

طراحی و اجرای مقدماتی آزمون تعیین سطح درک خوانداری بزرگسالان: شواهدی از دانشجویان کارشناسی ارشد دانشگاه گیلان^۱

سارانا قوامی لاهیج^۲

مریم دانای طوس^۳

عبدالرضا تحریری^۴

علی ربیع^۵

تاریخ دریافت: ۱۳۹۶/۱۲/۲۲

تاریخ پذیرش: ۱۳۹۷/۰۴/۲۵

چکیده

درک خوانداری، حاصل فرآیند تعامل خواننده با زبان نوشتار و در نهایت دریافت معنای متن است. پژوهش‌ها نشان می‌دهند که بین خوانندگان بزرگسال ماهر و ضعیف، از جنبه کاربرد راهبردهای زیربنایی درک خوانداری از جمله راهبردهای شناختی و فراشناختی تفاوت‌های بسیاری وجود دارد. باید توجه

^۱ شناسه دیجیتال (DOI): 10.22051/jlr.2018.19634.1522

^۲ دانشجوی دکتری زبان‌شناسی، دانشگاه آزاد اسلامی، تهران جنوب؛ st_s_ghavamilahij@azad.ac.ir

^۳ دکترای تخصصی زبان‌شناسی، دانشیار گروه زبان‌شناسی، هیأت علمی دانشگاه گیلان (نویسنده مسئول)؛

maryam.dana@guilan.ac.ir

^۴ دکترای تخصصی آموزش زبان انگلیسی، استادیار گروه زبان انگلیسی، هیأت علمی دانشگاه گیلان؛

atahriri@guilan.ac.ir

^۵ دکترای تخصصی زبان‌شناسی، استادیار گروه زبان‌شناسی، هیأت علمی دانشگاه آزاد اسلامی، تهران جنوب؛

a_rabi@azad.ac.ir

داشت، ابزاری مناسب برای سنجش سطح درک خوانداری بزرگسالان در داخل کشور و معیارهای مناسب درانتخاب متن‌های مورد نیاز چنین ابزاری وجود ندارد. بر این مبنای هدف پژوهش حاضر، ساخت ابزار سنجش سطح درک خوانداری بزرگسالان فارسی زبان بود. این مطالعه، پژوهشی ترکیبی بود که با هدف پاسخگویی به سه پرسش اصلی در زمینه معیارهای انتخاب متن برای آزمون‌های درک خوانداری، کاربری معیارهای منتخب در ساخت آزمون مناسب برای سنجش درک خوانداری بزرگسالان و بررسی روایی و پایایی آزمون ساخته شده انجام شد. یافته‌های پژوهش نشان داد متن‌های منتخب باید متناسب با الگوهای آزمون‌های بین‌المللی و اصول زبان‌شناسی متن باشد. برای بررسی روایی، ابزار ساخته شده در اختیار استاد‌های متخصص قرار گرفت و پس از انجام اصلاحات، روایی محتوایی آن تأیید شد. برای بررسی پایایی و تحلیل گویه‌ها، این ابزار در دو مرحله روی نمونه‌ای ۶۰ نفره از دانشجویان کارشناسی ارشد اجرا شد. یافته‌های به دست آمده، نشان‌دهنده سطح مطلوب پایایی آزمون (۰/۸۲) و (۰/۸۴) بود. پس از هنجاریابی، می‌توان از این ابزار در سنجش سطح درک خوانداری بزرگسالان و در برنامه‌ریزی‌های آموزشی برای انتخاب محتوای آموزشی بهره گرفت.

واژه‌های کلیدی: درک خوانداری، ساخت آزمون، بزرگسالان، دانشجویان کارشناسی ارشد

۱. مقدمه

خواندن به معنای جامع و علمی، ابزاری است که می‌توان با آن به گنجینه بی‌انتهای تجربه بشری دست یافت. بر این اساس، خواندن با فهم مطلب و پیامی پیوند می‌خورد که نویسنده با نشانه‌ها و رمزهای کلامی نوشته‌است. همچنین خواندن دربرگیرنده مجموعه پیچیده‌ای از مهارت‌ها شامل بازشناسی واژه‌ها و پاره‌گفته‌ها و هماهنگ‌سازی معنا با کل متن است. گریب و استولر (Grabe & Stoller, 2011, p. 13)، این مهارت‌ها را به گونه‌های مهارت‌های سطح بالا و مهارت‌های سطح پائین دسته‌بندی می‌کنند. مهارت‌های سطح پائین، معمولاً به پردازش‌های زبانی شامل بازشناسی واژه، تجزیه نحوی و رمزگشایی معنایی گفته می‌شود که مهارت‌محور هستند. این در حالی است که مهارت‌های سطح بالا، دربرگیرنده درک و راهبردهای آن است که در کنار هم مراحل خواندن را تکمیل می‌کنند. درک خوانداری، فرایندی چندمنظوره است. به گونه‌ای که در صورت شکل‌گیری ارتباط درست میان خواننده و متن و در پی آن فهم پیام مورد نظر، تحقق می‌یابد. هر

متن، مشتمل بر معنایی است که خواننده باید به کمک دانش و تجربه‌های پیشین خود به آن پی ببرد. هنگامی که متنی خوانده می‌شود، تجربیات خواننده بر درک یا برداشت متن تأثیر می‌گذارد و معنا استخراج می‌شود. طی فرآیند درک، از یک سوی مجموعه‌ای از اطلاعات انتقال می‌یابد و واژه‌هایی نیز به دایره واژگان خواننده افزوده می‌شود. از سوی دیگر، خواننده با یادگیری واژه‌های جدید، دانش نو و ساختارهای مختلف زبانی، می‌تواند دیگر مهارت‌های زبانی را در خود تقویت کند. بر مبنای همین وجوه چندگانه، پژوهشگرانی مانند آلن و هانکوک (Allen & Hancock, 2008, p. 125)، خواندن را مشتمل بر روابط پیچیده‌ای می‌دانند. این روابط درک و واژه، رابطه میان واژه‌ها و مفاهیم، دریافت ایده اصلی متن، تشخیص هدف نویسنده، ارزیابی متن و قضاوت درباره آن را شامل می‌شود. به باور آن‌ها، به دلیل همین پیچیدگی‌هاست که مهارت خواندن و درک خوانداری، توجه علاقه‌مندان بسیاری را به پژوهش در حوزه‌های گوناگون آن کشانیده‌است.

ضعف در پیش‌نیازهای درک متن و در نظر نگرفتن راهبردهای مناسب آن، به هنگام خواندن انواع متفاوت متن از جمله مشکلات دانش آموزان و دانشجویان در دوره‌های تحصیلی مختلف است (Atkinson, 2014, p. 3). یادگیری خواندن فرآیندی طولانی مدت است و در انتهای مراحل یادگیری خواندن است که خواننده بزرگسال به راحتی می‌تواند متن‌های گوناگون را بخواند و معنای مورد نظر را از آن برداشت کند. پژوهش‌ها نشان می‌دهند که بین خوانندگان ماهر و ضعیف از جنبه کاربرد راهبردهای زیربنایی درک خوانداری تفاوت بسیاری وجود دارد (Cain et al., 2000, p. 33). پژوهش‌های گوناگونی در پیوند با مشکلات خواندن و درک متن در سطح بین‌المللی انجام شده‌است. برای نمونه، می‌توان به برگزاری آزمون‌های پرلز^۱ و پیزا^۲ اشاره کرد. در سال‌های اخیر در کشور ما نیز تعداد این پژوهش‌ها رو به افزایش است. دلیل این مسئله را شاید بتوان در گزارش چند دوره آزمون استاندارد پرلز در ایران یافت که نشان از ضعف دانش آموزان ایرانی در بسیاری از زمینه‌های درک خوانداری دارد (Mullis et al., 2012). با وجود کسب چنین یافته‌های ضعیفی در آزمون‌های بین‌المللی و با وجود انجام پژوهش‌های مرتبط در داخل کشور (Jabbari & Khademi, 2013; Daneshgar, 2017)، هنوز ابزار مناسبی برای تعیین سطح درک خوانداری افراد به ویژه در گروه بزرگسال وجود ندارد. زندگی در جامعه امروزی نیازمند یادگیری و خواندن متن‌های گوناگون است. با وجود اهمیت این مسئله و برخلاف پیشرفت کمی تعداد دانشجویان دوره تحصیلات تکمیلی در کشور، هنوز معیار مشخصی برای تعیین سطح درک خوانداری بزرگسالان باسواد وجود ندارد. توسعه آموزش عالی در سال‌های اخیر، تحولی علمی

¹ PIRLS: progress in international reading literacy study

² PISA: program for international student assessment

است که با وجود آثار مثبت آن، کاستی‌هایی نیز دارد. یکی از مهم‌ترین این کاستی‌ها، نبود ابزاری مناسب برای سنجش سطح درک خوانداری دانشجویان و تهیه مواد آموزشی مناسب با این سطح است. این پژوهش، سه هدف مشخص را دنبال می‌کند. نخست، این مقاله در پی تعیین معیار(های) انتخاب متن و پرسش‌های مورد نظر برای آزمون تعیین سطح درک خوانداری بزرگسالان است. دوم، پژوهش پیش‌رو به دنبال طراحی و ساخت این آزمون بوده و سوم، این مقاله بر آن است تا روایی و پایایی آزمون ساخته‌شده را مشخص نماید. بر این مبنا، پرسش‌هایی پژوهش از این قرارند؛ نخست اینکه، «معیار(های) انتخاب متن و پرسش‌های مورد نظر برای آزمون تعیین سطح درک خوانداری بزرگسالان چه مواردی هستند؟». دوم آنکه، «کاربست معیارهای به‌دست‌آمده در پاسخ به این پرسش منجر به طراحی چه آزمونی برای تعیین سطح درک متن بزرگسالان می‌شود؟». در پایان، «آیا آزمون طراحی‌شده بر اساس این معیارها روایی و پایایی لازم را دارد؟»

۲. مبانی نظری

در این بخش، مفاهیم زیربنایی درک خوانداری و عوامل زیربنایی در طراحی آزمون‌های درک خوانداری مرور خواهند شد.

۲.۱. مفاهیم زیربنایی درک خوانداری

همان‌گونه که در بخش مقدمه اشاره شد، خواندن به طور کلی، فرآیندی پیچیده و چندوجهی است. خواندن از طریق دو فرآیند پایه انجام می‌شود: نخست، رمزگشایی نمادها و بازشناسی واژه‌ها و دوم، یکپارچه‌سازی معانی واژه‌ها در بافت متن (Gough & Tunmer, 1986, p. 7). نظریه ساده خواندن^۱ گاف و تانمر (Gough & Tunmer, 1986) نیز دقیقاً بر همین مبنا ارائه شده است. اصول این نظریه تا سال‌ها در سیستم‌های آموزشی، راهنمای تدریس درک خوانداری به شمار می‌آمد. بر مبنای این نظریه، خواندن فرآورده مشترک رمزگشایی و درک خوانداری است. در واقع، خواندن موفق، نیاز به هر دوی این اجزاء دارد که هیچ‌یک به تنهایی نمی‌تواند این مهم را انجام دهد. هر چند که اهمیت هر یک از این دو جزء به مرور زمان تغییر می‌کند. در مراحل اولیه یادگیری خواندن، مهارت رمزگشایی یکی از پیش‌بینی‌کننده‌های مهم این مهارت زبانی به شمار می‌آید. هر چند به مرور زمان و در بزرگسالی، درک خوانداری اهمیت بیشتری می‌یابد (Adlof et al., 2010, p. 333). در سال‌های پی‌درپی، بیشتر پژوهش‌ها در این زمینه فقط روی مهارت‌های بازشناسی واژه

¹ simple view of reading (SVR)

و رمز‌گشایی در کودکان و در سال‌های ابتدایی یادگیری سواد متمرکز شده بود. این در حالی است که در سال‌های اخیر این مطالعات بیشتر روی سنین بالاتر یعنی سنین بعد از چیرگی بر مهارت‌های رمز‌گشایی انجام می‌شود (Catts et al., 2006, p. 285; Cain et al., 2004, p. 51).

درک خوانداری از جنبه اکتساب و رشد وابسته به درک زبان گفتاری است و ارتباط بین این دو، دوسویه است. همچنین، همزمان با رشد، افراد در مواجهه با متن‌های چالشی‌تر قرار می‌گیرند. این مسئله، سبب افزایش دایره واژگان و غنی‌تر شدن دانش دایره‌المعارفی و تجربه می‌شود (Nation, 2005, p. 44). تفاوت‌های فردی در دانش واژگانی از سن پیش از مدرسه آغاز می‌شود. همچنین تأثیر وضعیت اجتماعی - اقتصادی بر توانایی کلامی به اثبات رسیده است. به ویژه، دانش واژگانی به گونه‌ای که پس از شروع مدرسه دایره واژگان کودکان به شیوه چشم‌گیری افزایش می‌یابد (Atkinson, 2014, p. 14).

