

## استفاده از سنجش شناختی تشخیص برای سنجش نوشتن در دانش آموزان پایه ششم ابتدایی

حسن تقیان<sup>۱\*</sup>، ابراهیم خدایی<sup>۲</sup>، عباس بازرگان<sup>۳</sup>، علی مقدم زاوه<sup>۴</sup> و مسعود کبری<sup>۵</sup>

### چکیده

در بین مهارت‌های چهارگانه زبانی، نوشتن یکی از مهم‌ترین مهارت‌های برقراری ارتباط است که برای آموزش آن در همه نظام‌های آموزشی از مقطع ابتدایی برنامه‌ریزی می‌شود. پیشرفت‌های روزافزون در حوزه سنجش آموزش زمینه را برای کسب اطلاعات تشخیصی مورد نیاز برای تعیین نقاط ضعف و قوت دانش آموزان فراهم کرده است. این تحقیق قصد دارد امکان استفاده از سنجش شناختی تشخیصی برای سنجش خصیصه‌های نوشتن دانش آموزان پایه ششم ابتدایی را بررسی کند. به این منظور سیاهه واری که توسط کیم (۲۰۱۵) برای سنجش نوشتن علمی در زبان انگلیسی پیشنهاد شده است و ماتریس کیو اولیه مورد بازبینی قرار گرفت. ۸۱۶ نفر از دانش آموزان پایه ششم ابتدایی، به سه تکلیف‌های نوشتن پاسخ دادند. برآورد پارامترهای مدل با استفاده از مدل بازپارامتری شده یک‌شکل محدودشده انجام شد. شاخص‌های نسبی و مطلق برازش مدل نشان داد مدل به صورت مناسبی برازش دارد. نتیجه تحقیق نشان داد در تکالیفی که دانش آموزان را ملزم می‌کند یک متن طولانی بنویسند، ۱۷ نشان‌گر می‌تواند خصیصه‌های نوشتن را با دقت مناسبی مورد سنجش قرار دهد؛ اما در تکالیفی که مستلزم نوشتن یک متن کوتاه است، ۹ نشان‌گر برای سنجش خصیصه‌های مذکور دارای برازش مناسبی است.

کلیدواژه‌ها: نوشتن، سنجش شناختی تشخیصی، پایه ششم ابتدایی، سنجش نوشتن، مدل بازپارامتری شده یک‌شکل محدودشده، مطالعات تربیتی

رتال جامع علوم انسانی

۱. نویسنده مسئول: دانشجوی دکتری، سنجش آموزش، دانشگاه تهران، تهران، ایران. taghian@gmail.com
  ۲. دانشیار، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. khodaie@ut.ac.ir
  ۳. استاد، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. abazarga@hotmail.com
  ۴. استادیار، دانشکده روان‌شناسی و علوم تربیتی، دانشگاه تهران، تهران، ایران. amoghadamzadeh@ut.ac.ir
  ۵. استادیار، پژوهشگاه مطالعات وزارت آموزش و پرورش، تهران، ایران. maskabiri@yahoo.com
- DOI: 10.22051/jontoe.2019.23583.2462  
[https://jontoe.alzahra.ac.ir/article\\_4367.html](https://jontoe.alzahra.ac.ir/article_4367.html)

## مقدمه

در بین مهارت‌های مرتبط با سواد، نوشتن یک مهارت مهم برای همه افراد به حساب می‌آید؛ زیرا بخش قابل توجهی از فرایند ارتباط افراد با دیگران از طریق نوشتن صورت می‌گیرد. می‌توان از نوشتن برای تشریح، ترغیب یا انتقال تجربیات استفاده کرد. در یک محیط علمی، از نوشتن برای ثبت آنچه فرد در مورد یک موضوع خاص می‌داند یا فکر می‌کند استفاده می‌شود. نوشتن ماهرانه، از مهم‌ترین دستاوردهای یادگیری است، زیرا در فرایند نوشتن، منابع چندگانه شناختی به کار گرفته می‌شوند تا نویسنده بتواند ایده‌ها را تولید و سازمان‌دهی کند و با به‌کارگیری فعالیت‌های فیزیکی متن خود را نوشته و آن را بازبینی کند (کولیکوویچ، ماسون و برون، ۲۰۰۸).

