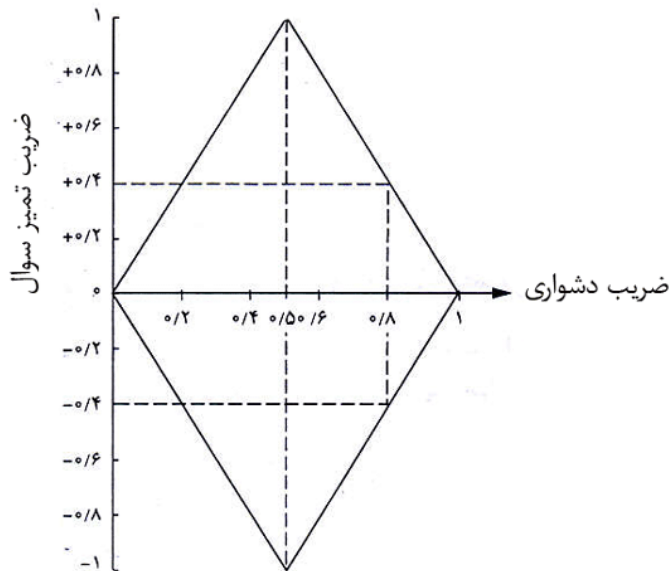
1. structure of the items

2. validity

3. discrimination

کنند و تنها می‌خواهیم هر کس که به تعداد مشخصی از پرسش‌ها پاسخ صحیح داد، به‌عنوان فرد قبول یا پذیرفته شده معرفی شود. در این حالت، صرفاً به میزان سهولت پرسش‌های موجود در بانک سوال، که میزان سهولت‌شان را داریم، مراجعه می‌کنیم و به فراخور هدف آزمون، افراد آزمون‌دهنده و شرایطشان پرسش‌هایی را انتخاب می‌کنیم. اما اگر به هر دلیلی بخواهیم آزمون ما تمایز میان ضعیف و قوی، باهوش و کم‌هوش را نشان دهد، ناچار خواهیم بود که در پیش‌آزمون یا آزمون، به محاسبه‌ی «ضریب تمیز»^۱ نیز بپردازیم و در مورد هر پرسش، به این شاخص نیز توجه کنیم. از آن‌جا که ضریب تمیز، بحث کلیدی در این پژوهش نمی‌باشد، از بیان نحوه‌ی محاسبه‌ی آن چشم‌پوشی می‌کنیم، اما به بیان رابطه‌ی ضریب سهولت/دشواری و ضریب تمیز می‌پردازیم.

سیف (۱۳۸۷: ۳۱۵) رابطه‌ی دو شاخص فوق را با شکل (۵) نشان می‌دهد و می‌گوید: «حدود چهار ضلعی این شکل، حداکثر و حداقل ضرایب دشواری و تمیز را نشان می‌دهد. برای مثال، اگر ضریب دشواری برابر یا $0/8$ باشد، ضریب تمیز بین $+0/4$ و $-0/4$ خواهد بود. دقت کنید که تنها در صورتی ضریب تمیز به حداکثر یا $+1$ می‌رسد که ضریب دشواری برابر با $0/5$ یعنی در حد وسط باشد. در نتیجه، سوال‌های خوب یک آزمون، سوال‌هایی هستند با ضریب دشواری متوسط و ضریب تمیز بزرگ».



شکل ۵. رابطه‌ی ضریب سهولت/دشواری و ضریب تمیز (منبع: سیف، ۱۳۸۷: ۳۱۵)

^۱. discrimination index

۳،۲،۵. توزیع گزینه‌ها^۱

دو شاخص ضریب دشواری و ضریب تمیز نمی‌توانند نشانگر مناسب/نامناسب بودن گزینه‌ها نیز باشند؛ زیرا این دو شاخص با تعداد پاسخ‌های صحیح در ارتباط هستند (فرهادی و همکاران، ۲۰۰۷: ۱۰۴ و ۱۰۵). هدف از قرار دادن گزینه‌های انحرافی در سوال‌ها، منحرف کردن آزمون‌شوندگانی است که جواب درست را نمی‌دانند. بنابراین، گزینه‌های انحرافی خوب، باید عملاً از عهده‌ی این کار برآیند (سیف، ۱۳۸۷: ۳۱۶).

در حقیقت، یکی از چندین عاملی که در پاسخگویی زبان‌آموزان به پرسش‌ها دخیل است و در نهایت می‌تواند بر ضریب سهولت و متعاقباً بر ضریب تمیز یک پرسش تأثیر مستقیم بگذارد، گزینه‌های آن پرسش است. مثلاً یک پرسش چهارگزینه‌ای باید دارای سه گزینه باشد که بتوانند به خوبی نشان دهند که آزمون‌دهنده‌ی قوی، پاسخ صحیح را از چند پاسخ اشتباه تشخیص داده است و واقعاً آن پرسش، چهارگزینه‌ای بوده و نه یک‌گزینه‌ای! در همین راستا، مثال زیر را از فرهادی و همکاران (۲۰۰۷: ۱۰۵) نقل می‌کنیم:

«فرض کنید که ۱۰۰ آزمون‌دهنده به ۳ پرسش زیر پاسخ داده‌اند و توزیع گزینه‌ها (یعنی تعداد انتخاب هر گزینه توسط آزمون‌دهندگان) به ترتیب زیر است:

توزیع گزینه‌ها				گزینه‌ی صحیح	شماره‌ی پرسش
گزینه‌ی د	گزینه‌ی ج	گزینه‌ی ب	گزینه‌ی الف		
۰	۳۰	۶۰	۱۰	ب	۱
۵	۵	۴۰	۵۰	الف	۲
۴۰	۸	۱۲	۴۰	د	۳

به رغم این که ضریب سهولت هر سه سوال، مناسب است، الگوی توزیع گزینه‌ها مناسب نیست. مثلاً گزینه‌ی «د» در پرسش ۱، خوب عمل نکرده و می‌شود گفت که این پرسش سه گزینه‌ای بوده است و نه چهارگزینه‌ای! یا در مورد پرسش ۲، که از ضریب سهولت ایده‌آل ۵۰٪ نیز برخوردار است، دو گزینه‌ی کم‌کاربرد (ج و د) وجود دارد که توسط تعداد کمی از آزمون‌دهندگان انتخاب شده‌اند. در مورد پرسش ۴ هم می‌بینیم که گزینه‌ی نادرست «الف» به اندازه‌ی گزینه‌ی صحیح «د» انتخاب شده است؛ پس گزینه‌ی «الف» یا بیش از اندازه گمراه‌کننده است یا مفهوم بسیار نزدیکی به گزینه‌ی صحیح، یعنی «د» دارد. به هر حال، اینها مسائلی است که طراحان آزمون پس از انجام پیش‌آزمون یا آزمون باید به آنها توجه نموده و اصلاح‌شان کنند».

^۱. choice distribution

۶. روش پژوهش

این پژوهش به صورت مقایسه‌ای صورت گرفته است. طرفین این مقایسه بدین شرح هستند:

الف) نتایج حاصل از تحلیل آماری پاسخ‌های زبان‌آموزان مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) به پرسش‌های یک آزمون «خواندن» که در شهریور ماه ۱۳۹۲ به‌عنوان آزمون پایان‌دوره برای زبان‌آموزان سطح پایه (مبتدی) برگزار شده بود. این آزمون شامل ۲۸ پرسش در ۳ بخش بوده است که در جدول زیر مشاهده می‌شود:

قسمت اول؛ پرسش‌های شماره‌ی ۱ تا ۱۰	واژگان	بخش اول
قسمت دوم؛ پرسش‌های شماره‌ی ۱۱ تا ۲۰		
متن به‌هم‌ریخته؛ پرسش شماره‌ی ۲۱	درک	بخش دوم
متن اول؛ پرسش‌های شماره‌ی ۲۲ تا ۲۴		
متن دوم؛ پرسش‌های شماره‌ی ۲۵ تا ۲۸		

تعداد ۱۵۱ زبان‌آموز زن و مرد، از ملیت‌های مختلف (مانند چینی، یمنی، عراقی، پاکستانی، سوریه‌ای، تاجیکستانی و...) در این آزمون شرکت کرده بودند. پس از برگزاری آزمون، ضریب سهولت هر پرسش براساس پاسخ‌های زبان‌آموزان استخراج شد و مشخص شد که هر پرسش طبق روش آماری مطرح شده در زمینه‌ی آزمون‌سازی، تا چه اندازه سخت یا آسان است. نتایج این تحلیل و میزان سهولت هر یک از پرسش‌ها در جدول شماره‌ی (۲) در بخش (۷) آمده است.