فرا تر از رشد مهارت‌های بازشناسی واژه، برای رسیدن به سطح مطلوب درک خوانداری مشارکت دیگر فرآیندها هم ضروری است. معنای تحت‌اللفظی به ندرت همه مفهوم یک نوشته را بیان می‌کند. بر پایه دیدگاه پرفتی و همکاران (Perfetti et al. 2005, p. 229)، بازنمون ذهنی^۱ ساخته شده از خواندن یک متن (همان مدل موقعیتی^۲ راوسون و کینش (Rawson & Kintsch, 2005) باید به اندازه‌ای غنی شود که فهم عمیق‌تری از معنای ظاهری متن ایجاد کند. مدل‌های موقعیتی، با تلفیق ساختارهای خرد^۳ و کلان^۴ متن با دانش دایره‌المعارفی در ذهن خواننده ساخته می‌شوند که از عوامل دیگری از جمله ظرفیت حافظه فعال و استنتاج‌سازی^۵ اثر می‌پذیرند (Atkinson, 2014, p. 10). پرفتی و همکاران (Perfetti et al., 2005, p. 234) ادعا می‌کنند که به‌روزرسانی بازنمون ذهنی متن، به وسیله کاربرد سه راهبرد در هنگام درک امکان‌پذیر است. این راهبردها مشتمل بر استنتاج‌سازی، نظارت بر درک^۶ و حساسیت به ساختار داستان هستند.

الف) استنتاج‌سازی، توانایی استفاده از دو یا چند بخش از اطلاعات متن برای رسیدن به بخش سوم اطلاعاتی است که به صورت ضمنی در متن مطرح شده است. استنتاج‌سازی می‌تواند تا اندازه تشخیص مرجع یک ضمیر ساده باشد. یا به اندازه‌ای پیچیده باشد که نیازمند رسیدن به معنای تلویحی نهفته در متن بوده که از طریق انتخاب یک واژه خاص توسط نویسنده شکل می‌گیرد و درک آن نیاز به دانش پیشین خواننده دارد. در واقع، می‌توان گفت در هنگام درک، خواننده با

¹ mental representation

² situation model

³ micro-structure

⁴ macro-structure

⁵ inference-making

⁶ comprehension monitoring

انجام استنتاج‌های گوناگون، جمله‌هایی با معنای پنهان و در ظاهر ناپیوسته را به مطلبی پیوسته تبدیل می‌کند. این توانایی خواننده، به نویسنده امکان می‌دهد که در هنگام نوشتن متن به اندازه لازم بنویسد. چرا که در بسیاری موارد نیاز به توضیح همه جزئیات در متن نیست، بلکه فقط زمینه و طرح کلی ارائه می‌گردد و خواننده با استنتاج‌هایش متن را گسترش خواهد داد. بسیاری از پژوهشگران ادعا می‌کنند که راهبرد استنتاج‌سازی، بخش اصلی و مرکزی درک خوانداری است (Tarchi, 2015, p. 81; Silagi et al., 2014, p. 407).

ب) نظارت بر درک، یکی از راهبردهای فراشناختی درک خوانداری است. به وسیله این راهبرد، شخص وضعیت فهم خود از اطلاعات را مورد بررسی قرار می‌دهد. این راهبرد، برای درست‌خوانی ضروری است. پژوهش‌ها نشان می‌دهد که خوانندگان ضعیف اغلب نمی‌توانند بر درک اطلاعات نظارت داشته باشند (Baker & Anderson, 1982, p. 282). هدف از نظارت بر درک، یافتن نقاط ضعف متن در هنگام خواندن است. وقتی خواننده ماهر از عدم فهم مطلب آگاهی می‌یابد، از راهبردهای اصلاحی برای حل این مشکل بهره می‌گیرد. این راهبردهای اصلاحی شامل دوباره‌خوانی، کمک گرفتن از دیگران، استفاده از منابع و متن‌های مرتبط، استنتاج‌های منطقی بر اساس متن و یا بهره‌گیری از دانش پیشین است.

پ) ساختار متن، نوع خاصی از دانش پیشین است که خوانندگان ماهر آن را دارند. خوانندگان ماهر اطلاعات مفید را تشخیص می‌دهند و می‌دانند که چگونه این اطلاعات با دیگر گزاره‌های متن در ارتباط اند. گذشته از این، آن‌ها می‌توانند این ارتباط را در متن‌هایی که به درستی سازمان‌دهی نشده‌اند، نیز برقرار کنند. این خوانندگان، نخست از طریق تجربه خود در زبان شفاهی و سپس به وسیله خواندن، الگوهای شناختی- ساختارهای بلاغی- که پیش‌تر کسب کرده‌اند، را برای درک به کار می‌برند. این الگوها آن‌ها را هدایت می‌کند تا مثلاً تشخیص دهند که یک متن ویژه دو چیز را مقایسه می‌کند، یا مشکلی را مطرح می‌کند و به آن پاسخ می‌دهد. خوانندگانی که می‌توانند ساختار متن را تشخیص دهند بهتر می‌توانند اطلاعاتی که برای درک موفق نیاز دارند را در متن بیابند. انواع متن‌ها روی توجه خواننده بر جنبه‌های گوناگون موقعیت ارتباطی متمرکز داشته و با فعالیت‌های ذهنی گوناگونی مرتبط است. هاتیم و میسون (Hatim & Mason, 1997, p. 33) معتقدند متن‌های تفسیری، مشتمل بر تحلیل و ترکیب مفاهیم است. متن‌های مباحثه‌ای، بر ارتباط مفاهیم یعنی تأیید یک نظریه و ارتباط آن با نظریه مخالف متمرکز است. این نوع متن‌ها با فرآیندهای ذهنی قضاوت سر و کار دارند. متن‌های روایی، زنجیره‌ای از رخ‌دادها را تشکیل می‌دهند که بین آن‌ها رابطه علت و معلولی برقرار است و به صورت زنجیروار پشت سر هم بیان می‌شوند.

۲.۲. عوامل مهم در طراحی آزمون‌های سنجش درک خوانداری

در بخش‌های پیشین، نظریه‌هایی که پیش‌زمینه‌های کسب مهارت در درک خوانداری بعد از فراگیری مراحل ابتدایی خواندن را نشان می‌دهد، شرح داده شد. در این زمینه، شیوه سنجش سطح مهارت خواندن را می‌توان از پیش‌نیازهای یک نظام آموزشی دانست. در دهه‌های اخیر، یکی از بخش‌های مهم زبان‌آموزی، آزمون‌سازی است. در واقع، آزمون‌سازی رابطه میان یک نظریه و بررسی‌های میدانی است. یافته‌های کیفی و کمی به دست آمده از چنین رویکرد عمل‌مداری در حوزه مطالعات زبان‌شناسی کاربردی، در راستای از بین بردن مشکلات و بازسازی ایرادهای آموزشی به کار برده می‌شود (Davis, 2001, p. 134). کیفیت هر گونه ارزیابی آموزشی، به کیفیت ابزار مورد استفاده بستگی دارد. اگر این ابزارها به خوبی طراحی شده باشند، می‌توانند از هدررفت وقت و هزینه جلوگیری نمایند. اما باید دید معیار طراحی آزمون متناسب با هدف این پژوهش چیست؟

کورته و همکاران (Corte et al., 2001, p. 535) سه عامل را در درک خوانداری دخیل می‌داند: مهارت خواننده، هدف خواندن و متن. تعامل درست این سه عامل در کنار هم، می‌تواند منجر به درک درستی از محتوای مورد بررسی شود. جمله‌ها باید به گونه‌ای در کنار هم قرار بگیرند که همپوشانی معنا مابین آن‌ها پدید آید و خواننده بازنمون منسجمی از محتوای متن داشته باشد. به باور کروسلی و همکاران (Crossley et al., 2011, p. 84) هنگامی که تهیه‌کنندگان مواد آموزشی می‌خواهند متن‌های قابل درک طراحی کنند، معمولاً دو دیدگاه کلی ساختاری و شمی^۱ را در نظر می‌گیرند.

دیدگاه ساختاری، برای تهیه متن مبتنی بر استفاده از ساختار و فهرست واژگان است که بر اساس سطح سواد خوانندگان و دایره واژگانی آن‌ها متن‌ها را طراحی می‌کند. دیدگاه دیگری که زیرمجموعه دیدگاه ساختاری به شمار می‌آید، استفاده از فرمول‌های خوانش‌پذیری^۲ است مانند فرمول سهولت خوانش فلش (Flesch, 1948)، فرمول دیل-چال (Dale-Chall, 1948)، فرمول فوگ (Gunning, 1952)، فرمول فرای (Fry, 1968) و موارد مشابه است. این فرمول‌های خوانش‌پذیری، آماری از متن به دست می‌دهند که بر اساس تعداد واژه‌ها و اندازه جمله‌ای متن است. بر خلاف دیدگاه ساختاری، دیدگاه شمی موضوعی تر است و فقط روی حس طبیعی نویسنده از درک‌پذیری متن و پردازش گفتمان پایه‌ریزی می‌شود (Crossley et al., 2011, p. 85).

¹ intuitive

² readability

با وجود اینکه فرمول‌های خوانش‌پذیری، به طور گسترده‌ای در بررسی‌های تحلیل متن‌های مناسب آموزشی به کار گرفته شده‌اند، اما دو معیار ساده تعداد واژه و اندازه جمله معیارهایی سطحی و بسیار ساده به شمار آیند. این فرمول‌ها بسیاری از اجزای زبانی و گفتمانی مانند انسجام^۱ و پیوستگی^۲ را نادیده می‌گیرند که از جنبه نظری بر روی دشواری درک متن تأثیرگذار هستند (Crossley et al., 2011, p. 88). بی‌تردید، هنگامی که متنی مشتمل بر واژه‌ها و جمله‌های طولانی‌تر باشد، خواندن آن به تلاش بیشتری نیاز دارد. زیرا واژه‌های طولانی در زبان چندان رایج نیستند و فراوانی کمتری دارند. جمله‌های طولانی‌تر هم، نیاز به چالش بیشتر حافظه فعال دارد و شاید سخت‌تر به نظر برسند. بنابراین نمی‌توان انکار کرد که پارامترهای تعداد واژه و اندازه جمله مهم هستند. هر چند فقط با در نظر گرفتن این دو عامل نمی‌توان سطح خوانش‌پذیری متن را نشان داد.

جنبه‌های زبان‌شناختی متن از دیدگاه‌های مختلفی بررسی شده‌است. از جمله می‌توان به دستور متن (Van Dijk, 1972)، زبان‌شناسی متن (De Beaugrande & Dressler, 1981) و تحلیل گفتمان (Brown & Yule, 1983) اشاره کرد. هدف دستور متن ون‌دایک (Van Dijk, 1972)، طراحی مدلی برای ساختارهای دستوری متن‌ها است (همانند دیدگاه گشتاری چامسکی). تحلیل گفتمان، تحلیل گفتار و متن را با توجه به بیناکنش‌ها و کارکردهای اجتماعی مورد بررسی قرار می‌دهد. زبان‌شناسی متن به چگونگی ایجاد متن و ویژگی‌های بافتی آن می‌پردازد. بوگراند و درسلر (De Beaugrande & Dressler, 1981) به عنوان پایه‌گذاران زبان‌شناسی متن، هفت ویژگی را برای «متن» برمی‌شمرند که از این قرارند: ۱- انسجام^۲ - پیوستگی^۳ - قابلیت پذیرش^۴ - هدفمندی^۵ - پیام‌رسانی^۶ - موقعیت‌مداری^۷ - بینامتنیت^۸ (Alborzi, 2013, p. 182).