بررسی پیشینه مربوط به سنجش توانایی نوشتن در جامعه دانش‌آموزان ایرانی نشان‌دهنده ضعف جدی دانش‌آموزان در این حوزه است. دانای طوسی (۱۳۹۱، به نقل از دانشگر، ۱۳۹۶) ضمن مقایسه ایران با کشورهایمانند انگلستان و آمریکا، بر ضعف مبانی نظری در اهداف آموزش نوشتن در برنامه درسی زبان فارسی تأکید کرده و ارائه مبهم و کلی هدف‌های آموزش زبان فارسی را نتیجه آن می‌داند که چارچوب مفهومی برای درس انشاء تدوین نشده است. شریفی و دانش‌پژوه (۱۳۸۳) در تحقیق خود نشان دادند عملکرد دانش‌آموزان پایه‌های مختلف در خواندن، املاء و دستور زبان و آئین نگارش در اغلب استان‌ها پایین‌تر از حد مورد انتظار (۶۰ درصد نمره آزمون) بوده است. همچنین گزارش فرهنگستان زبان و ادب فارسی از بررسی نمرات درس زبان و ادبیات فارسی سال‌های ۱۳۹۰ تا ۱۳۹۲، در دو آزمون استاندارد نشان می‌دهد با اینکه سؤالات امتحان نهایی در سطوح عالی یادگیری نیست، بیش از ۳۵ درصد کل دانش‌آموزان کشور نمره زیر ۱۲ کسب کرده‌اند کله نشان می‌دهد برای بیش از یک سوم جمعیت کل کشور اهداف آموزشی محقق نشده است (دانشگر، عابدینی و محمدی، ۱۳۹۴). موسوی، کیانی، اکبری و غفارثمر (۱۳۹۵) نیز علت عمده عدم مشارکت اساتید و دانشجویان ایرانی در تولید علمی داخلی و بین‌المللی را ضعف در مهارت‌های نگارشی آن‌ها دانسته‌اند. دانشگر (۱۳۹۶) در بررسی مهارت‌های زبانی دانش‌آموزان شهر تهران به این نتیجه رسید که



رشد مهارت‌های زبانی دانش‌آموزان با اهداف آموزش کشور همخوانی نداشته و نشان داد نمره‌های دو مهارت خواندن و نوشتن حاکی از آن است که دانش‌آموختگان در این دو مهارت به کم‌ترین درصد توانایی دست‌یافته‌اند و پیشنهاد می‌کند شیوه سنجش و ارزیابی مراحل مختلف آموزشی در ساختار رسمی کشور بازبینی شود و با توجه و دقت بیشتری به جزئیات پرداخته شود تا امکان رفع کاستی‌ها و ارتقاء آموزش فراهم آید.

هرچند سیاست‌های دفتر برنامه‌ریزی و تألیف کتب درسی در کشور ایران نیز اهداف آموزش زبان فارسی را در سه حیطه شناختی، مهارتی و عاطفی تعریف می‌کند (دفتر برنامه‌ریزی و تألیف کتب درسی، ۱۳۸۶)؛ اما ضعف در ارزشیابی صحیح و دقیق، بررسی اجرای سیاست‌های مذکور را با دشواری زیادی مواجه کرده است. از سویی بررسی تحقیق سیاست‌ها و اهداف برنامه درسی به آزمون‌هایی نیاز دارد که مهارت‌های زبانی دانش‌آموزان را به صورت دقیق بسنجد و از سویی در آزمون‌های رایج در نظام ارزشیابی آموزش و پرورش، از چنین سنجشی غفلت شده و سنجش‌های موجود قدرت اندازه‌گیری توانایی مهارت‌های زبانی دانش‌آموزان را ندارد (دانشگر، ۱۳۹۶).

سنجش تشخیصی زبان عبارت است از سنجشی که بتواند ضعف‌ها، مشکلات و حوزه‌هایی را مشخص کند که دانش‌آموزان در آن حوزه‌ها نیاز به کمک دارند (آلدرسن<sup>۱</sup>، ۲۰۰۵). اهمیت تشخیص نقاط قوت و ضعف دانش‌آموزان برای معلمان در تشخیص نیازهای مختلف دانش‌آموزان کلاس، برنامه‌ریزی آموزشی بر اساس نیازهای دانش‌آموزان و دریافت بازخورد در حین تدریس خود را نشان می‌دهد.

به‌رغم اهمیت سنجش شناختی تشخیصی در سنجش آموزش، آزمون‌های زیادی برای تهیه بازخورد تشخیصی در امور آموزشی طراحی نشده است (آلدرسن، ۲۰۰۵؛ گورین<sup>۲</sup>، ۲۰۰۷). این درحالی است که آزمون‌های تشخیصی با محتوای خصیصه‌های شناختی برای تشخیص ضعف‌ها و قوت‌های یادگیرندگان نیاز جدی آموزش است. در اغلب تحقیقات با توجه به کمبود آزمون‌های تشخیصی، از مدل‌های شناختی تشخیصی برای تهیه بازخوردهای تشخیصی از آزمون‌های غیر تشخیصی موجود استفاده شده است.