ب) نظرات ۳۰ نفر از مدرّسان مرکز آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) در خصوص میزان سختی یا آسانی هر یک از پرسش‌ها که در این مقاله از این نظرات با عنوان برآوردها یا همان پیش‌بینی‌ها یا حدس‌های مدرّسان یاد می‌شود. به این منظور، پرسش‌نامه‌ای طراحی شد و از مدرّسان خواسته شد که در مورد هر پرسش، یکی از گزینه‌های «واقعاً سخت»، «خیلی سخت»، «سخت»، «آیده‌آل»، «آسان»، «خیلی آسان» و «واقعاً آسان» را انتخاب کنند. مشخصات مربوط به مدرّسان شرکت‌کننده در این پژوهش در جدول شماره‌ی (۱) آمده است.

جدول ۱. اطلاعات مربوط به مدرّسان شرکت‌کننده در تکمیل پرسش‌نامه

گذراندن واحد «آزمون‌سازی» یا واحد مشابه در مقطع ارشد یا دکترا		میانگین سابقه‌ی تدریس زبان‌های دیگر در آموزشگاه‌ها، مدارس و ...	سابقه‌ی تدریس زبان فارسی به صورت خصوصی		میانگین سابقه‌ی تدریس در مرکز زبان فارسی قزوین	جنسیت	
بله	خیر	۲/۶ سال	بله	خیر	۴,۸ دوره (ترم)	مرد	زن
۲۸ نفر	۲ نفر		۱۳ نفر	۱۷ نفر		۹ نفر	۲۱ نفر

بنابراین، هدف پژوهش حاضر این بوده است که نشان دهد حدس‌ها یا پیش‌بینی‌های مدرّسان درباره‌ی میزان سختی یا آسانی پرسش‌های یک آزمون، تا چه اندازه می‌تواند با واقعیت آزمون، که همان ضرایب سهولت برآمده از تحلیل پاسخ‌های آزمون‌دهندگان است، انطباق داشته باشد و لزوم یا عدم لزوم پیش‌آزمون و تحلیل‌های آماری پس از آن و نیز پس از آزمون اصلی را اثبات نماید.

۷. ارائه‌ی داده‌ها و یافته‌های پژوهش

ضرایب سهولتِ پرسش‌های آزمون مورد بحث در این پژوهش، پس از محاسبه‌ی ضریب سهولت هر پرسش، به همراه توصیف کیفی/واژگانی براساس پیوستار موجود در شکل (۴)، در جدول شماره‌ی (۲) نشان داده شده است.

جدول ۲. توصیف آماری آزمون خواندن سطح پایه

شماره‌ی پرسش	پاسخ‌های درست	پاسخ‌های نادرست	ضریب سهولت	توصیف براساس پیوستار استاندارد
۱	۸۸	۶۳	۰,۵۸	آسان
۲	۱۰۹	۴۲	۰,۷۲	خیلی آسان
۳	۸۴	۶۷	۰,۵۵	ایده‌آل
۴	۹۱	۶۰	۰,۶۰	آسان
۵	۶۷	۸۴	۰,۴۴	سخت
۶	۱۳۸	۱۳	۰,۹۱	واقعاً آسان
۷	۸۱	۷۰	۰,۵۳	ایده‌آل
۸	۱۰۰	۵۱	۰,۶۶	خیلی آسان
۹	۳۹	۱۱۲	۰,۲۵	خیلی سخت
۱۰	۹۱	۶۰	۰,۶۰	آسان
۱۱	۶۲	۸۹	۰,۴۱	سخت
۱۲	۹۱	۶۰	۰,۶۰	آسان
۱۳	۶۸	۸۳	۰,۴۵	ایده‌آل
۱۴	۹۹	۵۲	۰,۶۵	خیلی آسان
۱۵	۵۷	۹۴	۰,۳۷	سخت
۱۶	۷۹	۷۲	۰,۵۲	ایده‌آل
۱۷	۶۰	۹۱	۰,۳۹	سخت
۱۸	۵۸	۹۳	۰,۳۸	سخت
۱۹	۹۵	۵۶	۰,۶۲	آسان
۲۰	۵۱	۱۰۰	۰,۳۳	خیلی سخت
۲۱	۷۶	۷۵	۰,۵۰	ایده‌آل
۲۲	۵۵	۹۶	۰,۳۶	خیلی سخت
۲۳	۵۳	۹۸	۰,۳۵	خیلی سخت
۲۴	۶۵	۸۶	۰,۴۳	سخت
۲۵	۳۰	۱۲۱	۰,۱۹	خیلی سخت
۲۶	۱۲۰	۳۱	۰,۷۹	خیلی آسان
۲۷	۴۹	۱۰۲	۰,۳۲	خیلی سخت
۲۸	۱۱۸	۳۳	۰,۷۸	خیلی آسان

توصیف تعداد پرسش‌های آزمون براساس توصیف کیفی بر روی پیوستار موجود در شکل (۴)، در جدول شماره‌ی (۳) نیز مشاهده می‌شود.

جدول ۳. شمایی کلی از پرسش‌های آزمون
براساس توصیف کیفی/واژگانی بر روی پیوستار موجود در شکل (۴)

۰	واقعاً سخت
۶ پرسش	خیلی سخت
۶ پرسش	سخت
۵ پرسش	ایده‌آل
۵ پرسش	آسان
۵ پرسش	خیلی آسان
۱ پرسش	واقعاً آسان

در ادامه، به بررسی پرسش‌های مربوط به هر بخش از آزمون می‌پردازیم و نشان می‌دهیم که پیش‌بینی‌ها یا حدس‌های مدرّسان در مورد هر پرسش و همچنین، ضریب سهولت هر پرسش براساس پاسخ‌های زبان‌آموزان، چه‌قدر است و این ضریب سهولت در آزمون‌سازی، چه توصیفی را از پیوستار موجود در شکل (۴) به خود اختصاص می‌دهد.

همان‌گونه که در جدول شماره‌ی (۴) مشاهده می‌شود، در مورد پرسش‌های شماره‌ی ۲، ۶ و ۸ اختلاف فاحشی میان آنچه که آمار می‌گوید و آنچه که مدرّسان حدس زده‌اند، وجود دارد. در مورد پرسش شماره‌ی ۲، با ضریب سهولت ۰,۷۲ (خیلی آسان)، مشاهده می‌شود که بیشتر مدرّسان گزینه‌های ایده‌آل و سخت را انتخاب کرده‌اند. در مورد پرسش شماره‌ی ۶ نیز تفاوت، بسیار آشکار و معنی‌دار است؛ ۴۳,۳٪ از مدرّسان، این پرسش را که از جمله پرسش‌های واقعاً آسان با ضریب سهولت ۰,۹۱ بوده و جزو پرسش‌هایی است که از لحاظ ضریب تمیز نیز خوب عمل نمی‌کند، سخت دانسته‌اند و ۳۳,۳٪ از آنان نیز آن را ایده‌آل خوانده‌اند. در ۱۰ پرسش این بخش، از لحاظ آمار، دو پرسش شماره‌ی ۳ و ۷ جزو پرسش‌های ایده‌آل با ضرایب سهولت ۰,۵۵ و ۰,۵۳ محسوب می‌شوند، اما ۵۰٪ از مدرّسان، پرسش شماره‌ی ۷ را ایده‌آل دانسته‌اند و ۴۳,۳٪ از آنان تمایل به سخت یا خیلی سخت دانستن آن داشته‌اند. پرسش شماره‌ی ۳ را حتی یک نفر نیز ایده‌آل ندانسته و بیش از دو سوم مدرّسان آن را متمایل به آسان و خیلی آسان تلقی کرده‌اند.

پرسش‌های شماره‌ی ۱، ۴ و ۱۰ به ترتیب با ضرایب سهولت ۰,۵۸، ۰,۶۰ و ۰,۶۰ از منظر آمار، به‌عنوان پرسش‌هایی آسان، برآورد شده‌اند. اما قریب به ۴۰٪ از مدرّسان، پرسش شماره‌ی ۱ و ۷۷,۷٪ از آنها پرسش

شماره‌ی ۴ را رو به دشواری و واقعاً سخت بودن تشخیص داده‌اند. پرسش شماره‌ی ۱۰ را نیز ۷۰٪ از مدرّسان سخت و رو به دشوار بودن قلمداد نموده‌اند.