از میان ویژگی‌های بالا، دو ویژگی «انسجام» و «پیوستگی»، ویژگی‌های متن‌محور هستند. به این معنا که مربوط به خود متن می‌شوند و از آراء هلییدی و حسن (Halliday & Hasan, 1976) تأثیر پذیرفته‌اند. پنج ویژگی دیگر، رویکردی کاربردی نسبت به متن دارند و درباره نحوه ارتباط و تعامل متن و خواننده با یک‌دیگر است (Amiri Khorasani & Alinejad, 2015). منظور از پیوستگی، تفسیرپذیری متن و منظور از انسجام، استفاده درست از ابزارهای

¹ cohesion

² coherence

³ acceptability

⁴ intention

⁵ informativity

⁶ situationality

⁷ intertextuality

زبانی مثل عناصر ارجاعی (ضمایر، عناصر اسمی، عناصر اشاره و عناصر ملکی)، تکرار، حذف، جایگزینی، روابط گزاره‌ها و نشانگرهای موضوعی است. پیوستگی بیشتر با عناصر گفتمانی در ارتباط است. برای نمونه، دو جمله «اینجا بهار است. پرنده‌ها آواز می‌خوانند» پیوسته است اما هیچ رابطی جمله اول را به جمله دوم پیوند ندهد است (Ehrlich, 1991, p. 169).

پیوستگی و انسجام در مدل درک خوانداری کینش و وندایک (Kintsch & van Dijk, 1978) به صورت درک معنا در نموداری پیوسته از گزاره‌ها تعبیر می‌شود. گزاره یک واحد منطقی-معنایی^۱ است که از یک محمول^۲ و یک یا چند موضوع^۳ تشکیل شده است. دو گزاره در صورتی پیوسته به شمار می‌آیند که موضوع مشترکی داشته باشند. این قانون همپوشانی موضوع‌ها سبب ایجاد پیوستگی محلی ساختارهای خرد متن می‌شود. در سطح کلی‌تر، گزاره‌ها در ساختارهای کلان سازمان‌دهی می‌شوند. به دلیل ظرفیت محدود حافظه فعال، پردازش‌ها در چرخه‌های متوالی انجام می‌شوند. هنگامی که بخشی پردازش می‌شود، برخی گزاره‌ها انتخاب می‌شوند و در محفظه موقت نگهداری می‌شوند که بخشی از حافظه فعال است. در صورتی که موضوع مشترکی مابین یکی از گزاره‌ها در چرخه فعلی و یکی از گزاره‌ها در محفظه موقت باشد، این گزاره‌ها برای ایجاد انسجام به شیوه مستقیم و خودکار عمل می‌کنند. در این میان، اگر موضوع مشترکی وجود نداشته باشد، به دنبال گزاره‌هایی می‌گردد که پیش‌تر پردازش شده و در حافظه بلند مدت ذخیره شده‌است و آن را دوباره در حافظه فعال بارگزاری و پردازش می‌کند. در صورتی که جستجو موفق نباشد، برای فهم معنا از استنتاج‌سازی بر مبنای دانش دایره‌المعارفی استفاده می‌شود. به نظر می‌رسد از دیدگاه کینش و وندایک (Kintsch & van Dijk, 1978) بارگزاری دوباره یک گزاره در حافظه فعال و ایجاد استنتاج نیاز به زمان و منابع شناختی بیشتری دارد. در نتیجه به باور آن‌ها، پردازش متن‌هایی که نیاز به استنتاج‌سازی‌های مکرر دارند، مشکل‌تر است. آنچه از مرور این نظریه‌ها می‌توان دریافت این است که انسجام و پیوستگی در تعیین ساختار متن، وجه مشترک همه نظریه‌هاست.

جدا از انسجام و پیوستگی، در تحلیل یک متن، سه ساختار کلی مورد توجه قرار می‌گیرد: ساختار خرد متن، ساختار کلان متن و آبرساختار^۴. ساختار خرد، سازمان‌دهی درست بین جمله‌ها، بندها و پاره‌گفته‌ها است. تشخیص ساختارهایی مانند ارجاع، ساخت محذوف^۵ و عناصر پیوندی را تحلیل ساختار خرد متن می‌نامند. ساختار کلان، توصیفی از مفهوم کلی متن است که در قالب

^۱ logicosemantic

^۲ predicate

^۳ argument

^۴ superstructure

^۵ elliptical form

روابط بین پاراگراف‌ها از جنبه مفهومی نمود می‌یابد. آبرساختار، قالب نگارشی خاص هر متن است. برای نمونه، قالب نگارش متن یک مقاله علمی با داستان و یا نامه‌ای دوستانه متفاوت است و حتی قالب نگارش یک نامه رسمی با نامه دوستانه فرق دارد (Demirgünes, 2017, p. 582).

عناصر انسجام متن بر اساس آراء هلیدی و حسن (Halliday & Hasan, 1976) پنج نوع هستند که عبارتند از ارجاع، جایگزینی، حذف، ربط دهنده‌ها و انسجام واژگانی. این عناصر در قالب سه مقوله انسجام واژگانی، دستوری و پیوندی دسته‌بندی می‌شوند. انسجام واژگانی به صورت مترادف‌ها، مشابهت معنایی و شمول معنایی نمود می‌یابد. انسجام دستوری در قالب ارجاع، حذف و جایگزینی در متن شکل می‌گیرند. انسجام پیوندی در قالب ربط دهنده‌ها به چهار دسته افزایشی^۱ (و، همچنین، به علاوه، به همین ترتیب و موارد مشابه)، تبیینی^۲ (ولی، علی‌رغم، بلکه، برعکس و مواردی از این قبیل)، علی (زیرا، به دلیل اینکه و موارد مشابه) و زمانی^۳ (آنگاه، آن وقت، پس از آن و مواردی از این قبیل) دسته‌بندی می‌شوند.

با وجود مطالب اشاره شده، باید مشاهده کرد در آزمون‌های استاندارد خواندن در سطح بین‌المللی چه معیارهایی برای انتخاب متن‌ها و پرسش‌ها در نظر گرفته می‌شود. این آزمون‌های استاندارد خواندن در سطح بین‌المللی مانند آزمون پرلز و پیزا برای سنجش سطح درک خوانداری افراد در زبان اول آن‌ها اجرا می‌شوند و بخش درک متن آزمون تافل که برای سنجش سطح درک خوانداری در زبان انگلیسی (به عنوان زبان خارجی) اجرا می‌شوند. در ادامه، روند انتخاب متن و پرسش این آزمون‌ها به صورت مختصر ارائه می‌شود:

۲.۱. روند انتخاب متن و سؤال در آزمون پرلز

آزمون پرلز، برای سنجش خواندن و درک خوانداری دانش‌آموزان ۹ تا ۱۰ ساله، به وسیله مرکز بین‌المللی پیشرفت تحصیلی^۴ و در سطح کشورهای جهان از سال ۲۰۰۱ ارائه شده‌است. این آزمون هر پنج سال یک‌بار برگزار می‌گردد. همچنین، این آزمون درک خوانداری دانش‌آموزان را در کاربرد چهار راهبرد «بازیابی اطلاعاتی که به صورت روشن بیان شده (معنای تحت‌اللفظی)»، «استنتاج‌سازی‌های مستقیم»، «تفسیر و یکپارچه‌سازی مفاهیم» و «ارزیابی محتوای متن از منظر انتقادی» مورد بررسی قرار می‌دهد.

¹ additive

² adversative

³ temporal

⁴ IEA: international association for the evaluation of educational achievement

موفقیت در انجام تکالیف این آزمون، نیاز به راهبردهای شناختی و فراشناختی دارد تا علاوه بر تشخیص معانی و مفاهیم، خواننده بتواند فهم خود از معنا را مورد ارزیابی قرار دهد. همچنین، دانش و تجربه‌های پیشین که خوانندگان به متن وارد می‌کنند، آن‌ها را برای فهم زبان و متن مجهز می‌سازد (Kintsch, 2005, p. 52).

تهیه متن در آزمون‌هایی که به صورت بین‌المللی اجرا می‌شوند، امر بسیار پیچیده‌ای است. در واقع، به سبب تفاوت در دوره‌های تحصیلی و فرهنگ کشورهای شرکت‌کننده انتخاب متن‌ها با محدودیت‌هایی روبه‌رو است. برای نمونه، ترجمه شعر امری بسیار دشوار است و نمایش‌نامه‌ها نیز برای سن‌های پائین مناسب نیستند. سبک اصلی متن‌های مورد استفاده در آزمون پرلز، بیشتر از گونه اطلاعاتی و با موضوع‌های علمی، تاریخی، جغرافیایی یا اجتماعی تهیه می‌شوند. این متن‌ها، در نحوه سازمان‌دهی محتوا متنوع و متفاوت هستند. برای نمونه، حقایق تاریخی بر اساس زمان، دستورالعمل‌ها به صورت مرحله‌به‌مرحله و مبحث‌ها به صورت منطقی ارائه می‌شوند. این متن‌ها به‌طور ویژه، برای این آزمون به وسیله نویسندگان نوشته می‌شوند که از شیوه نگارش برای مخاطب با سن پایین آگاهی دارند و در فراخوانی بین‌المللی، توسط کشورهای شرکت‌کننده تهیه می‌شوند (Mullis et al., 2016). برای انتخاب متن‌های آزمون پرلز به‌طور کلی چهار ویژگی ارائه شده است: نخست، وضوح، رعایت ساختارهای انسجام و پیوستگی متن. دوم، محتوای متناسب با فرهنگ کشورهای مختلف. سوم، محتوای جالب و سرگرم‌کننده برای گستره وسیعی از دانش‌آموزان و چهارم، پایه مناسب برای ارزیابی طیف کامل مراحل درک متن.

به سبب محدودیت‌های زمانی و شرایط آزمون، برخی محدودیت‌ها در تعداد واژه‌های متن‌ها اعمال می‌شود تا دانش‌آموزان در مدت آزمون، زمان کافی برای خواندن کل متن‌ها و پاسخ به پرسش‌های درک متن داشته باشند. متن‌های پرلز به صورت الکترونیکی (ای پرلز^۱) و یا چاپی، مشتمل بر ۴۰۰ تا ۸۰۰ واژه هستند. هر چند می‌تواند با توجه به ویژگی‌های متن تاحدودی متغیر باشد. همچنین، به سبب شرایط سنی، آزمودنی‌ها بیشتر متن‌ها با واژگان عینی نگارش می‌شوند و از ساختارهای پیچیده دستوری پرهیز می‌شود.

۲.۲.۲ روند انتخاب متن و سؤال در آزمون پیزا

نخستین بار آزمون پیزا، به وسیله سازمان همکاری و توسعه اقتصادی^۲ در سال ۲۰۰۰ به اجرا درآمد و هر سه سال یک‌بار برای دانش‌آموزان ۱۵ ساله برگزار می‌شود. سواد خواندن در این آزمون، به

^۱ E-PIRLS

^۲ OECD: organization for economic co-operation and development

معنای فهم، کاربرد و بازنمایی متن به منظور رسیدن به هدف‌ها، پیشرفت دانش و توانایی و مشارکت در اجتماع است. در پیوند با طراحی متون برای این آزمون اعتقاد بر این است که هر متن باید مشتمل بر اطلاعات کافی برای خواننده‌ها باشد تا در درک معنا دچار مشکل نشوند. هر نوع ارزیابی درک، نیاز به متن‌های ویژه‌ای دارد و هرگز طبقه‌بندی مشترکی برای انتخاب متن‌ها وجود ندارد. متن‌های برگزیده این آزمون گونه‌های مختلف متون‌های توصیفی روایی، توضیحی، استدلالی و دستورالعمل را شامل می‌شود که برای طراحی تکالیف درک خوانداری به کار برده می‌شود. آن چه در انتخاب متون این آزمون نیز گزارش شده است رعایت ویژگی‌های انسجام و پیوستگی در متن است که از طریق شاخص‌های زبان‌شناسی متن مشخص می‌شود (Adams & Cresswell, 2016, p. 32). اولویت‌های طراحی پرسش‌های این آزمون مشتمل بر «تکالیف سهولت و کارآیی خواندن»، «تکالیف مهارت سریع خواندن و پیدا کردن نکات کلیدی»، «استنتاج‌سازی» و «ارزیابی و انعکاس (سنجش فراشناخت)» است.