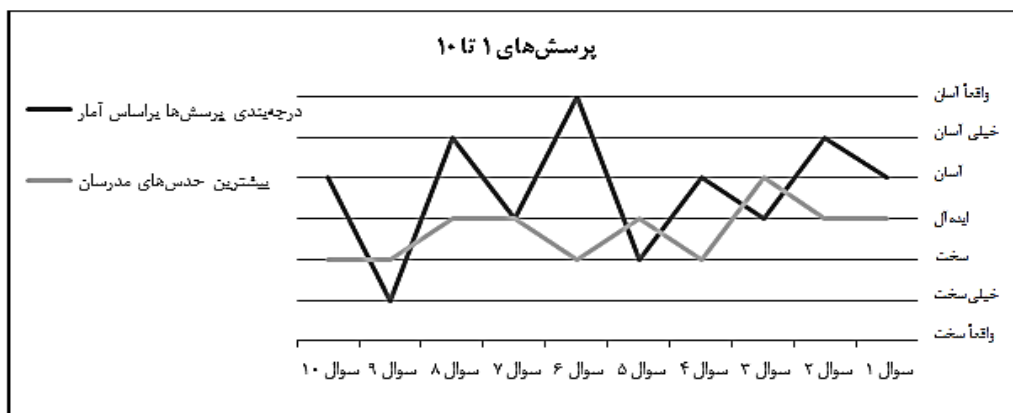
تنها در مورد پرسش‌های شماره‌ی ۵ و ۹ است که می‌توان تا حدودی بیشترین انطباق نظر مدرّسان و آمار را مشاهده نمود. این دو پرسش به ترتیب با ۰,۴۴ و ۰,۲۵ ضریب سهولت، سخت و خیلی سخت بوده‌اند. البته همچنان ۴۰٪ از مدرّسان، گزینه‌ی ایده‌آل و ۲۰٪ گزینه‌ی آسان را به پرسش شماره‌ی ۵ اختصاص داده‌اند. پرسش شماره‌ی ۹ تقریباً بیشترین میزان انطباق آمار و نظر مدرّسان را به خود اختصاص داده، اما با وجود این، ۲۰٪ نیز آن را ایده‌آل گفته‌اند.

در مجموع، در قسمت اول از بخش واژگان، یک پرسش واقعاً آسان، ۲ پرسش خیلی آسان، ۳ پرسش آسان، ۲ پرسش ایده‌آل، و ۲ پرسش خیلی سخت از منظر آمار وجود داشته، اما بیشترین تمایل برآوردهای مدرّسان به سخت یا خیلی سخت دانستن مجموعه‌ی پرسش‌ها بوده است.

جدول ۴. برآوردهای مدرّسان در مورد پرسش‌های شماره‌ی ۱ تا ۱۰ (بخش واژگان؛ قسمت اول)

توصیف مدرّسان شماره‌ی پرسش	واقعاً سخت	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعاً آسان	بدون نظر	ضریب سهولت براساس آمار ↓
پرسش شماره‌ی ۱	۳,۳٪	۳,۳٪	۳۳,۳٪	۵۳,۳٪	۶,۷٪	۰	۰	۰	۰,۵۸ (آسان)
پرسش شماره‌ی ۲	۰	۰	۱۳,۳٪	۴۳,۳٪	۴۰٪	۳,۳٪	۰	۰	۰,۷۲ (خیلی آسان)
پرسش شماره‌ی ۳	۰	۳,۳٪	۲۰٪	۰	۶۳,۳٪	۱۳,۳٪	۰	۰	۰,۵۵ (ایده‌آل)
پرسش شماره‌ی ۴	۱۰٪	۱۶,۷٪	۵۰٪	۱۶,۷٪	۶,۷٪	۰	۰	۰	۰,۶۰ (آسان)
پرسش شماره‌ی ۵	۳,۳٪	۳,۳٪	۳۰٪	۴۰٪	۲۰٪	۰	۰	۳,۳٪	۰,۴۴ (سخت)
پرسش شماره‌ی ۶	۰	۱۳,۳٪	۴۳,۳٪	۳۳,۳٪	۱۰٪	۰	۰	۰	۰,۹۱ (واقعاً آسان)
پرسش شماره‌ی ۷	۰	۱۰٪	۳۳,۳٪	۵۰٪	۶,۷٪	۰	۰	۰	۰,۵۳ (ایده‌آل)
پرسش شماره‌ی ۸	۶,۷٪	۱۰٪	۲۳,۳٪	۳۶,۷٪	۱۳,۳٪	۳,۳٪	۰	۶,۷٪	۰,۶۶ (خیلی آسان)
پرسش شماره‌ی ۹	۳,۳٪	۱۳,۳٪	۵۳,۳٪	۲۰٪	۰	۰	۰	۱۰٪	۰,۲۵ (خیلی سخت)
پرسش شماره‌ی ۱۰	۶,۷٪	۱۰٪	۵۳,۳٪	۱۳,۳٪	۳,۳٪	۰	۰	۱۳,۳٪	۰,۶۰ (آسان)

نمودار شماره‌ی (۱) براساس مقایسه‌ی دو متغیر «بیشترین حدس‌های مدرّسان درباره‌ی هر پرسش» و «ضریب سهولت پرسش‌ها براساس آمار» ترسیم شده است که به‌خوبی نشان می‌دهد مدرّسان بیشتر تمایل به ایده‌آل یا رو به سختی دانستن پرسش‌ها داشته‌اند.



نمودار ۱. مقایسه‌ی نتایج آماری و برآوردهای مدرّسان (پرسش‌های ۱ تا ۱۰)

در خصوص پرسش‌های بخش دوم واژگان که حاوی پرسش‌هایی به صورت کلوز تست بوده است، مشاهده می‌کنیم که در مورد پرسش شماره‌ی ۱۷ با ضریب سهولت ۰,۳۹ (سخت)، ۷۳,۳٪ از مدرّسان راهی مخالف را پیموده‌اند و تمایل به ایده‌آل و رو به آسان بودن این پرسش داشته‌اند (جدول شماره‌ی ۵).

بیشترین انطباق در مورد پرسش شماره‌ی ۱۵ با ضریب سهولت ۰,۳۷ (سخت) بوده است که بیش از نیمی از مدرّسان رأی به سخت و خیلی سخت بودن این پرسش داده‌اند.

دو پرسش شماره‌ی ۱۳ و ۱۶ به ترتیب با ضرایب سهولت ۰,۵۲ و ۰,۴۵ جزو پرسش‌های ایده‌آل در این گروه از پرسش‌ها می‌باشند که تنها ۲۶,۷٪ از مدرّسان پرسش ۱۳ را ایده‌آل دانسته‌اند و بیش از ۶۰٪ آن را روی طیف سخت تا واقعاً سخت تلقی نموده‌اند. پرسش شماره‌ی ۱۶ را نیز قریب به ۴۰٪ از مدرّسان سخت یا متمایل به سخت خوانده‌اند.

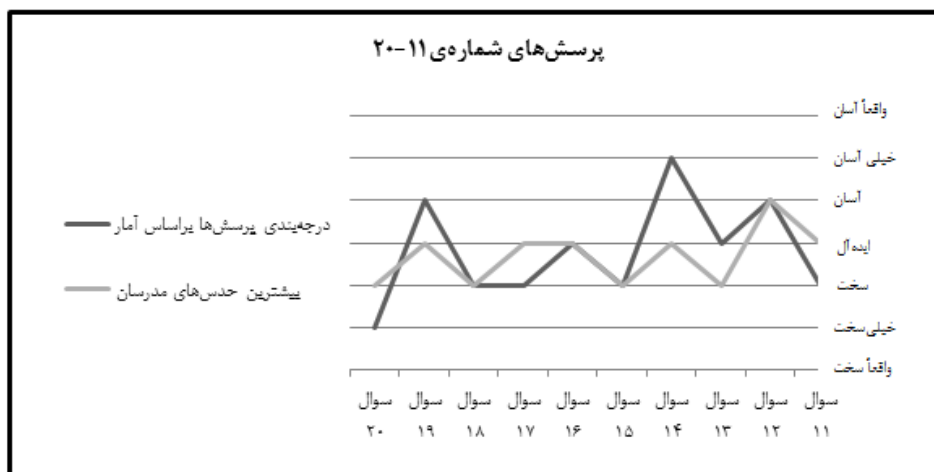
در خصوص پرسش‌های آسان (پرسش‌های شماره‌ی ۱۲ و ۱۹) و خیلی آسان (پرسش ۱۴) پراکندگی نسبتاً فراوانی دیده می‌شود. مثلاً هرچند که ۴۳,۳٪ از مدرّسان رأی به آسانی پرسش شماره‌ی ۱۲ داده‌اند، بقیه‌ی مدرّسان این پرسش را از خیلی سخت تا واقعاً آسان درجه‌بندی کرده‌اند. در مورد پرسش شماره‌ی ۱۹ نیز وضع به همین منوال است و با وجود این که ۲۳,۳٪ از مدرّسان آن را آسان خوانده‌اند، بقیه آن را ایده‌آل یا رو به سخت بودن برآورد کرده‌اند. پرسش خیلی آسان (از لحاظ آماری) شماره‌ی ۱۴ را نیز تنها ۶,۷٪ از مدرّسان، خیلی آسان دانسته‌اند و بقیه آن را روی طیفی از درجات ایده‌آل تا واقعاً سخت فرض کرده‌اند.

پرسش‌های شماره‌ی ۱۱، ۱۸، و ۲۰ از منظر آمار به ترتیب با ۰,۴۱، ۰,۳۸ و ۰,۳۳ پرسش‌هایی سخت و خیلی سخت بوده‌اند، اما ۵۳,۳٪ از مدرّسان، پرسش شماره‌ی ۱۱ را ایده‌آل گفته‌اند. پرسش‌های شماره‌ی ۱۸ و ۲۰ نیز روی پیوستاری از آسان تا واقعاً سخت توسط مدرّسان درجه‌بندی شده‌اند.

جدول ۵. برآوردهای مدرّسان در مورد پرسش‌های شماره‌ی ۱۱ تا ۲۰ (بخش واژگان؛ قسمت دوم)

توصیف مدرّسان شماره‌ی پرسش	واقعیاً سخت!	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعیاً آسان!	بدون نظر	ضریب سهولت براساس آمار ↓
پرسش شماره‌ی ۱۱	۰	۳،۳٪	۳۶،۷٪	۵۲،۳٪	۶،۷٪	۰	۰	۰	۰،۴۱ (سخت)
پرسش شماره‌ی ۱۲	۰	۶،۷٪	۶،۷٪	۳۳،۳٪	۴۲،۳٪	۳،۳٪	۳،۳٪	۳،۳٪	۰،۶۰ (آسان)
پرسش شماره‌ی ۱۳	۱۳،۳٪	۱۳،۳٪	۳۶،۷٪	۲۶،۷٪	۳،۳٪	۶،۷٪	۰	۰	۰،۴۵ (ایده‌آل)
پرسش شماره‌ی ۱۴	۳،۳٪	۶،۷٪	۲۶،۷٪	۳۳،۳٪	۲۲،۳٪	۶،۷٪	۰	۰	۰،۶۵ (خیلی آسان)
پرسش شماره‌ی ۱۵	۰	۱۰٪	۵۰٪	۳۶،۷٪	۳،۳٪	۰	۰	۰	۰،۳۷ (سخت)
پرسش شماره‌ی ۱۶	۶،۷٪	۶،۷٪	۲۶،۷٪	۴۶،۷٪	۶،۷٪	۰	۰	۶،۷٪	۰،۵۲ (ایده‌آل)
پرسش شماره‌ی ۱۷	۳،۳٪	۳،۳٪	۲۰٪	۳۳،۳٪	۲۶،۷٪	۱۰٪	۳،۳٪	۰	۰،۳۹ (سخت)
پرسش شماره‌ی ۱۸	۲۲،۳٪	۱۶،۷٪	۲۶،۷٪	۲۶،۷٪	۳،۳٪	۰	۰	۳،۳٪	۰،۳۸ (سخت)
پرسش شماره‌ی ۱۹	۰	۶،۷٪	۱۰٪	۵۲،۳٪	۲۲،۳٪	۶،۷٪	۰	۰	۰،۶۲ (آسان)
پرسش شماره‌ی ۲۰	۶،۷٪	۱۳،۳٪	۳۳،۳٪	۳۳،۳٪	۱۰٪	۰	۳،۳٪	۰	۰،۳۳ (خیلی سخت)

نمودار شماره‌ی (۲) که براساس مقایسه‌ی دو متغیر «بیشترین حدس‌های مدرّسان درباره‌ی هر پرسش» و «ضریب سهولت پرسش‌ها براساس آمار» ترسیم شده، نشان می‌دهد بیشتر نظرات مدرّسان حول گزینه‌های سخت و ایده‌آل بوده است، در صورتی که از منظر آمارهای پس از آزمون و با توجه به عملکرد زبان‌آموزان، ۲ پرسش ایده‌آل، ۲ پرسش آسان، و یک پرسش خیلی آسان وجود داشته است. اما در مجموع، نمودار مربوط به قسمت دوم از بخش واژگان، نسبت به بخش اول، اختلافات کمتری را نشان می‌دهد.



نمودار ۲. مقایسه‌ی نتایج آماری و برآوردهای مدرّسان (پرسش‌های ۱۱ تا ۲۰)

بخش دوم این آزمون دارای یک پرسش با عنوان «متن به‌هم‌ریخته» بوده (پرسشی نسبتاً طولانی) که همان‌گونه که در جدول شماره‌ی (۶) مشاهده می‌شود، این پرسش یک پرسش کاملاً ایده‌آل از جنبه‌ی آماری، با ضریب سهولت ۰,۵۰ بوده است و می‌تواند ضریب تمیز بسیار بالایی نیز داشته باشد، اما تنها ۱۳,۳٪ از مدرّسان، این پرسش را ایده‌آل دانسته‌اند و نزدیک به ۸۰٪ از مدرّسان آن را سخت و حتی واقعاً سخت حدس زده‌اند.

جدول ۶. برآوردهای مدرّسان در مورد پرسش شماره‌ی ۲۱ (بخش دوم؛ متن به‌هم‌ریخته)

توصیف مدرّسان ← شماره‌ی پرسش	واقعیاً سخت!	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعیاً آسان!	بدون نظر	ضریب سهولت براساس آمار ↓
پرسش شماره‌ی ۲۱	۱۶,۷٪	۲۶,۷٪	۳۶,۷٪	۱۳,۳٪	۳,۳٪	۳,۳٪	۰	۰	۰,۵۰ (ایده‌آل)

سومین بخش از آزمون مورد بحث در این پژوهش، شامل دو متن بوده است که پرسش‌های شماره‌ی ۲۲ تا ۲۸ را شکل می‌دهند. در مورد متن اول، پرسش‌های ۲۲ تا ۲۴، با ضرایب سهولتی بین ۰,۳۵ تا ۰,۴۳ جزو پرسش‌های خیلی سخت و سخت بوده‌اند، اما مجموعاً ۱۰۳٪ از مدرّسان رأی به ایده‌آل بودن این سه پرسش، مخصوصاً پرسش‌های شماره‌ی ۲۳ و ۲۴ داده‌اند و حتی نزدیک به ۳۰٪ از آنها این پرسش‌ها را آسان یا خیلی آسان حدس زده‌اند (جدول شماره‌ی ۷).