۲.۲.۳. روند انتخاب متن و پرسش در بخش درک مطلب آزمون نافل آی‌بی‌تی (TOFEL IBT, 2010)

بخش درک مطلب آزمون نافل، توانایی آزمودنی‌ها برای فهم متن‌ها در سطح دانشگاهی به زبان انگلیسی را مورد ارزیابی قرار می‌دهد. آزمودنی‌ها ۳ تا ۵ متن با اندازه تقریبی ۷۰۰ واژه را می‌خوانند و به پرسش‌های مربوط پاسخ می‌دهند. اگرچه نوع گویه‌ها و مهارت‌های فرعی در بیشتر آزمون‌های خواندن مشابه است، اما تعداد و نوع آن‌ها متغیر است و بسته به قضاوت نویسنده و طراح می‌تواند متغیر باشد. به طور متوسط هر متن ۱۰۰ واژه‌ای نیاز به سه پرسش دارد. متن‌ها مشتمل بر همه اطلاعاتی است که خواننده برای پاسخ‌گویی به پرسش‌ها به آن‌ها نیاز دارد. پرسش‌ها اغلب برای ارزیابی توانایی خواننده در درک اطلاعات واقعی، معنای واژه‌ها درون بافت و هدف نویسنده است. پرسش‌های این بخش در قالب ارزیابی‌های زیر مطرح می‌شود:

- پرسش‌های مربوط به ایده اصلی: تشخیص عنوان، تشخیص موضوع، تشخیص ایده اصلی
- تشخیص ساختار مفاهیم: ارتباط پاراگراف‌ها با یکدیگر
- تشخیص جزئیات متن
- تشخیص مرجع ضمیر
- استنتاج‌سازی که نیاز به دانش دایره‌المعارفی خواننده دارد
- واژگان
- تشخیص لحن و هدف نویسنده

دی و پارک (Day & Park, 2005)، در پژوهشی با موضوع طراحی پرسش‌های درک خوانداری با بررسی همه نظریه‌ها و انواع پرسش‌های درک خوانداری الگویی برای سنجش انواع سطوح راهبردهای درک خوانداری طراحی کردند (جدول ۱).

جدول ۱: طراحی پرسش‌ها و سطح‌های سنجش راهبردهای درک خوانداری دی و پارک (Day&Park, 2005)

انواع سطوح راهبردهای درک خوانداری						نوع پرسش‌ها
پاسخ‌های شخصی ^۴	ارزیابی	پیش‌بینی ^۳	استنتاج‌سازی	سازمان‌دهی مجدد ^۲	تحت‌الفظی ^۱	
						بله/خیر
						دو شقی ^۵
						درست/نادرست
						پرسش‌های تشریحی
						چندگزینه‌ای

- بر این اساس انواع سطوح سنجش راهبردهای درک خوانداری عبارت‌اند از:
- درک تحت‌الفظی: درک تحت‌الفظی به معنای فهمیدن مستقیم معنای متن است. مانند درک حقایق موجود در متن، تشخیص معنای واژگان، پیدا کردن تاریخ‌ها و مکان‌ها. این پرسش‌ها در بیشتر مواقع به راحتی و به روشنی قابل پاسخ‌گویی هستند.
 - سازمان‌دهی دوباره: این پرسش‌ها به گونه‌ای طراحی می‌شوند که برای پاسخ‌گویی نیاز به درآمیختن اطلاعات از نقطه‌های گوناگون متن دارد. این گونه پرسش‌ها در مقایسه با پرسش‌های تحت‌الفظی چالشی‌تر به شمار می‌آیند.
 - استنتاج‌سازی: این پرسش‌ها، فراتر از دریافت معنای تحت‌الفظی متن هستند. بنابراین، معمولاً نیاز به زمان بیشتری برای پاسخ‌گویی دارند. زیرا اطلاعات مورد نیاز برای پاسخ‌گویی به این پرسش‌ها مستقیماً در متن بیان نشده‌است. استنتاج، نیازمند درآمیختن معنای تحت‌الفظی با دانش پیشین و شمّ زبانی خواننده است.
 - پیش‌بینی: چهارمین گونه از پرسش‌های درک خوانداری، پیش‌بینی است. پیش‌بینی مشتمل بر استفاده از اطلاعات متن و دانش پیشین آزمودنی به شیوه‌ای اصولی است برای تشخیص آنچه پیش یا پس از متن اتفاق می‌افتد.

¹ literal

² reorganization

³ prediction

⁴ personal response

⁵ alternative

- ارزیابی: در این گونه پرسش‌ها، خواننده قضاوتی کلی و جامع در مورد جنبه‌های گوناگون متن از جمله لحن و ساخت متن انجام می‌دهد. برای پاسخ‌گویی به این پرسش‌ها، آزمودنی باید از معنای تحت‌اللفظی متن و دانش پیشین در مورد موضوع و موضوع‌های مرتبط با متن استفاده کند تا بتواند ارزیابی عادلانه‌ای از متن ارائه دهد.
- پاسخ شخصی: در این گونه پرسش‌ها، از خواننده درخواست می‌شود تا احساس خود را نسبت به متن و موضوع ارائه کند. پاسخ‌های این بخش متغیر است و هیچ پاسخی نادرست نیست، اما باید در چهارچوب متن مورد نظر باشد.

۳. پیشینه تجربی پژوهش

در این بخش پیشینه‌ای از پژوهش‌های تجربی انجام شده در خارج و داخل کشور در زمینه درک خوانداری ارائه می‌شود.

۳.۱. پیشینه پژوهش‌های خارجی در پیوند با آزمون‌های تعیین سطح درک خوانداری

دی و پارک (Day & Park, 2005) گونه‌های مختلف پرسش‌هایی را که می‌توان برای سطح‌های شش‌گانه ارزیابی درک خوانداری - که در بالا به آن‌ها اشاره شد - به کار برد، را به پنج دسته طبقه‌بندی کرده‌اند. این پرسش‌ها مشتمل بر پرسش‌های بله/خیر، پرسش‌های دو شقی، پرسش‌های درست/نادرست، پرسش‌های تشریحی و پرسش‌های چندگزینه‌ای هستند. در ادامه، هر یک از این گونه پرسش‌ها با جزئیات بیشتری شرح داده می‌شوند.

- پرسش‌های بله/خیر، پرسش‌هایی ساده و رایج در میان پرسش‌های درک خوانداری به ویژه برای سنجش آزمودنی‌های با سنین پایین‌تر هستند. برای هر پاسخ، ۵۰ درصد احتمال درست بودن وجود دارد. هر شش سطح ارزیابی درک خوانداری را می‌توان با سؤالات بله/خیر ارائه کرد. این گونه پرسش‌ها، ساده‌ترین نوع طراحی پرسش‌های درک هستند.
- پرسش‌های دو شقی دو یا چند پرسش بله/خیر است که با یک «یا» به هم پیوند می‌خورند: برای نمونه، آیا این مقاله در مورد طراحی پرسش‌های درک خوانداری است یا در مورد راهبردهای درک خوانداری؟ پرسش‌های دوشقی را می‌توان برای سطح‌های تحت‌اللفظی، سازمان‌دهی دوباره، استنتاج‌سازی و پیش‌بینی به کار برد. هر چند برای پرسش‌های بخش ارزیابی و پرسش‌های شخصی کاربرد چندانی ندارد.
- پرسش‌های درست/نادرست، استفاده فراوانی در کتاب‌های خواندن دارند. نقطه ضعف این گونه پرسش‌ها ۵۰ درصد احتمال پاسخ‌گویی است که برای آزمون‌های مهمی مانند تعیین

سطح مناسب نیستند. همچنین طراحی جمله‌های نادرست باید با دقت ویژه‌ای دنبال شود تا بتواند نادرست فهمیدن متن توسط آزمودنی را به خوبی نشان دهد.

- پرسش‌های تشریحی برای بررسی معنای تحت‌اللفظی متن بسیار مناسب هستند. همچنین، برای پاسخ به پرسش‌های ارزیابی و پرسش‌های شخصی نیز کاربرد بسیاری دارند. هر چند به طور کلی، پاسخ‌گویی به این نوع پرسش‌ها، چه برای دانش‌آموزان با سن کم و چه برای بزرگسالان بسیار خسته‌کننده‌اند. همچنین برای تصحیح‌کننده نیز تا حدودی مشکل‌آفرین است.
- پرسش‌های چندگزینه‌ای در مقایسه با چهار نوع دیگر، بهترین گونه پرسش‌ها برای ارزیابی درک خوانداری به شمار می‌آیند. این نوع پرسش‌ها، در بیشتر موارد فقط یک پاسخ درست دارند. همه انواع سطح‌هایی که در بالا شرح داده شد، را می‌توان در قالب این گونه پرسش‌ها مطرح کرد. طراحی پرسش‌های چندگزینه‌ای نیاز به دقت بسیار بالایی دارد. در واقع، تمام گزینه‌ها باید به گونه‌ای طراحی شوند که تا اندازه‌ای بتوانند توجه خواننده را جلب کرده و تفکر او را نسبت به متن برانگیزند.

۲.۳. پیشینه پژوهش‌های ایرانی درباره آزمون‌های تعیین سطح درک خوانداری

در سال‌های اخیر در کشور ما تعداد پژوهش‌ها در زمینه درک خوانداری رو به افزایش است. دلیل آن را شاید بتوان در گزارش چند دوره آزمون پرلز در ایران یافت. بر پایه یافته‌های این آزمون‌ها، وضعیت دانش‌آموزان ایرانی در بخش درک خوانداری و در پاسخ‌گویی به پرسش‌هایی که جنبه استنباطی و استنتاجی داشته، بسیار سطح پائین بوده‌است. ایران در آزمون پرلز در سواد خواندن در سال ۲۰۰۱ رتبه ۳۲ از ۳۵، در سال ۲۰۰۶ رتبه ۴۰ از ۴۵ و در سال ۲۰۱۱ رتبه ۳۸ از ۴۹ را داشته‌است. عملکرد دانش‌آموزان ایرانی در پرلز ۲۰۰۱ از ۴۱۴ به ۴۵۷ در سال ۲۰۱۱ رسیده‌است که با وجود پیشرفت هنوز به میانگین مقیاس پرلز (۵۰۰) نرسیده‌است (Mullis et al., 2012).

عزیزیان و عابدی (Azizian & Abedi, 2006)، آزمونی تشخیصی برای سنجش سطح خواندن دانش‌آموزان پایه سوم ابتدایی شهر اصفهان طراحی کردند. این آزمون شامل ۸ خرده‌آزمون در سه زمینه درک خواندن، درک مطلب و آگاهی واج‌شناختی بود. با وجود جامعیت این آزمون معیارهای انتخاب متن‌ها مشخص نیست. رقیب‌دوست و اصغرپور (Raghibdoust & Asgharpour, 2009)، در پژوهشی، رابطه بین دانش پیشین و انگیزش خواندن را با درک خواندن فارسی‌آموزان مورد بررسی قرار دادند. به این منظور، آزمونی با چهار متن و برای هر متن پنج پرسش ساخته شد. همچنین دو پرسش‌نامه انگیزش خواندن و دانش پیشین نیز توسط پژوهشگران طراحی شد. پایایی آزمون درک خواندن ۰/۸۰ در مرحله مقدماتی محاسبه شد. در این آزمون،

معیار انتخاب متن‌ها، تناسب سطح دشواری واژگانی و اطلاعاتی متن‌ها با کتاب درسی آزمودنی‌های پژوهش در نظر گرفته شد که همگی از غیرفارسی‌زبانان بودند. یافته‌های تحلیل داده‌های آن‌ها نشان داد که هر دو متغیر انگیزش خواندن و دانش پیشین رابطه معناداری با درک خواندن داشتند که نقش دانش پیشین در این زمینه قوی‌تر بود.

فنسولی (Ghonsooly, 2010)، آزمونی برای بررسی بسندگی زبان فارسی برای فارسی‌آموزان خارجی طراحی کرد. این آزمون همانند آزمون تافل برای زبان انگلیسی طراحی شده بود که شامل آزمون‌های شنیدن، خواندن و دستور بود. یافته‌های این پژوهش نشان داد که فارسی‌آموزان خارجی در بخش دستور، بیشتر از سایر بخش‌ها مشکل دارند. شکبیا و اسدزاده (Shakiba & Asadzade, 2013)، در پژوهش خود، تأثیر استفاده از انسجام متن در سطح خرد و کلان را بر بهبود درک مطلب دانش‌آموزان دوره متوسطه مورد بررسی قرار دادند. بر اساس یافته‌های آن‌ها متن‌های دارای انسجام خرد و کلان بالا، بیشترین تأثیر را بر درک مطلب دانش‌آموزان داشتند. جباری و خادمی (Jabbari & Khademi, 2013)، در پژوهشی، آزمونی تشخیصی برای سنجش خواندن و درک مطلب دانش‌آموزان ابتدایی طراحی کردند. این آزمون مشتمل بر ۱۱ خرده آزمون سنجش عوامل زیربنایی خواندن و درک بود. اگر چه این خرده آزمون‌ها را می‌توان یک گام مناسب برای سنجش مهارت‌های زیربنایی خواندن دانست، اما در این پژوهش نیز ملاک‌های انتخاب متون به درستی معرفی نشده بود.