جدول ۷. برآوردهای مدرّسان در مورد پرسش‌های شماره‌ی ۲۲ تا ۲۴ (بخش سوم؛ متن اول)

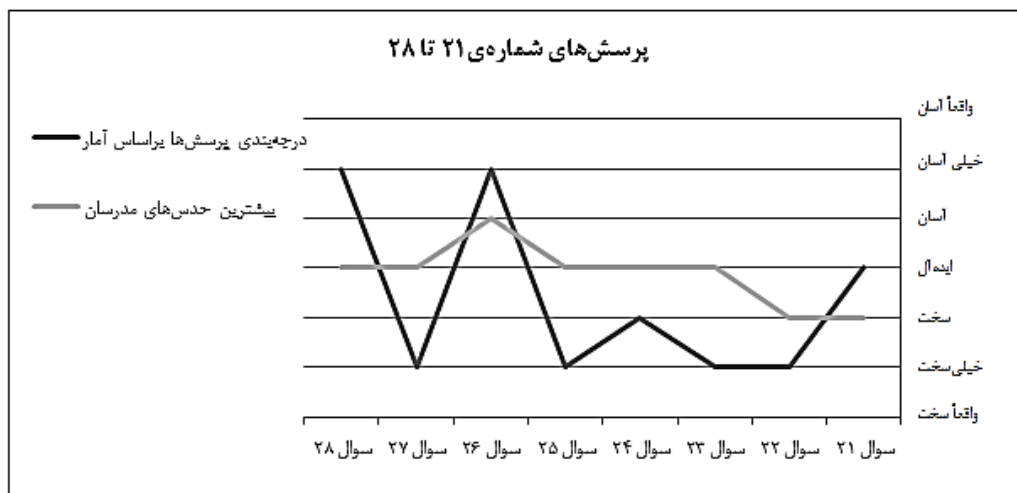
توصیف مدرّسان ← شماره‌ی پرسش	واقعیاً سخت!	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعیاً آسان!	بدون نظر	ضریب سهولت براساس آمار ↓
پرسش شماره‌ی ۲۲	۲۰٪	۲۶,۷٪	۳۳,۳٪	۲۰٪	۰	۰	۰	۰	۰,۳۶ (خیلی سخت)
پرسش شماره‌ی ۲۳	۶,۷٪	۱۰٪	۳۶,۷٪	۴۰٪	۳,۳٪	۳,۳٪	۰	۰	۰,۳۵ (خیلی سخت)
پرسش شماره‌ی ۲۴	۰	۳,۳٪	۲۶,۷٪	۴۳,۳٪	۲۰٪	۳,۳٪	۰	۳,۳٪	۰,۴۳ (سخت)

پرسش‌های شماره‌ی ۲۵ تا ۲۸ مربوط به دومین متن هستند که اطلاعات آماری و حدس‌های مدرّسان درباره‌ی این پرسش‌ها در جدول شماره‌ی (۸) آمده است. همان‌گونه که مشاهده می‌شود، طبق آمار، دو پرسش شماره‌ی ۲۶ و ۲۸ با ضریب سهولت ۰,۷۹ و ۰,۷۸ از پرسش‌های خیلی آسان محسوب می‌شوند، اما تنها تقریباً ۲۰٪ از مدرّسان پرسش شماره‌ی ۲۶، و ۶,۷٪ پرسش شماره‌ی ۲۸ را خیلی آسان پیش‌بینی کرده‌اند.

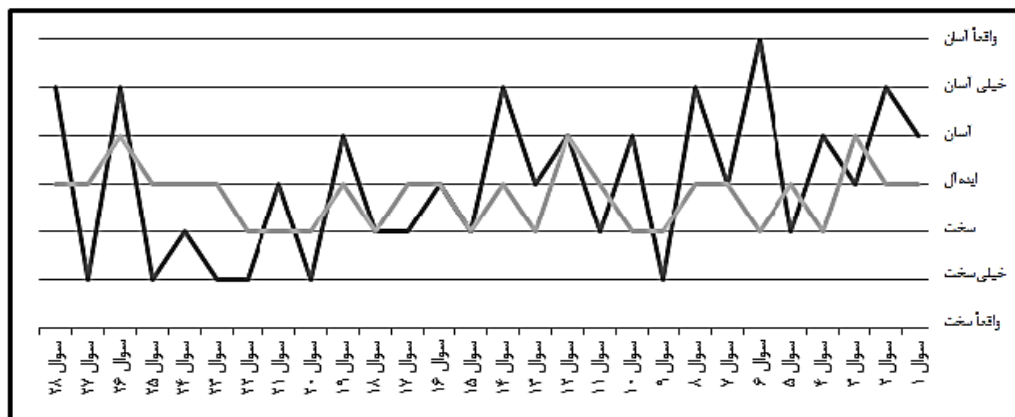
مجموعاً در مورد چهار پرسش موجود در این بخش، بیشترین نظر مدرّسان دال بر ایده‌آل بودن این پرسش‌ها بوده است (با ۱۵۹,۹٪ در مجموع چهار پرسش) که این نشان از اختلاف پیش‌بینی‌ها و تحلیل‌های آماری (واقعیت آزمون) دارد. نمودار شماره‌ی (۳)، توصیفی از وضعیت موجود در خصوص پرسش‌های شماره‌ی ۲۱ تا ۲۸ می‌باشد و نمودار شماره‌ی (۴) نیز در یک نگاه کلی، وضعیت همه‌ی پرسش‌ها را در ارتباط با «حدس و پیش‌بینی مدرّسان» و «ضرایب سهولت برآمده از آمار» ارائه می‌نماید.

جدول ۸. برآوردهای مدرّسان در مورد پرسش‌های شماره‌ی ۲۵ تا ۲۸ (بخش سوم؛ متن دوم)

توصیف مدرّسان شماره‌ی پرسش	واقعیاً سخت!	خیلی سخت	سخت	ایده‌آل	آسان	خیلی آسان	واقعیاً آسان!	بدون نظر	ضریب سهولت براساس آمار ↓
پرسش شماره‌ی ۲۵	۱۰٪	۱۳,۳٪	۳۰٪	۳۳,۳٪	۱۳,۳٪	۰	۰	۰	۰,۱۹ (خیلی سخت)
پرسش شماره‌ی ۲۶	۰	۰	۱۳,۳٪	۳۳,۳٪	۳۳,۳٪	۲۰٪	۰	۰	۰,۷۹ (خیلی آسان)
پرسش شماره‌ی ۲۷	۶,۷٪	۱۶,۷٪	۲۶,۷٪	۴۰٪	۱۰٪	۰	۰	۰	۰,۳۲ (خیلی سخت)
پرسش شماره‌ی ۲۸	۰	۳,۳٪	۶,۷٪	۵۳,۳٪	۳۰٪	۶,۷٪	۰	۰	۰,۷۸ (خیلی آسان)



نمودار ۳. مقایسه‌ی نتایج آماری و برآوردهای مدرّسان (پرسش‌های ۲۱ تا ۲۸)



نمودار ۴. مقایسه‌ی نتایج آماری (خط پررنگ) و برآوردهای مدرّسان (خط کم‌رنگ) درباره‌ی همهی پرسش‌ها

در مجموع، یافته‌های این مقایسه نشان می‌دهند که

(الف) در مورد بعضی از پرسش‌ها، حدس‌های تعدادی از مدرّسان نزدیک به واقعیت بوده، اما در همان حالت نیز نیم یا بیش از نیمی از آنان گزینه‌های دیگری را برای همان پرسش انتخاب کرده بودند.

(ب) بعضی از پرسش‌ها که طبق تحلیل آماری، اصلاً ایده‌آل نبوده‌اند، توسط مدرّسان جزو پرسش‌های ایده‌آل قلمداد شده بودند.

(پ) بیشترین حدس‌های مدرّسان را توصیف‌هایی همچون سخت و خیلی سخت شامل می‌شد و در مجموع، مدرّسان بیشتر به سخت یا خیلی سخت دانستن پرسش‌ها تمایل داشتند؛ در مقابل، گزینه‌های مربوط به آسانی، چندان در انتخاب‌ها به چشم نمی‌خوردند.

(ت) پراکندگی توصیف‌های انتخابی توسط مدرّسان در بعضی پرسش‌ها به اندازه‌ای گسترده است که می‌توان مشاهده نمود مدرّسان با درصدهایی نزدیک به هم، توصیف‌های گوناگونی را برای یک پرسش برگزیده‌اند.

۸. نتیجه‌گیری

این پژوهش در پی آن بود که دو موضوع را در مورد یک دوره آزمون خواندن سطح پایه برای فارسی‌آموزان غیر فارسی‌زبان بررسی کند: (۱) تشخیص مدرّسان در خصوص این که پرسش‌های این آزمون تا چه اندازه سخت، آسان یا ایده‌آل هستند که ما از این تشخیص با عنوان‌های «قضاوت»، «برآورد» یا «پیش‌بینی و حدس» مدرّسان، بدون توجه به نتایج حاصل از بررسی آماری داده‌ها و پاسخ‌های فارسی‌آموزان، یاد کردیم. در واقع، نوعی قضاوت شمی و براساس تجربه (با توجه به این که میانگین تجربه‌ی تدریس این مدرّسان

نزدیک به پنج ترم بوده است)؛ (۲) ضرایب سهولت هر یک از پرسش‌های آزمون مذکور براساس تحلیل آماری پاسخ‌های زبان‌آموزان به هر پرسش و از طریق فرمول ضریب سهولت. براساس معیارهای ارائه شده در منابع آزمون‌سازی، ضریب سهولت بر روی پیوستاری از صفر تا یک قرار گرفت و به هر بخش از این پیوستار، توصیفی واژگانی (مانند واقعاً سخت، خیلی سخت، سخت، ایده‌آل و...) تعلق گرفت.

هدف دوم، اما مهم‌تر این پژوهش آن بود که لزوم یا عدم لزوم سپری کردن مرحله‌ی پیش‌آزمون و تحلیل آماری (همان حالت رفت و برگشتی که در شکل شماره‌ی (۲) نشان داده شده است)، از منظر میزان صحت و نزدیکی تشخیص مدرسان به واقعیت آماری پرسش‌ها اثبات شود. یعنی این موضوع که در رویکرد علمی نوین به مسأله‌ی آزمون‌های زبان گفته می‌شود که قبل از برگزاری آزمون باید پیش‌آزمون و تحلیل آماری صورت گیرد تا پرسش‌ها براساس توانایی زبان‌آموزان (پاسخ‌های عینی و واقعی خود زبان‌آموزان) درجه‌بندی گردند و پس از آزمون نیز تحلیل‌ها ادامه یابند، ضرورتی دارد یا خیر؟ آیا می‌توان به تجربه و شمّ مدرّسان اکتفا کرد و از پیش‌آزمون و تحلیل آماری بی‌نیاز شد؟

نتایج این پژوهش نشان داد که از این نظر که نمی‌توان به قضاوت‌های شمّی و توصیف‌های زبانی صرف اکتفا کرد، پیش‌آزمون و تحلیل آماری، جزء لاینفک فرایند آزمون‌سازی محسوب می‌شود و از رویکرد علمی-آماری به آزمون‌سازی نباید غافل شد. بنابراین، می‌توان با استناد به مطالب این پژوهش، این‌گونه عنوان کرد که با اتکا به پیش‌بینی‌ها یا حدس‌های مدرّسان، ممکن است

- پرسش‌هایی خیلی سخت یا خیلی آسان، به‌عنوان پرسش‌هایی ایده‌آل محسوب شده و در آزمون اصلی از آنها استفاده شود.

- پرسش‌هایی آسان یا خیلی آسان (و مسلماً با ضریب تمیز کم) به‌عنوان پرسش‌هایی سخت شناخته شوند و از گردونه‌ی آزمون حذف شوند.

- پرسش‌هایی کاملاً ایده‌آل و مناسب به‌عنوان پرسش‌هایی خیلی سخت یا خیلی آسان محسوب شده و کنار گذاشته شوند.

- آزمون‌هایی برگزار شوند که پرسش‌هایشان یا خیلی آسان هستند یا خیلی سخت، و ضریب تمیز مناسبی ندارند و این باعث شود که آزمون‌ها دوباره به همان نظام و رویکرد سنتی بازگردند و آسانی آزمون، ملاکی برای تدریس خوب مدرّسان، و سختی آن ملاکی برای مناسب نبودن پرسش‌ها تلقی شود.

پس همان‌گونه که ویر (۱۹۹۰: ۴۱) نیز اشاره می‌کند که آزمون آزمایشی قبل از برگزاری آزمون اصلی باید با شرایط ویژه‌ای برگزار شود تا پرسش‌های آزمون از هر لحاظ تحت کنترل قرار گیرند، می‌توان گفت نتایج این پژوهش نیز مؤید همین نکته است. همین‌طور، از آن‌جا که سطح دشواری سؤالات نیز در میزان

روایی یک آزمون اهمیت دارد و سؤالات خیلی ساده یا خیلی دشوار، روایی آزمون را به خطر می‌اندازد (زیرا چنین سؤالاتی، یا بالاتر و یا پایین‌تر از سطح توانایی آزمون‌دهندگان هستند) و معلومات واقعی آزمون‌دهنده را اندازه‌گیری نمی‌کنند (فرهادی و همکاران، ۲۰۰۷: ۱۵۵)، ضرورت دارد که به پیش‌آزمون و همچنین تحلیل‌های آماری پس از آزمون، توجه ویژه‌ای شود.

واضح و آشکار است که دستیابی به همهی موارد فوق، صرفاً با وجود پیش‌آزمون و تحلیل‌های آماری و به‌طور کلی، با بهره‌گیری از رویکرد علمی-آماري و نظام‌مند به آزمون‌سازی محقق می‌شود. در پایان، برای انواع برآوردها، حدس‌ها و پیش‌بینی مدرّسان در خصوص سختی یا دشواری پرسش‌ها می‌توان دلایل احتمالی زیر را ذکر کرد که می‌توانند به‌عنوان موضوعاتی برای پژوهش‌های آتی نیز مورد توجه قرار گیرند:

- شناخت مدرّسان مختلف از زبان‌آموزان کلاس‌های خودشان (از جنبه‌های مختلف مانند هوش، ملیت و... . مثلاً یک مدرّس ممکن است معیار قضاوتش زبان‌آموزانی از ملیت‌هایی باشد که فارسی را بهتر از ملیت‌های دیگر می‌آموزند)

- توجه مدرّسان به روش تدریس و محتوای تدریس خودشان

- توجه به سطح منابع درسی از لحاظ واژگان و...

- تمایل احتمالی برخی از مدرّسان به بالا بردن آمار قبولی در کلاس‌هایشان و سعی در آسان‌سازی هر چه بیشتر آزمون

- پیشینه‌ی تحصیلی مدرّسان از این لحاظ که خودشان در طول دوران یادگیری زبان‌هایی دیگر، آزمون‌های سخت یا آسانی را تجربه کرده‌اند.

- عدم تعریف واحد و یکپارچه از مفهوم سطوح زبانی (مانند پایه، میانی، پیشرفته و...) در نزد مدرّسان و مسئولان برگزاری آزمون به دلایل مختلف مانند نبود یک پیکره‌ی زبانی در زبان فارسی

- عدم وجود منابع مناسب، واحد و یکپارچه در همهی کلاس‌ها برای همهی زبان‌آموزان

- روال‌ها و سیاست‌گذاری‌های واحدهای آموزش‌دهنده‌ی زبان فارسی

- تأثیر عواطف و احساسات موجود میان مدرّسان و زبان‌آموزان، و واهمه‌ی مدرّسان از شناخته شدن به‌عنوان مدرّس سختگیر یا آسانگیر

تقدیر و سپاسگزاری

بر خود لازم می‌دانیم که صمیمانه از همهی مدرّسان گروه آموزش زبان فارسی دانشگاه بین‌المللی امام خمینی (ره) که در کمال تواضع و با صبر و حوصله‌ای بی‌اندازه، ما را در انجام این پژوهش یاری رساندند، قدردانی و سپاسگزاری نماییم.

منابع:

- بوخنسکی، ا. (۱۳۷۹). *فلسفه‌ی معاصر اروپایی*. ترجمه‌ی شرف‌الدین خراسانی. تهران: نشر علمی فرهنگی. بازیابی شده از: <http://emamhossein2.mihanblog.com/post/258> (تاریخ انتشار در تارنما: دوشنبه ۱۳۹۲/۰۳/۰۶، نویسنده: حسین مرادی).
- آلوستانی مفرد، م. (۱۳۸۵). نظریه‌های تصویری زبان و بازی‌های زبانی ویتگنشتاین. روزنامه‌ی رسالت، شماره‌ی ۵۹۶۱ (۱۳۸۵/۰۶/۲۲)، ص ۱۸ (اندیشه‌ی غرب)؛ بازیابی شده از: <http://www.magiran.com/n1204170>.
- جعفرپور، ع. (۱۳۷۵). *مقدمه‌ای بر آزمون‌شناسی زبان*. شیراز: انتشارات دانشگاه شیراز.
- جلیلی، س. ا. (۱۳۹۰). *مبانی تهیه و تدوین آزمون مهارتی فارسی (آمفا)*. رساله‌ی کارشناسی ارشد. دانشگاه علامه طباطبائی.
- چپمن، ش. (۱۳۸۴). *از فلسفه به زبان‌شناسی*. ترجمه‌ی حسین صافی. تهران: نشر گام نو.
- سیف، ع. ا. (۱۳۸۷). *اندازه‌گیری، سنجش، و ارزشیابی آموزشی (ویرایش پنجم)*. تهران: نشر دوران.
- ممقانی، ه. (۱۳۹۰). *ارزیابی روایی و پایایی آزمون وزارت علوم، تحقیقات و فناوری (MSRT) با هدف دستیابی به یک آزمون مهارت برای زبان فارسی*. رساله‌ی کارشناسی ارشد. دانشگاه علامه طباطبائی.
- موسوی، ث. (۱۳۷۹). *بررسی راهکارهای نظری آزمون‌های بسندگی و تهیه‌ی یک آزمون بسندگی استاندارد زبان فارسی برای غیرفارسی‌زبانان*. رساله‌ی کارشناسی ارشد. دانشگاه شیراز.
- نوروززاده، ن. (۱۳۹۰). کاربرد آزمون‌سازی انطباقی در تعیین سطح زبان انگلیسی دانشجویان. *رشد آموزش زبان*، دوره‌ی ۲۶، شماره ۲، زمستان ۱۳۹۰، صص: ۱۲-۶.