آقاگل‌زاده و میرزایی حصاریان (Aghagolzade & Mirzaie Hesarian, 2013) در پژوهش خود با ارائه انگاره‌ای در چارچوب دستور نظام‌مند نقش‌گرای هلیدی به معرفی مدلی برای سطح‌بندی متون آموزشی زبان فارسی به غیرفارسی‌زبانان پرداختند. پژوهش آن‌ها بر روی چند متن که از کتاب آموزش خواندن و درک مطلب فارسی به غیرفارسی‌زبانان انتخاب شده بود، انجام شد. یافته‌ها نشان دادند افزایش تراکم واژگانی سبب پیچیده‌تر شدن و افزایش پیچیدگی دستوری منجر به ساده‌تر شدن متن می‌شود که این یافته‌ها نیازمند اندیشیدن هستند. بر اساس دستور نظام‌مند نقش‌گرای هلیدی، در متون با سطح پیچیدگی پایین‌تر، بندهای مستقل و بندهای همپایه بیشتر از بندهای وابسته و ناهم‌پایه دیده می‌شود. همچنین بر اساس این نظریه جمله‌های معلوم از جمله‌های مجهول خوانایی بیشتری دارند.

در زمینه ملاک‌های سطح‌بندی متون منتخب آموزشی، میرزایی حصاریان و همکاران (Mirzaei Hesarian et al., 2016)، در پژوهشی با نام «سطح‌بندی متون پزشکی فارسی برای غیرفارسی‌زبانان: انگاره‌ای برپایه دستور نظام‌مند- نقش‌گرای هلیدی»، تلاش کردند تا مدلی برای انتخاب متون مناسب سطح زبانی فارسی آموزان خارجی ارائه کنند. بر اساس یافته‌های آن‌ها

استعاره دستوری، اسم‌سازی از فعل و صفت، صفت‌سازی (فاعلی و مفعولی) از فعل، موجز‌گویی، واژه‌های تخصصی، مرتبگی (همپایه و ناهمپایه) و واژگون‌مرتبگی در سطح گروه، بند و بند مرکب، از عوامل مؤثر در دشواری متون پزشکی فارسی بودند.

در طرح ملی «بررسی مهارت‌های چهارگانه نوشتن، خواندن، شنیدن و سخن گفتن دانش‌آموختگان دوره متوسطه» که توسط دانشگر (Daneshgar, 2017) انجام شد، آزمونی برای سنجش و ارزیابی سطح مهارت‌های زبانی و ادبی دانش‌آموختگان دوره متوسطه (امزا) طراحی شد. ساختار این آزمون بر اساس چهارچوب اروپایی و مرجع مشترک زبان‌ها^۱ طراحی شده بود، اما معیار مشخصی برای انتخاب متون ارائه نشده بود. نتیجه این پژوهش نشان داد فارغ‌التحصیلان دوره متوسطه در تمام مهارت‌های زبانی به‌ویژه خواندن و درک، ضعیف و پائین‌تر از سطح معیار (زبان آموز مسلط بر مبنای طبقه‌بندی چهارچوب مشترک اروپا) هستند. همچنین بهترین یافته‌ها در پیوند با مهارت سخن گفتن به دست آمد. این نتیجه نیز با توجه به عدم تدریس رسمی مهارت‌های گفتاری در مدارس تا اندازه‌ای قابل بحث است.

بر پایه بررسی پیشینه مطالعات انجام شده در زبان فارسی می‌توان دریافت تاکنون الگوی مشخص و ثابتی در زبان فارسی برای شاخص‌های خوانش‌پذیری متن ارائه نشده است. شاید بتوان رساله دکتری میرزایی حصاریان (Mirzaie Hesarian, 2016) را نخستین تلاش، برای معرفی چنین الگویی دانست. با وجود تلاش بسیار در معرفی این الگو، همچنان حلقه گمشده این الگو در گزینش متن‌ها پابرجا است.

بررسی‌های گوناگون آزمون‌های ویژه سنجش درک خوانداری بزرگسالان، یافته‌هایی را نشان می‌دهد. در این زمینه می‌توان به بخش درک مطلب آزمون استعداد تحصیلی ورود به دوره دکترا اشاره کرد که از متون تخصصی هر رشته برگرفته شده است و میزان درک متن را می‌سنجد. این در حالی است که در داخل کشور هنوز آزمون استاندارد ویژه سنجش سطح درک خوانداری در گروه بزرگسال طراحی نشده است. آنچه در آزمون پیشرفت تحصیلی وودکاک-جانسون^۲، با نام خرده آزمون درک مطلب برای بزرگسال مشاهده شد، متشکل بر ۱۶ جمله به صورت جای خالی بود. آزمودنی باید با توجه به بافت، واژه مناسب را برای پر کردن جای خالی، جاگذاری کند (سنجش توانایی استنتاج‌سازی). این آزمون هنوز جنبه‌های هنجاریابی و بررسی روایی و پایایی را در داخل کشور پشت سر نگذاشته است.

¹ Common European Framework of Reference for Languages

² Woodcock-Johnson III

مرور کلی پیشینه پژوهش، نبود ابزاری مناسب برای سنجش سطح درک خوانداری بزرگسالان و نیز نداشتن معیارهای مناسب برای برگزیدن متن‌های موردنیاز برای چنین ابزاری را آشکارتر کرد. بنابراین، پژوهش حاضر تلاش کرده‌است این خلاء را در پیشینه مربوط پر کند. همچنین، این مقاله بر آن است به صورت خاص، با تعیین معیار(های) انتخاب متن و پرسش‌های مورد نظر برای آزمون تعیین سطح درک خوانداری بزرگسالان، چنین آزمونی را ساخته و سپس روایی و پایایی آن را تعیین کند. در ادامه مقاله، روش پژوهش برای دست‌یابی به این هدف‌ها مطرح می‌شود.

۴. روش پژوهش

مطالعه حاضر، پژوهشی ترکیبی بوده‌است. بخش کیفی پژوهش، مشتمل بر بررسی معیارهای انتخاب متن و پرسش‌های مورد نظر برای ساخت آزمون تعیین سطح درک خوانداری بزرگسالان، و نیز مرحله‌های ساخت آن بود. بخش کمی پژوهش، شامل اجرای آزمایشی آزمون ساخته‌شده برای تعیین روایی و پایایی آن بود. برای پاسخ‌گویی به پرسش نخست پژوهش در زمینه معیارهای انتخاب متن‌ها و پرسش‌های آزمون با استفاده از روش‌های توصیفی و مروری، ملاک‌های طراحی متن استخراج شد. برای پاسخ به پرسش دوم، آزمونی بر اساس ملاک‌های به‌دست‌آمده از مرحله پیش طراحی شد. برای پاسخ‌گویی به پرسش سوم پژوهش، ابتدا روایی محتوایی آزمون طراحی‌شده توسط ۵ کارشناس موضوعی بررسی شد. برای اطمینان از پایایی آن روی نمونه آزمایشی اجرا و مقدار پایایی با استفاده از فرمول ضریب آلفای کرونباخ محاسبه شد. در جریان محاسبه پایایی، گویه‌ها از نظر سهولت، شاخص تمیز و چگونگی توزیع گزینه‌های انحرافی تحلیل شدند. نمونه آماری این پژوهش ۶۰ نفر از دانشجویان کارشناسی ارشد دانشگاه گیلان بودند. این تعداد مشتمل بر ۱۹ دانشجوی مرد و ۴۱ دانشجوی زن بودند که به شیوه نمونه‌گیری در دسترس^۱ انتخاب شدند. این آزمودنی‌ها در سال تحصیلی ۹۶-۹۷ در مقطع کارشناسی ارشد رشته‌های زبان‌شناسی، آموزش زبان انگلیسی، تربیت بدنی و شهرسازی در دانشگاه گیلان مشغول به تحصیل بودند.

۵. یافته‌های پژوهش

در این بخش، تلاش می‌شود که به پرسش‌های پژوهش بر مبنای یافته‌های به‌دست‌آمده پاسخ داده شود. به این منظور، یافته‌ها در سه بخش تعیین ملاک و معیار انتخاب متن‌ها و طراحی پرسش‌ها، طراحی و ساخت آزمون و تعیین روایی و پایایی آزمون ارائه می‌شود.

^۱ convenience sampling

۵. ۱. ملاک و معیار انتخاب متن‌ها و طراحی پرسش‌ها

در انتخاب متن‌های آزمون حاضر، به ویژگی تناسب متن‌ها با سطح دانش و آگاهی آزمودنی‌های بزرگسال (دیپلم به بالا) توجه شد. بررسی آزمون‌های خواندن در سطح بین‌المللی مشخص کرد که موضوع متن‌ها باید از بین موضوعات ادبی، اجتماعی، تاریخی و علمی انتخاب شود. در تعیین سطح دشواری متن‌ها گرچه از فرمول‌های رایج تعیین سطح خوانش‌پذیری استفاده نشد، اما تلاش شد تا با توجه به معیارهای مناسب در زمینه پیچیدگی متن که مهم‌ترین آن‌ها انسجام و پیوستگی است، متن‌های مناسب انتخاب شوند.

بر این اساس، برای انتخاب متن‌ها، تلفیقی از معیارهای معرفی شده توسط انجمن بین‌المللی ارزشیابی پیشرفت سواد خواندن برای آزمون دوره پیشرفته پرلز (PIRLS, 2016) و معیارهای مؤسسه آزمون‌سازی تحصیلی^۱ (Educational Testing Service, 2010) مورد توجه قرار گرفت. این معیارها عبارتند از:

- تعداد واژه‌های ۳۰۰ تا ۶۰۰ واژه
- متناسب بودن متن با معیارهای فرهنگی
- داشتن وضوح کافی (نداشتن ابهام واژگانی، معنایی و نحوی)
- استفاده از ابزارهای انسجام واژگانی، انسجام دستوری و انسجام پیوندی مانند کاربرد درست حروف ربط و ضمائر ارجاعی مناسب، استفاده از جایگزینی‌ها و حذف‌های مناسب
- عدم استفاده از ساختارهای پیچیده نحوی مثل عدم استفاده مکرر از جملات لانه‌گیری شده^۲
- عدم نیاز به دانش تخصصی زیاد برای درک متن‌ها

برای طراحی پرسش‌های آزمون، همه راهبردهای زیربنایی درک متن که شرح آن در پیشینه پژوهش گذشت (استنتاج‌سازی، نظارت بر درک، ساختار متن و موارد مشابه) در نظر گرفته شد. در این زمینه، ۶ سطح برای طراحی سؤالات استفاده شد که این سطح‌ها بر مبنای تلفیقی از مدل دی و پارک (Day & Park, 2005) و طراحی سؤالات تافل برای آزمون‌های تعیین سطح خوانداری تهیه شد. همان‌گونه که دی و پارک (Day & Park, 2005) بیان کردند و پیش از این در بخش پیشینه به آن اشاره شد، پرسش‌های چندگزینه‌ای بهترین نوع سؤالات برای ارزیابی درک خوانداری به شمار می‌آیند. بر این مبنای انواع سطح‌های راهبردهای درک خوانداری را می‌توان در

^۱ مؤسسه آزمون‌سازی تحصیلی (ETS)، مجری برگزاری آزمون‌های تافل در سراسر جهان است.

^۲ embeddedness structures

قالب این نوع پرسش‌ها مطرح کرد. بنابراین پرسش‌های این آزمون نیز به صورت چهارگزینه‌ای طراحی شده‌اند. خلاصه طراحی سؤالات آزمون در جدول (۲) قابل مشاهده است.