- Alderson, J. C.** (2000). *Assessing Reading*. Cambridge University Press.
- Birjandi, P. & Mosallanejad, P.** (2010). *An Overview of Testing and Assessment*. Esfahan: Sepahan Publication.
- Brown, H. D.** (2003). *Language Assessment, Principles and Classroom Practices*. White Plains, NY: Pearson Education.
- Brown, H. D.** (2010). *Language Assessment Principles and Classroom Practices (2nd edition)*. USA: Pearson Education, Inc.
- Brown, J. D.** (2008). *Testing in Language Programs*. USA: McGraw-Hill.
- Coombe, C. et al.** (2007). *A Practical Guide to Assessing English Language Learners*. US: The University of Michigan Press.
- Farhady, H. et al.** (2007). *Testing Language Skills from Theory to Practice*. Tehran: SAMT.
- Hughes, A.** (2003). *Testing for Language Teachers (2nd edition)*. Cambridge University Press.
- McNamara, T. & Roever, C.** (2006). *Language Testing: The Social Dimension*. USA: Blackwell Publishing.
- Mousavi, S. A.** (1997). *A Dictionary of Language Testing*. Iran: Rahnama Publications.
- Weir, C. J.** (1990). *Communicative Language Testing*. UK: Prentice Hall International Ltd.

پیوست: آزمون خواندن سطح پایه

بخش یکم: واژگان

الف) برای هر جای خالی بهترین گزینه را انتخاب کنید.

- ۱- یک دوست خوب کسی است که هرگز شما را تنها نمی‌گذارد و اگر با او مشورت کنید شما را می‌کند.
 الف) راهنمایی (ب) آماده (ج) پیشنهاد (د) آشنا
- ۲- من هر روز صبح دست و صورتم را و بعد، با حوله می‌کنم.
 الف) می‌شوم - خشک (ب) می‌شویم - خشک (ج) می‌شویم - پهن (د) می‌شورم - پهن
- ۳- کودکان معمولاً بعد از انجام کارهای اشتباه به پدر و مادرشان می‌دهند که دیگر آن کار را در آینده تکرار نکنند.
 الف) قرار (ب) تصمیم (ج) قول (د) ترجیح
- ۴- سیما یک دانش‌آموز حواس‌پرت و است! امروز دوباره کیفش را در خانه جا و بدون کیف به مدرسه رفت.
 الف) سر به هوا - گرفت (ب) فراموش کار - گذاشت (ج) کم حرف - گذاشت (د) تنبل - گذشت
- ۵- خیلی خسته‌ام و خوابم وقتی به خانه برسم، روی تخت خواب دراز چشم‌هایم را می‌بندم و یک چرت کوتاه می‌زنم.
 الف) می‌آید - می‌کنم (ب) می‌آید - می‌کشم (ج) می‌برد - می‌ریزم (د) می‌برد - می‌دهم
- ۶- هفته‌ی پیش وقتی ماشینم در خیابان شد، چند نفر کمک کردند و آن را دادند تا به تعمیرگاه ببرم.
 الف) پنجر - دستور (ب) تعمیر - فشار (ج) جریمه - نشان (د) خراب - هل
- ۷- انسان‌های بداخلاق و را کسی دوست ندارد، زیرا خوش‌اخلاقی و خوش‌رویی از انسان‌های خوب است.
 الف) پاکیزه - ویژگی‌های (ب) اخمو - ویژگی‌های (ج) سخت‌گیر - اندازه‌های (د) اخمو - اندازه‌های
- ۸- سارا: این روزها حال خوبی ندارم! احساس می‌کنم غمگین و هستم.
 سوسن: به نظرم تو دچار افسردگی شده‌ای و باید به یک مراجعه کنی.
 الف) نگران - روان‌شناس (ب) خجالتی - ستاره‌شناس (ج) ترسو - زیست‌شناس (د) شاداب - روان‌شناس
- ۹- ساسان: وقت داری امروز با هم تئاتر برویم؟
 مهیار: آره، اسم امروز چیست؟
 ساسان: اسمش را نمی‌دانم، اما ساعت ۷ می‌آیم که با هم برویم. حتماً خوش‌ت
- الف) نمایش - دنبالت - می‌آید (ب) همایش - دنبالت - می‌گذرد
 ج) نمایش - کنارت - می‌گذرد (د) همایش - کنارت - می‌آید
- ۱۰- چند روز پیش یک نفر را در این شهر کشتند. پلیس زیادی برای قاتل کرده است، اما موفق نشده است او را بگیرد.
 الف) جستجوی - پیدا کردن - بعضی وقت‌ها (ب) تلاش - دستگیری - هنوز
 ج) تلاش - جستجوی - کم‌کم (د) سعی - دستگیری - گاهی

ب) برای هر جای خالی یک گزینه‌ی مناسب انتخاب کنید.

بعضی از شغل‌ها وجود دارند که فقط اسم دارند! یعنی این شغل‌ها معروف هستند و مردم فکر می‌کنند داشتن این شغل یعنی زندگی بهتر. یکی از این شغل‌ها نویسندگی است. نویسندگی یکی از کارهایی است که شاید (۱۱)..... خوبی نداشته باشد، اما بسیاری از مردم همیشه یک نویسنده را یک انسان (۱۲)..... و آگاه می‌دانند. آنها فکر می‌کنند یک نویسنده اطلاعات زیادی دارد و کتاب‌های زیادی خوانده است. (۱۳)..... معمولاً مردم نویسنده‌ها را دوست دارند؛ البته یک نویسنده‌ی خوب و ماهر باید هم آن ویژگی‌ها را (۱۴)..... . او باید در میان مردم زندگی کند و با آنها (۱۵)..... داشته باشد تا بتواند مطلبی را بنویسد که زندگی واقعی مردم را (۱۶)..... . نویسنده باید زبان مادری خودش را هم خوب بداند و سعی کند در نوشته‌های خودش غلط و اشتباه نداشته باشد. یکی دیگر از ویژگی‌های یک نویسنده‌ی خوب (۱۷)..... نوشتن است، یعنی جوری ننویسد که فقط خودش بفهمد! بلکه باید خیلی ساده و مشخص بنویسد. شاید (۱۸).....ترین چیز برای نویسنده‌ها این باشد که با فرهنگ‌ها و سنت‌های مختلف مردم دنیا آشنا باشند تا بتوانند از آنها در کتاب‌ها، داستان‌ها و مطالب خودشان (۱۹)..... کنند. نویسندگی انواع مختلفی دارد: روزنامه‌نگار، داستان‌نویس و... . عده‌ای (۲۰)..... هم مطالب علمی و تحقیقی می‌نویسند؛ کار این گروه از نویسندگان همیشه سخت‌تر است، زیرا باید درباره‌ی چیزهایی که می‌نویسند اول تحقیق کنند و بعد در نوشته‌هایشان دقت کنند تا مطلب اشتباهی ننویسند.