جدول ۲: پراکندگی پرسش‌های آزمون تعیین سطح درک خوانداری

انواع سطوح راهبردهای درک خوانداری						نوع پرسش‌ها
ارزیابی (تشخیص لحظی متن و احساس نویسنده)	پیش‌بینی (تشخیص موضوع پاراگراف‌های پیش و پس از متن)	استنتاج‌سازی	سازمان‌دهی مجدد	نحت الفظی	پرسش‌ها ایده اصلی	
۲ سؤال	۲ سؤال	۸ سؤال	۱۰ سؤال	۵ سؤال	۳ سؤال	چهارگزینه‌ای

۵.۲. طراحی و ساخت آزمون سنجش درک خوانداری بر اساس معیارهای

پیش‌گفته

به این منظور، ابتدا هشت متن توضیحی با موضوع‌های مختلف و مناسب با معیارهایی که در بخش پیشین بیان شد از کتاب‌های خواندن و درک مطلب برای فارسی‌آموزان سطح پیشرفته انتخاب شد. این متن‌ها به وسیله محمدباقر میرزایی حصاریان (Mirzaei Hesarian, 2011)، نوشته شده و ترجمه‌ای از متون آزمون‌های تافل آن‌تی‌سی^۱ و تولیمو^۲ است. این متن‌ها بر مبنای معیارهای انسجام از جمله عناصر انسجام و ازگانی مشتمل بر رعایت استفاده از هم‌معناها و عدم استفاده پیوسته از واژگان مشابه، انسجام پیوندی شامل رعایت استفاده مناسب از پیوندهای افزایی، تباینی، علی و زمانی و انسجام دستوری شامل استفاده از عناصر ارجاعی مناسب و رابط‌های پیوندهای پاره‌گفته‌ها مناسب‌سازی شد. سپس این متون، به همراه پرسشنامه‌ای محقق‌ساخته شامل شش پرسش نگرش‌سنجی به صورت بسته-پاسخ در زمینه میزان علاقه‌مندی به متون، تناسب متن‌ها با سطح تحصیلی، طول متن‌ها، ساختار دستوری جمله‌ها و پاره‌گفته‌ها، استفاده از واژه‌های غیرتخصصی و استفاده درست از علامت‌های نگارشی در اختیار چند متخصص آموزش زبان فارسی و زبان‌شناسی و ۱۰ دانشجوی کارشناسی ارشد قرار داده شد. بر پایه دیدگاه‌های دریافت شده، تعداد چهار متن با میانگین ۳۰۰ واژه در هر متن، از بین این متون انتخاب شدند و پس از

^۱ NTC

^۲ TOLIMO

مناسب‌سازی محتوای متن‌ها طبق نظرات دریافت شده، پرسش‌های درک‌خوانداری متناسب با آن‌ها طراحی شد.

تعداد متن‌ها و پرسش‌های مطرح شده برای هر متن، همانند بخش درک‌متن آزمون استاندارد تافل آی‌بی‌تی است. برای طراحی پرسش‌ها از مدل تلفیقی آزمون‌های استاندارد و الگوی دی و پارک (Day & Park, 2005) استفاده شد. این مدل، به سنجش تشخیص ایده‌اصلی متن و پاراگراف‌ها، واژگان، انواع توانایی‌های استنتاجی از جمله ارجاع ضمائر و استنتاج‌های تفصیلی، تحلیل درست متن، تشخیص جمله‌های پیشین و پسین می‌پردازد. آزمون محقق‌ساخته از ۳۰ سؤال چهارگزینه‌ای تشکیل شده است که آزمودنی در مدت ۳۵ دقیقه به سؤالات آن پاسخ می‌دهد.

ملاک نمره‌گذاری این آزمون، همانند بخش درک‌متن آزمون تافل آی‌بی‌تی است که ۳۰ نمره‌ای بوده و به هر پاسخ درست یک نمره اختصاص می‌یابد. این آزمون برای تعیین سطح درک‌خوانداری سه سطح قوی، متوسط و ضعیف در نظر می‌گیرد. در چارچوب این ملاک نمره‌گذاری، افراد با نمره ۲۱ به بالا در گروه ادراک‌کننده قوی، افراد با نمرات ۱۵ تا ۲۰ در گروه ادراک‌کنندگان سطح متوسط و افراد با نمره ۱۴ و پایین‌تر در گروه ادراک‌کنندگان ضعیف قرار می‌گیرند.

۳.۵. اجرای مقدماتی آزمون

به منظور بررسی مقدماتی آزمون از جنبه پایایی و تحلیل گویه‌ها، این آزمون در مرحله نخست، روی گروه نمونه ۳۰ نفره از دانشجویان کارشناسی ارشد دانشگاه گیلان انجام شد. این آزمون، در یک جلسه ۳۵ دقیقه‌ای در یکی از کلاس‌های دانشگاه به صورت چاپی و با مداد و بر روی کاغذ اجرا شد.

۴.۵. بررسی مقدماتی روایی محتوایی و پایایی آزمون

هنگامی که یک ابزار، بتواند آن‌چه را که برای اندازه‌گیری‌اش در نظر گرفته شود را به‌طور شایسته‌ای اندازه بگیرد، روایی دارد. در صورتی که ابزار پژوهش از روایی مناسب برخوردار نباشد، همه جنبه‌ها و یافته‌های پژوهش می‌توانند تحت تأثیر قرار گیرند (Litosseliti, 2010, p. 55). یکی از انواع روایی، روایی محتوایی^۱ است. منظور از روایی محتوایی، تحلیل منطقی محتوای یک آزمون است و تعیین آن بر اساس قضاوت کارشناسان انجام می‌گیرد. به منظور سنجش روایی

¹ content validity

محتوایی، آزمون طراحی شده در اختیار دو استاد زبان شناس، یک استاد آموزش زبان انگلیسی و دو مدرس آموزش زبان فارسی قرار داده شد. سپس روایی محتوایی آن به تأیید اساتید رسید. همچنین، بر اساس دیدگاه‌های دریافت شده، اصلاحاتی در جمله‌بندی متون و ۱۲ مورد از پرسش‌ها انجام شد.

پایایی یک ابزار اندازه‌گیری، به دقت آن اشاره می‌کند. یک آزمون در صورتی پایا است که اگر آن را به فاصله‌های زمانی کوتاه و چندین بار بر روی یک گروه انجام دهیم، یافته‌های به‌دست آمده نزدیک به هم باشد. در واقع پایایی، تکرارپذیری یافته‌های یک آزمون در فاصله‌های زمانی مشابه است. شیوه رایج برای بررسی پایایی آزمون‌ها روش بازآزمایی است. این روش فقط هنگامی پذیرفتنی است که ویژگی مورد اندازه‌گیری در اندازه‌ی زمان تغییر نکند (مانند هوش). این در حالی است که نگهداشتن شرایط یکسان برای مواردی که کنش انسان در آن‌ها دخالت دارد، بسیار مشکل است. زیرا انسان‌ها با استفاده از تجربه موارد بسیاری را فرا می‌گیرند و در صورت تکرار آزمون مشابه با گروهی مشابه و به دلیل کسب تجربه تا اندازه‌ی بسیاری یافته‌ها تحت تأثیر قرار خواهد گرفت (Litosseliti, 2010, p. 55). در این میان، رایج‌ترین روش بررسی پایایی، استفاده از فرمول ضریب آلفای کرونباخ^۱ است. ضریب آلفای کرونباخ، یک روش اعتباریابی پایایی آزمون است که همسانی درونی آزمون موردنظر را ارزشیابی می‌کند. ضریب آلفای کرونباخ، نشان می‌دهد که پرسش‌های آزمون تا چه اندازه توانایی سنجش ویژگی یکسانی را دارند.

برای بررسی پایایی آزمون طراحی شده در این پژوهش، از روش ضریب آلفای کرونباخ استفاده شد. میزان ضریب آلفای آزمون در گروه نمونه ۳۰ نفره از دانشجویان کارشناسی ارشد که در یک جلسه ۳۵ دقیقه‌ای به سؤالات پاسخ دادند، ۰/۸۲ محاسبه شد. این میزان نشان می‌دهد آزمون از پایایی مناسبی برخوردار است (به جدول (۳) نگاه کنید).

جدول ۳: آمارهای پایایی آزمون تعیین سطح درک خوانداری

تعداد آیتم‌ها	آلفای کرونباخ بر اساس آیتم‌های استاندارد شده	آلفای کرونباخ
۳۰	۰/۸۲۱	۰/۸۲۰
انحراف معیار	واریانس	میانگین
۵/۹۵۲۳۰	۳۵/۴۳۰	۱۷/۸۶۶۷

¹ Cornbach's alpha

الف) سهولت گویه‌ها^۱: سهولت گویه‌ها، نشان دهنده میزان آسان بودن پرسش‌های آزمون است. گویه‌ها نباید خیلی ساده و یا خیلی سخت باشند. برای اندازه‌گیری سهولت آزمون برای هر گویه فرمول زیر به کار برده می‌شود:

$$I.F. = \frac{\sum c}{N} \quad N = \text{تعداد شرکت کننده} \quad c = \text{تعداد پاسخ‌های درست}$$

سهولت گویه‌ها از صفر تا یک متغیر است. گویه‌ای با سهولت صفر بسیار سخت و گویه‌ای با سهولت یک بسیار ساده است. گویه‌ ایده‌آل، گویه‌ای با سهولت ۰/۵ است. محدوده سهولت گویه‌ها بین ۰/۳۷ تا ۰/۶۳ است. به این معنا که گویه‌ای با سهولت کمتر از ۰/۳۷ بسیار سخت و گویه‌ای با سهولت بالای ۰/۶۳ بسیار ساده است (Farhadi, Jafarpour & Birjandi, 2004, p.100).

بر پایه اطلاعات به دست آمده از انجام آزمون ۳۰ گویه‌ای روی نمونه مطالعه، ۵ گویه دارای سهولت بالاتر از ۰/۶۳ بودند. این پرسش‌ها حذف و های جدیدی جایگزین شدند (به جدول (۴) نگاه کنید).

ب) شاخص تمیز^۲: به طور کلی، یکی از اهداف آزمون‌های تشخیصی، فراهم کردن اطلاعات لازم در مورد عملکرد بهتر آزمودنی‌های قوی نسبت به آزمودنی‌های با عملکرد ضعیف است. یعنی آزمون بتواند بین آزمودنی‌های قوی و ضعیف تمایز قائل شود. قدرت تمیز یک آزمون وابسته به قدرت تمیز گویه‌های آن است. شاخص تمیز، یک نشانه و عامل تمایزدهنده بین آزمودنی‌های قوی و ضعیف به شمار می‌رود. این شاخص، مشخص می‌کند که هر گویه تا چه اندازه می‌تواند بین آزمودنی‌ها تمایز برقرار کند. برای نمونه، اگر همه آزمودنی‌ها به پرسشی درست پاسخ دهند، نشان می‌دهد که پرسش نه تنها بسیار ساده است، بلکه قابلیت تمایز بین افراد را ندارد. بنابراین بین سهولت گویه و شاخص تمیز رابطه وجود دارد. گویه‌ای که دارای سهولت بسیار زیاد یا بسیار کم است، نمی‌تواند قدرت تمیز مناسبی داشته باشد. بنابراین شاخص تمیز، شاخصی است که از مقایسه عملکرد آزمودنی‌های قوی و ضعیف در هر یک از گویه‌های آزمون به دست می‌آید. شاخص تمیز بین -۱ تا +۱ متغیر است که هر چه این عدد به صفر نزدیک‌تر باشد، توانایی تمیز آن کمتر است. به طور کلی گویه‌های با شاخص ۰/۴ و بالاتر از قدرت تمیز بهتری برخوردارند. هر چه شاخص تمیز یک گویه بزرگتر باشد، توانایی آن در تمایز گذاشتن بین

¹ item facility

² discrimination index

آزمودنی‌های قوی و ضعیف بیشتر است (Farhadi, Jafarpour & Birjandi, 2004, p. 102).

فرمول محاسبه شاخص تمیز به صورت زیر است:

$$ID = \frac{CH-CL}{\frac{1}{2}N}$$

CH= تعداد پاسخ‌های درست گروه قوی

CL= تعداد کل آزمودنی‌ها N= تعداد پاسخ‌های درست گروه ضعیف

اطلاعات به دست آمده از نمونه اولیه آزمودنی‌ها بر اساس نمرات کسب شده به دو گروه قوی و ضعیف تقسیم‌بندی شدند. بر پایه این اطلاعات، شاخص تمیز در ۱۴ گویه اندکی کمتر از سطح معیار بود که بازسازی محتوای گویه‌ها و گزینه‌های آن‌ها انجام پذیرفت (به جدول (۴) نگاه کنید).

جدول ۴: تحلیل پیش آزمون درک خوانداری از لحاظ سهولت آزمون و شاخص تمیز

شماره سوال	سهولت گویه	شاخص تمیز
۱	۰/۶	۰/۲۶
۲	۰/۶۶	۰/۶۶
۳	۰/۶	۰/۱۳
۴	۰/۶۶	۰/۴
۵	۰/۶۳	۰/۳۳
۶	۰/۵۶	۰/۶
۷	۰/۵۳	۰/۴
۸	۰/۶۳	۰/۴۶
۹	۰/۵۶	۰/۳۳
۱۰	۰/۶۶	۰/۴
۱۱	۰/۶	۰/۴
۱۲	۰/۶	۰/۴
۱۳	۰/۶۳	۰/۶
۱۴	۰/۵۶	۰/۲
۱۵	۰/۵	۰/۲
۱۶	۰/۶۳	۰/۳۳
۱۷	۰/۶۳	۰/۴۶
۱۸	۰/۵	۰/۲
۱۹	۰/۵۳	۰/۶۶
۲۰	۰/۵۶	۰/۲
۲۱	۰/۶۶	۰/۲۶
۲۲	۰/۴۶	۰/۱۳
۲۳	۰/۷۳	۰/۲۶
۲۴	۰/۵۶	۰/۳۳
۲۵	۰/۶	۰/۲۶
۲۶	۰/۵۳	۰/۴
۲۷	۰/۶۶	۰/۴
۲۸	۰/۵	۰/۳۳
۲۹	۰/۶	۰/۱۳
۳۰	۰/۶۳	۰/۳۳

ج) توزیع گزینه‌ها: سهولت گویه و شاخص تمیز دو پارامتر برای تناسب و قابلیت پذیرش پرسش‌های آزمون هستند. بر اساس این مشخصه‌ها می‌توان درباره‌ی درستی آزمون پیش‌داوری کرد. هر چند نکته‌ای که به وسیله‌ی این دو شاخص مشخص نمی‌شود، کفایت گزینه‌های انحرافی^۱ است. معیار دیگری که در تحلیل گویه‌های آزمون مورد توجه قرار می‌گیرد، توزیع گزینه‌های انحرافی است. به این معنا که فراوانی انتخاب هر گزینه توسط آزمودنی‌ها، که از طریق این اطلاعات آزمون‌گر به کفایت گزینه‌های انحرافی پی می‌برد. گاهی گویه‌ای علاوه بر داشتن شاخص تمیز و سهولت مناسب دارای گزینه‌های انحرافی نامناسب است. در این صورت، گزینه‌هایی که توجه هیچ‌یک از آزمودنی‌ها را به خود جلب نکرده‌اند، باید حذف یا تعدیل شوند. هدف از قراردادن گزینه‌ی انحرافی در پرسش‌ها منحرف کردن آزمودنی‌هایی است که پاسخ درست پرسش را نمی‌دانند. بنابراین، در صورتی یک پرسش به خوبی عمل می‌کند که افراد ضعیف بیشتر از افراد قوی، گزینه‌های انحرافی آن پرسش را انتخاب نمایند (Farhadi, Jafarpour & Birjandi, 2004, p. 105). با توجه به اطلاعات به‌دست‌آمده از نمونه‌ی اولیه‌ی آزمودنی‌ها، گزینه‌های انحرافی در ۱۳ گویه، کمتر توجه آزمودنی‌ها را به خود جلب کرده بود که با تغییر جمله‌بندی گزینه‌های انحرافی بازسازی و تعدیل شدند.

۵.۵. بررسی دوباره پایایی و تحلیل گویه‌ها بعد از انجام تغییرات به‌دست‌آمده از تحلیل گویه‌ها

پس از انجام مراحل تحلیل گویه‌های آزمون، آزمون تصحیح‌شده پس از فاصله‌ی زمانی دو ماه از اجرای آزمون اول در گروهی ۳۰ نفره جدید از دانشجویان کارشناسی ارشد در یک جلسه‌ی ۳۵ دقیقه‌ای اجرا شد. ضریب آلفای کرونباخ دوباره محاسبه شد. این بار نیز ضریب آلفا ۰/۸۴ به دست آمد که نشان‌دهنده‌ی پایایی مناسب آزمون برای سنجش سطح درک خوانداری افراد بود (به جدول (۵) نگاه کنید). همچنین مراحل تحلیل گویه‌ها دوباره روی آزمون انجام شد که در هر سه معیار سهولت گویه‌ها، شاخص تمیز و توزیع گزینه‌هایی انحرافی مناسب تشخیص داده شد.

¹ choice distribution

² distracters

جدول شماره ۵: آمارهای پایایی آزمون تعیین سطح درک خوانداری در مرحله دوم

تعداد آیتم‌ها	آلفای کرونباخ بر اساس آیتم‌های استاندارد شده	آلفای کرونباخ
۳۰	۰/۸۲۸	۰/۸۲۶
انحراف معیار	واریانس	میانگین
۵/۹۱۷۴۴	۳۵/۰۱۶	۱۷/۴۶۶۷

۶. نتیجه‌گیری

پژوهش حاضر با سه هدف کلی انجام شد. این اهداف مشتمل بر تعیین معیارهای انتخاب متن‌ها و پرسش‌ها در آزمون‌های درک خوانداری، طراحی آزمون تعیین سطح درک خوانداری و تعیین روایی و پایایی آزمون محقق‌ساخته بودند. بر پایه پژوهش‌های انجام شده، مهم‌ترین ملاک در تعیین متن برای چنین ابزاری رعایت معیارهای انسجامی متن مشتمل بر انسجام واژگانی، دستوری و پیوندی و عدم استفاده از ساختارهای نحوی پیچیده در جمله‌ها است. در طراحی آزمون تعیین سطح درک خوانداری برای بزرگسالان تلاش شد تا آزمونی با پوشش تمامی مؤلفه‌های درک خوانداری اعم از سنجش انواع توانایی‌های استنتاجی، تحلیل درست متن، تشخیص پیام اصلی نویسنده و درک ساختار منطقی متن بر پایه الگویی تلفیقی از آزمون‌های بین‌المللی و مدل سطح‌بندی سنجش راهبردهای درک خوانداری دی و پارک (Day & Park, 2005) طراحی شود. نتیجه بررسی‌های اولیه روی آزمون طراحی شده نشان داد روایی و پایایی آن در اندازه مطلوبی قرار دارد. همچنین، پس از انجام مراحل تحلیل گویه‌ها اصلاح‌های لازم روی آزمون ساخته شده صورت گرفته است. تاکنون ابزار مناسبی برای تعیین سطح درک متن افراد بزرگسال ایرانی وجود نداشته است. بر این مبنای، این آزمون با توجه به سهولت در اجرا و شیوه نمره‌گذاری، می‌تواند ابزاری مطلوب برای تعیین سطح درک خوانداری بزرگسالان باسواد ایرانی باشد. این آزمون در گام پسین باید روی نمونه گسترده‌تری از جامعه هدف اجرا شود تا مراحل روایی سازه و هنجاریابی آن انجام شود و بتوان با اطمینان بیشتری به یافته‌های آن استناد کرد. همچنین، تعیین سطح درک خوانداری افراد بزرگسال می‌تواند نقش مهمی در تدوین برنامه‌های کمک آموزشی برای کاهش مشکلات ناشی از این ضعف مهارتی، بازی می‌کند.

فهرست منابع

- آقاگل‌زاده، فردوس و میرزایی حصاریان، محمدباقر (۱۳۹۲). «در جستجوی انگاره‌ای آموزشی بر پایه پیچیدگی‌های واژگانی و دستوری در دستور نظام مند نقش‌گرای هلیدی: بررسی موردی کتاب خواندن و درک فارسی آموزان غیرایرانی». *پژوهش‌نامه آموزش زبان فارسی به غیر فارسی‌زبانان*. سال ۲. شماره ۲. صص ۸۷-۱۱۸.
- البرزی، پرویز (۱۳۹۲). *مبانی زبان‌شناسی متن*. تهران: امیرکبیر.
- امیری خراسانی، احمد و علی‌نژاد، حلیمه (۱۳۹۴). «بررسی عناصر انسجام متن در نفثة‌المصدر بر اساس نظریه هلیدی و حسن». *متن پژوهی/ادبی*. سال ۱۹. شماره ۶۳. صص ۷-۳۱.
- جباری، سوسن و خادمی، محسن (۱۳۹۲). «ساخت آزمون تشخیصی خواندن و درک مطلب جهت دانش‌آموزان دوره ابتدایی». *پژوهش‌های برنامه‌درسی*. سال ۳. شماره ۲. صص ۳۱-۵۱.
- دانشگر، مریم (۱۳۹۶). «مهارت‌های دانش‌آموختگان نظام متوسطه با محوریت درس زبان و ادبیات فارسی». *جستارهای زبانی*. سال ۸. شماره ۱. صص ۲۳۱-۲۵۶.
- رقیب‌دوست، شهلا و اصغری‌پور ماسوله، مهرداد (۱۳۸۸). «رابطه انگیزش خواندن و دانش پیشین با درک خواندن زبان آموزان غیرفارسی‌زبان». *زبان‌پژوهی*. سال ۱. دوره ۱. صص ۵۱-۷۲.
- شکیبا، ابوالقاسم و اسدزاده، حسن (۱۳۹۲). «بررسی انسجام متن در سطح خرد و کلان بر بهبود درک مطلب». *پژوهش و نگارش کتب دانشگاهی*. سال ۱۷. شماره ۳۰. صص ۵۰-۶۶.
- عزیزیان، مرضیه و عابدی، محمدرضا (۱۳۸۴). «ساخت و هنجاریابی آزمون تشخیصی سطح خواندن برای دانش‌آموزان پایه سوم دبستان». *روانپزشکی و روانشناسی بالینی ایران*. سال ۱۱. شماره ۴. صص ۳۷۹-۳۸۷.
- قتسولی، بهزاد (۱۳۸۹). «طراحی و روایی سنجی آزمون بسندگی زبان فارسی». *پژوهش ادبیات معاصر جهان*. سال ۱۵. شماره ۵۷. صص ۱۱۵-۱۲۹.
- میرزایی حصاریان، محمدباقر (۱۳۹۰). *خواندن و درک مطلب متون فارسی (برای فارسی‌آموزان سطح پیشرفته)*. قزوین: انتشارات دانشگاه بین‌المللی امام خمینی (ره).
- میرزایی حصاریان، محمدباقر (۱۳۹۵). *طراحی انگاره سطح‌بندی متون پزشکی فارسی برای خارجی‌ها: چارچوب دستور نظام‌مند-نقش‌گرای هلیدی*. رساله دکتری. دانشگاه تربیت مدرس.
- میرزایی حصاریان، محمدباقر، فردوس آقاگل‌زاده، عالیبه کرد زعفرانلو و ارسلان گلفام (۱۳۹۵). «سطح‌بندی متون پزشکی فارسی برای غیر فارسی‌زبانان: انگاره‌ای برپایه دستور نظام‌مند-نقش‌گرای هلیدی». *پژوهش‌نامه آموزش زبان فارسی به غیر فارسی‌زبانان*. سال ۵، شماره ۱۲. صص ۴۷-۷۴.

References

- Adams, R., & Cresswell, J. (2016). PISA for development technical strand: an enhancement of PISA cognitive instruments. *OECD Education Working Papers* (No. 126.). Paris: OECD Publishing.
- Adlof, S. M., Catts, H. W., & Lee, J. (2010). Kindergarten predictors of second versus eighth grade reading comprehension impairments. *Journal of Learning Disabilities*, 43(4), 332-345.
- Aghagolzadeh, F. & Mirzaei Hesarian, M. (2013). In search of a Model based on lexical and grammatical complexities in Haliday's systemic functional grammar: the case study of Persian language reading for non-Persian learners. *Journal of Teaching Persian to Speakers of Other Languages*, 2(2), 87-118 [In Persian].
- Alborzi, P. (2013). *ABC of Text linguistics*. Tehran: Amirkabir publication [In Persian].
- Allen, K. D., & Hancock, T. E. (2008). Reading comprehension improvement with individualized cognitive profiles and metacognition. *Literacy Research and Instruction*, 47(2), 124-139.
- Azizian M., Abedi, M. (2006). Construction and standardization of reading level diagnostic test for third Grade primary school children, *IJPCP*. 11(4), 379-387 [In Persian].
- Amiri khorasani, A. & Alinejad, H. (2015). The Analysis of Text cohesion elements in Nafthath al-Masdoor based on Halliday and Hasan's theory. *Literary Text Research*, 19 (63), 7-31 [In Persian].
- Atkinson, L. (2014). *A longitudinal investigation of the social, cognitive and social cognitive predictors of reading comprehension*. (Unpublished PhD. dissertation, University of Roehampton, Roehampton, London.
- Retrieved from <http://hdl.handle.net/10142/608675>.
- Baker, L. & Anderson, R. I. (1982). Effects of inconsistent information on text processing: evidence for comprehension monitoring. *Reading Research Quarterly*, 17, 281-293.
- Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge university press.
- Cain, K., & Oakhill, J. V. & Bryant, P. E. (2000). Phonological skills and comprehension failure: A test of phonological processing deficit hypothesis. *Reading and Writing*, 13 (1-2), 31-56.
- Cain, K., Oakhill, J.V., & Lemmon, K. (2004). Individual differences in the inference of word meanings from context: the influence of reading comprehension, vocabulary knowledge, and memory capacity. *Journal of Educational Psychology*, 96, 671-681.
- Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: a case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49 (2), 278--293.
- Corte, E., Verschaffel, L., & Ven, A. (2001). Improving text comprehension strategies in upper primary school children: a design experiment. *British Journal of Educational Psychology*, 71(4), 531-559.
- Crossley, S. A., Allen, D. B., & McNamara, D. S. (2011). Text readability and intuitive simplification: a comparison of readability formulas. *Reading in a Foreign Language*, 23(1), 86.
- Dale, E. & Chall, J. (1948). A formula for predicting readability. *Educational Research Bulletin*, 22, 11-20.
- Daneshgar, M. (2017). Literary and lingual skills evaluation (case study: Tehran secondary school students). *Language Related Research*, 8 (1), 231-256 [In Persian].
- Davis, A. (2001). The logic of testing languages for specific purposes. *Language Testing*, 18(2), 133-147.
- Day, R., & Park. J.S. (2005). Developing reading comprehension questions. *Reading in a foreign language*, 17(1), 60-67.
- De Beaugrande, R., & Dressler, W. U. (1981). *Introduction to text linguistics*. United Kingdom: Routledge.