۱۱- الف) هزینه	ب) تلاش	ج) گنجایش	د) درآمد
۱۲- الف) کارگر	ب) دانا	ج) رایگان	د) راهرو
۱۳- الف) زیرا	ب) به علت این که	ج) به همین دلیل	د) اگر
۱۴- الف) داشته باشد	ب) آورده باشد	ج) دیده باشد	د) برداشته باشد
۱۵- الف) معاینه	ب) رابطه	ج) علاقه	د) مراجعه
۱۶- الف) یاد می‌دهد	ب) تمرین می‌کند	ج) ادامه می‌دهد	د) نشان می‌دهد
۱۷- الف) تُند	ب) کُند	ج) ساده	د) داستان
۱۸- الف) ضروری	ب) نهایی	ج) مرغوب	د) عمومی
۱۹- الف) پیدا	ب) استفاده	ج) درست	د) فایده
۲۰- الف) یکدیگر	ب) درباره	ج) باز	د) دیگر

بخش دوم: متن به هم ریخته

- ۲۱- این یک متن به هم ریخته است. کدام گزینه ترتیب درست آن را نشان می‌دهد؟
- ۱) تعداد این گیرنده‌ها در بعضی از اندام‌ها مانند زبان، دست و نوک انگشتان بسیار زیاد است.
- ۲) این سرعت‌های مختلف به این دلیل است که در اعضای بدن انسان گیرنده‌های حسی گرما و سرما وجود دارد.
- ۳) تعداد گیرنده‌ها در دیگر قسمت‌های بدن مانند ران و پشت کمر کم‌تر است.
- ۴) به همین دلیل، این اندام‌ها سرما و گرما را بیشتر و سریع‌تر احساس می‌کنند.
- ۵) تعداد کم گیرنده‌ها باعث می‌شود در اندام‌هایی مثل ران و پشت کمر، سرما و گرما به خوبی احساس نشود.
- ۶) اعضای بدن ما سرما و گرما را با سرعت‌های مختلفی احساس می‌کنند.

الف) ۵-۴-۳-۶-۱-۲ ب) ۵-۴-۳-۱-۲-۶ ج) ۵-۳-۴-۶-۱-۲ د) ۵-۳-۴-۱-۲-۶

بخش سوم: درک متن (۲ متن)

متن زیر را بخوانید و به پرسش‌های آن پاسخ دهید:

پول توجیبی پولی است که کودکان یا نوجوانان از پدر یا مادرشان می‌گیرند و آن را برای خودشان خرج می‌کنند. معمولاً والدین این پول را روزانه به فرزندان پرداخت می‌کنند، اما پرداخت پول توجیبی به صورت هفتگی یا ماهانه نیز رایج است. درست است که این پول زیاد نیست، ولی می‌تواند راهی برای یادگیری مدیریت باشد! البته والدین هم باید به کودکان کمک کنند. وقتی کودکان و نوجوانان پول توجیبی دریافت می‌کنند باید یاد بگیرند که آن را چگونه خرج کنند. مثلاً چیزهایی بخرند که مهم‌تر هستند و بیشتر به آنها نیاز دارند. در واقع، آنها می‌توانند مدیریت مالی و پولی را از دوران کودکی و نوجوانی و با استفاده از پول توجیبی بیاموزند. اگر یک کودک یا نوجوان بتواند پول‌هایش را مدیریت کند، می‌تواند از همان پول‌هایی که دارد مقداری را نیز پس‌انداز کند و از این پس‌اندازها در آینده استفاده کند.

امروزه مقدار پول توجیبی که بچه‌ها از والدین می‌گیرند خیلی زیاد شده است. یکی از دلایل‌های بالا رفتن مقدار پول توجیبی، گران شدن اجناس و کالاهایی است که بچه‌ها می‌خواهند بخرند. دلیل دیگر هم تبلیغات زیاد است. کودکان و نوجوانان همیشه دوست دارند چیزهایی را بخرند که تلویزیون آنها را تبلیغ می‌کند؛ مانند اسباب‌بازی‌ها، خوردنی‌ها و تنقلات. تبلیغات باعث می‌شود که کودک بخواهد این اشیاء را بخرد و همه‌ی پول‌هایش را برای خرید آنها خرج کند. به همین دلیل والدین مجبورند پول بیشتری به بچه‌هایشان بدهند. این علاقه به خرید، بعضی وقت‌ها مشکلاتی را هم به وجود می‌آورد؛ مثلاً کودکی که پول کافی برای خرید همه‌ی آن چیزها ندارد، شاید بدون اجازه از کیف پدر یا مادرش پول بردارد یا حتی از مغازه‌ها دزدی کند! گاهی نیز کودک از این که او پول کافی برای خرج کردن ندارد و دوستانش پول دارند خجالت می‌کشد. این هم باعث می‌شود والدین پول توجیبی بیشتری به او بدهند تا جلوی دوستانش خجالت نکشد. البته وضعیت همه‌ی بچه‌ها یکسان نیست و در دنیا کودکانی هستند که نه تنها پول از کسی نمی‌گیرند، بلکه مشغول انجام کارهای سخت و ناسالم با دست‌مُزد خیلی کم هستند. برای این کودکان پول توجیبی اصلاً معنی ندارد. آنها حتی به مدرسه هم نمی‌روند.



۲۲- بهترین عنوان برای اولین بند (پاراگراف) چیست؟

- (الف) پول توجیبی چیست؟
- (ب) آموختن مدیریت به وسیله‌ی پول توجیبی
- (ج) پس‌انداز کردن پول‌های توجیبی
- (د) مدیریت پول توجیبی به وسیله‌ی والدین

۲۳- طبق متن، کدام گزینه دلیل زیاد شدن مقدار پول‌های توجیبی نیست؟

- (الف) مشکلات کودکان
- (ب) تلویزیون
- (ج) بالا رفتن قیمت‌ها
- (د) خجالت کشیدن از دوستان

۲۴- کدام گزینه به معنی واژه‌ی دست‌مُزد در خط شانزدهم نزدیک‌تر است؟

- (الف) قیمت
- (ب) جایزه
- (ج) پول توجیبی
- (د) حقوق

متن زیر را بخوانید و به پرسش‌های آن پاسخ دهید:

دنیای حیوانات همیشه دنیایی عجیب و جالب بوده است. یکی از چیزهای عجیب در دنیای حیوانات، فداکاری آنها برای به دنیا آوردن فرزندان‌شان یا برای نگهداری از آنهاست. مثلاً وقتی بچه‌های خرگوش متولد می‌شوند، خرگوش مادر موها و پشم بدن خودش را می‌کند و با آنها در گوشه‌ای از لانه، یک تَشک نرم درست می‌کند. سپس نوزادان را روی این تشک نرم قرار می‌دهد تا بدن بی‌مو و ضعیف‌شان روی زمین سِفْت لانه اذیت نشود. حیوان دیگری که برای فرزندانش فداکاری می‌کند هشت‌پا است. این حیوان تخم‌هایش را در مکانی امن، مثلاً بین سنگ‌ها در زیر دریا مخفی می‌کند و با حرکت دادن پاهایش باعث می‌شود آب و اکسیژن کافی به تخم‌ها برسد. هشت‌پا تا زمان تولد نوزادانش در کنار تخم‌ها می‌ماند و حتی برای سیر کردن شکم خود نیز از آن جا تکان نمی‌خورد. پس از شش ماه نوزادها به دنیا می‌آیند، اما مادر فداکار به خاطر خستگی و گرسنگی می‌میرد. مورچه‌ها هم دنیای عجیبی دارند. دانشمندان فهمیده‌اند که مورچه‌ها وقتی به شدت بیمار می‌شوند، لانه را ترک می‌کنند. دانشمندان در آزمایشگاه، یک ماده‌ی سمّی را به لانه‌ی مورچه‌ها وارد کردند تا بعضی از آن مورچه‌ها را بیمار کنند. آنها دیدند که مورچه‌های کارگر بیمار از لانه‌شان دور شدند و در تنهایی مردند. علت این کار مورچه‌های کارگر این بود که سایر مورچه‌ها را مریض نکنند. نکته‌ی جالب‌تر این بود که آنها این کار را داوطلبانه انجام دادند، یعنی هیچ مورچه‌ای مجبورشان نکرد که این کار را انجام دهند.

۲۵- کدام گزینه به معنی واژه‌ی سِفْت در خط چهارم نزدیک‌تر است؟

الف) کُلُفت ب) قوی ج) محکم د) خاکی

۲۶- هشت‌پا فقط وقتی می‌خواهد غذا بخورد نوزادانش را ترک می‌کند.

الف) درست ب) نادرست

۲۷- طبق متن، کدام گزینه درباره‌ی خرگوش‌ها، هشت‌پاها و مورچه‌ها درست است؟

الف) دانشمندان می‌خواستند همه‌ی مورچه‌ها را بیمار کنند.

ب) بعد از تولد خرگوش‌ها، موهای مادرشان می‌ریزد.

ج) مورچه‌های مریض مجبور نیستند لانه را ترک کنند.

د) هشت‌پا تخم‌هایش را روی سنگ‌ها می‌گذارد.

۲۸- واژه‌ی آنها در خط نهم به چه کسی اشاره می‌کند؟

الف) مورچه‌های کارگر

ب) مورچه‌ها

ج) مورچه‌های بیمار

د) دانشمندان