- Demirgünes, S. (2017). Microstructural (cohesion and coherence) text generation problems of Syrian refugee students learning Turkish. *Universal Journal of Educational Research*, 5(4), 581-590.
- Ehrlich, M. F. (1991). Metacognition and reading comprehension: theoretical and methodological problems. *Learning and instruction: european research in an international context*, 3, 351-363.
- Farhadi, H., Jafarpour, A. & Birjandi, P. (2004). *Testing language skills form theory to practice*. Tehran: SAMT [In Persian].
- Flesch, R. (1948). A new readability yardstick. *Journal of applied psychology*, 32(3), 221-233.
- Fry, E. (1968). A readability formula that saves time. *Journal of Reading*, 11 (7), 265-271.
- Ghonsooly, B. (2010). Development and validation of a Persian proficiency test. *Research in Contemporary World literature*, 15 (57), 115-129 [In Persian].
- Gough, P. B., & Tunmer, W. E. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7(1), 6-10.
- Grabe, W., & Stoller, F. L. (2011). *Teaching and researching reading (2nd ed.)*. New York: Routledge.
- Gunning, R. (1952). *The technique of clear writing*. New York: McGraw-Hill Book Company.
- Halliday, M. A., & Hasan, R. (1976). *Cohesion in English*. England: Longman.
- Hatim, B., & Mason, I. (1997). *The translator as communicator*. New York and London: Routledge.
- Jabbari, S. & Khademi. (2013). Developing a diagnostic test for the elementary students. *Journal of Curriculum research*, 3(2), 33-51. [In Persian].
- Kintsch, E. (2005). Comprehension theory as a guide for the design of thoughtful questions. *Topics in Language Disorders*, 25(1), 51-64.
- Kintsch, W. & Van Dijke, T.A. (1978). Toward a model of text comprehension and production. *Psychological Review*, 85, 363-394.
- Litosseliti, L. (2010). *Research methods in linguistics*. UK: Bloomsbury Publishing.
- Mirzaei Hesarian, M. (2011). *Persian reading comprehension for advanced learners*. Qazvin: Imam Khomeini International University Publication [In Persian].
- Mirzaei Hesarian, M. (2016). *Planning a Model of Leveling Persian Medical Texts for foreigners: Systemic-functional Grammar framework of Halliday*. Unpublished Doctoral dissertation, University of Tarbiat Modarres, Tehran [In Persian].
- Mirzaei Hesarian, M.; Aghagolzadeh, F.; kord zafaranlu, A. & Golfam, A. (2016). Persian medical texts leveling for Non-Persian language speakers: a model based on Halliday's systemic-functional grammar. *Journal of Teaching Persian to Speakers of Other Languages*, 5 (12), 47-74.
- Mullis, I. V., Martin, M. O., Foy, P., & Drucker, K. T. (2012). *PIRLS 2011 International Results in Reading*. Amsterdam: International Association for the Evaluation of Educational Achievement.
- Mullis, I. V., Martin, M. O., & Sainsbury, M. (2016). PIRLS 2016 reading framework. *PIRLS*, 11-29.
- Nation, K. (2005). *Children's reading comprehension difficulties*. In M. J. Snowling, & C. Hulme(Eds.), *The Science of Reading: A Handbook* (pp. 248-266). Malden: Blackwell Publishing.
- Perfetti, C. A., Landi, N., & Oakhill, J. (2005). *The acquisition of reading comprehension skill*. In M. J. Snowling, & C. Hulme (Eds.), *The Science of Reading: A Handbook* (pp. 227-247). Malden: Blackwell Publishing.
- Raghibdoust, Sh. & Asgharpour, M. (2009). The relation between motivation and prior knowledge with reading comprehension to non-Persian language learners. *Zabanpa-zhuhi*, 1 (1), 51-72. [In Persian].
- Rawson, K. A., & Kintsch, W. (2005). Rereading effects depend on time of test. *Journal of educational psychology*, 97(1), 70-80.

- Shakiba, A. & Asadzahe, H. (2013). The impact of local and global coherent in the interpretation of student's comprehension at high school level. *Journal of University Textbooks: Research and Writing*, 17 (30), 50-66. [In Persian].
- Silagi, M. L., Romero, V. U., Mansur, L. L., & Radanovic, M. (2014). Inference comprehension during reading: influence of age and education in normal adults. *Codas*, 26 (5), 407-414.
- Tarchi, C. (2015). Fostering reading comprehension of expository texts through the activation of readers' prior knowledge and inference-making skills. *International Journal of Educational Research*, 72, 80-88.
- Van Dijk, T. A. (1972). Foundations for typologies of texts. *Semiotica*, 6(4), 297-323.
- Woodcock, R. W., Mather, N., McGrew, K. S., Schrank, F. A., Johnson, M. B., & Wendling, B. J. (2001). *Woodcock-Johnson III normative update: tests of cognitive abilities*. United States: Riverside Publishing.



Development and Preliminary Implementation of Adults Reading Comprehension Placement Test: Evidence from MA Students of University of Guilan

Sārānā Ghavāmi Lāhij¹
Maryam Dānāye Tous*²
Abdorrezā Tahriri³
Ali Rabi⁴

Received: 13/03/2018

Accepted: 16/07/2018

Abstract

Reading comprehension is the process of reader's interaction with a written language and making meaning from the text. Studies show that there is a great difference between skilled and poor readers in terms of using the fundamental strategies of reading comprehension, i.e., cognitive and metacognitive strategies. Due to the lack of appropriate tool for measuring adults' reading comprehension level in Iran and the lack of appropriate criteria for selecting the texts for such a tool, the aim of this study was to make a placement tool for evaluating adult Persian speakers' reading comprehension level. It was a mixed methods study and the aim was to answer three main questions about finding a text selection criterion for reading comprehension tests, using the selected criteria for making reading comprehension placement test and studying the validity and reliability of the devised test. Findings showed that text selection criteria should be fitted with patterns of international tests and the linguistic principles of text. The content validity of the text was approved by experts, after implementing their comments on the texts and questions. To ensure the test reliability and to do item analysis, the test was distributed among a sample of 60 MA students of University of Guilan at two stages. The reliability of the test was computed and Cronbach's alpha were 0.84 and 0.82 respectively, showing the appropriate reliability of the tool. After normalization, this tool can be used to evaluate adult reading comprehension and in educational planning, it can be used for selecting the educational content.

¹ PhD Candidate of Linguistics, Faculty of Literature and Foreign Languages, Islamic Azad University, South Tehran Branch, Iran; st_s_ghavamilahij@azad.ac.ir

² Associate Professor of Linguistics, Department of English Language and Literature, Faculty of Literature and Humanities, University of Guilan, Rasht, Iran (Corresponding Author); maryam.dana@guilan.ac.ir

³ Assistant Professor of Teaching English as a Foreign Language, Department of English Language and Literature, Faculty of Literature and Humanities, University of Guilan, Rasht, Iran; atahriri@guilan.ac.ir

⁴ Assistant Professor of Linguistics, Faculty of Literature and Foreign Languages, Islamic Azad University, South Tehran Branch, Iran; a_rabi@azad.ac.ir

Introduction: Reading comprehension is the process of reader's interaction with a written language and drawing meaning from the text. Generally, reading comprehension is a complex and multidimensional process which is done through two core processes. The first is decoding the symbols and recognizing the words, and the second is integrating the meaning of words in the context of the text (Gough & Tunmer, 1986; in: Atkinson, 2014). Learning reading comprehension is a long-term process; so it is at the end of the learning process that the adult reader can easily read different texts and draw the meaning from them. Studies show that there is a great difference between skilled and poor readers in terms of using the fundamental strategies of reading comprehension, i.e., cognitive and metacognitive strategies (Cain, Oakhill & Bryant, 2000). Weakness in the prerequisites of reading comprehension and failure in selecting the appropriate comprehension strategies are some of the important problems of students at different educational levels when reading different types of texts. Some international studies have been done on the reading comprehension such as PIRLS¹ and PISA² tests. During the recent years, the number of such studies has increased in Iran. Perhaps it could be due to the Iranian students' low performance in the PIRLS test at different time intervals which shows their weakness in reading comprehension. Despite such weak results in international tests, and doing some related research in Iran, still there is no appropriate tool for determining the reading comprehension of Iranian people, especially adults. Living in the modern society needs learning and reading various texts. Despite the importance of this issue and the quantitative progress in the number of graduate students, there is no specific criterion to determine the educated adults' level of reading comprehension. The development of higher education is a great scientific evolution that despite its positive effects has some shortcomings as well. One of the most important shortcomings is the lack of an appropriate placement tool for evaluating students' reading comprehension in order to prepare suitable educational material. The aim of this study was to develop a placement tool for evaluating adult Persian speakers' reading comprehension. The study followed three main objectives, i.e. finding text selection criteria and the related questions for reading comprehension tests, using the selected criteria for developing reading comprehension placement test and finally determining the validity and reliability of the designed test.

Questions: There were three main questions in this study: 1. What are the text and question selection criteria for developing a reading comprehension placement test? 2. Which reading comprehension placement test could be designed for adults, implementing the above-mentioned criteria? 3. Does the designed test have validity and reliability?

Method: It was a mixed methods study. The qualitative part included finding the text and question selection criteria for developing the adult Persian speakers' reading comprehension placement test and the steps of its development. Also, the quantitative part of the study included the pilot study of the mentioned test to determine its validity and reliability. The content validity of the test was checked by 5 experts. To examine its reliability, the test was distributed among two groups of MA students of

¹ The Progress in International Reading Literacy Study

² The Program for International Student Assessment

the University of Guilan, who were selected using convenience sampling method (each group 30 students) with two months interval and the level of reliability was computed using Cronbach's alpha. Along with calculation of the reliability of the test, the test items were analyzed in terms of item facility, item discrimination, and the distractors' distribution.

Findings: To select the texts, a combination of criteria introduced by the International Institute for the study of Reading Literacy for PIRLS test and Educational Testing Service (ETS) has been used. It was tried to match the selected texts in accordance with patterns of international tests and linguistic characteristics of the texts. These criteria included lexical and grammatical cohesion and also coherence of the texts. Taking into account all of the strategies underlying reading comprehension (i.e., inference-making, comprehension monitoring, text structure, etc.) the test questions were designed at 6 levels. These levels were selected based on the integration of Day and Park (2005) model and the design of TOFEL tests for reading comprehension placement tests. The content validity of the test was approved by 5 linguists, English language teaching, and Persian language teaching experts, after implementing their comments on the texts and questions. To ensure the test reliability, Cronbach's alpha coefficient was used. First, the test was distributed among a sample of 30 MA students of University of Guilan and the level of Cronbach's alpha was 0.82. Also, different levels of item analysis were conducted, including item facility, item discrimination, and the distractors' distribution. To make sure of the reliability of the test, after revising the items, and with two months interval, the test was distributed among a new sample of 30 MA students of the University of Guilan. Again, the reliability of the test was computed and Cronbach's alpha was 0.84, showing the appropriate reliability of the tool.

Conclusion: This test could be used to assess adult Persian speakers' level of reading comprehension and also its results could be used to select appropriate educational material. In the next step of this research, this tool should be distributed among a larger sample to determine its construct validity and also to compute its norms so that its results could be cited with more confidence.

Keywords: reading comprehension, placement test, test development, adults, MA students