1. Alderson
2. Gorin



آلدرسن (۲۰۰۵) بیان می‌کند که سنجش تشخیصی زبان باید به گونه‌ای متفاوت با سنجش مبتنی بر عملکرد مهارت‌های زبانی انجام شود. هرچند یک نمره کلی از مهارت‌های زبانی می‌تواند برای جایابی دانش‌آموزان در یک پیوستار تک بعدی مناسب باشد، اما سنجش تشخیصی زبان باید بتواند اطلاعات جزئی‌تری از حوزه‌های مرتبط با زبان را آشکار کند.

هرچند آلدرسن (۲۰۰۵) برای سنجش مهارت‌های زبانی، استفاده از سنجش غیرمستقیم و مهارت‌های سطح پایین‌تر مانند واژگان، زمان افعال و ... را پیشنهاد می‌کند، اما به نظر می‌رسد این کار در سنجش مهارت‌های تولیدی<sup>۱</sup> مانند نوشتن، امکان‌پذیر نباشد، زیرا سنجش دانش‌آموزان از طریق مطالبی انجام می‌شود که نوشته‌اند و لازم است سنجش نوشتن بر اساس سطوح بالاتری از یادگیری انجام شود. این سؤال همچنان وجود دارد که چگونه می‌توان ابزاری برای سنجش تشخیصی عملکرد نوشتن ساخت.

بررسی پیشینه مربوط به سنجش نوشتن نشان می‌دهد، برای سنجش نوشتن از دو رویکرد کل‌نگر<sup>۲</sup> و تحلیلی<sup>۳</sup> بیشتر از رویکردهای دیگر استفاده شده است. در رویکرد کل‌نگر به هر بخش از نوشته بر اساس برداشت کلی ارزیاب، از نقاط ضعف و قوت متن، یک نمره کلی اختصاص می‌یابد (جاویز، گرت، بیکوسکی و فریس<sup>۴</sup>، ۲۰۰۳). در روش تحلیلی که سنجش خصیصه‌های چندگانه نیز نام دارد، از ارزیاب خواسته می‌شود بر اساس معیارهای چندگانه‌ای که از قبل طراحی شده است، به متن نوشته شده امتیاز دهد.

به نظر می‌رسد هر دو رویکرد فوق برای اهداف تشخیصی نامناسب باشند (کنج<sup>۵</sup>، ۲۰۱۱). رویکرد کل‌نگر برای ارزیابی متن‌هایی مناسب است که دارای همگنی درونی هستند، اما برای سنجش متن‌هایی که پیچیدگی درونی دارند نامناسب است. رویکرد تحلیلی ممکن است برای ارائه اطلاعاتی مفید بر اساس ملاک‌های چندگانه مناسب باشد، اما تحقیقات تجربی نشان داده است که ارزیابان نمی‌توانند به میزان کافی بین ملاک‌های عملکرد مختلف تمایز قائل شوند؛ به عبارت دیگر گرایش ارزیاب - که می‌تواند به عنوان اثر هاله<sup>۶</sup> مطرح شود - باعث می‌شود که

1. productive skills
2. holistic
3. Analytic
4. Jarvis, Grant, Bikowski and Ferris
5. Knoch
6. Halo effect



ادراک کلی از متن یا بخشی از متن، بر ارزیابی سایر بخش‌های متن تأثیر بگذارد (انگلهارد<sup>۱</sup>، ۲۰۱۳). این اثر هاله باعث ایجاد نیم‌رخ‌هایی یکنواخت می‌شود؛ به عبارت دیگر دانش‌آموزان به عنوان یک فرد خوب یا بد در همه ملاک‌های مختلف شناخته می‌شود. این اثر هاله می‌تواند باعث تضعیف ظرفیت تشخیصی رویکرد تحلیلی در سنجش نوشتن شود.

عدم توان تشخیصی دو رویکرد کلی‌نگر و تحلیلی، در سنجش نوشتن، باعث ایجاد لیست‌های واری برای تشخیص مشکلات نوشتن در برخی تحقیقات شده است (بانرجی و وال<sup>۲</sup>، ۲۰۰۶؛ کیم<sup>۳</sup>، ۲۰۱۵؛ استراثر، لاپادات و مک‌میلان<sup>۴</sup>، ۲۰۱۳). سیاهه واری تشخیصی، متن را به تعداد زیادی از نشانگرهای جزئی متناسب با جنبه‌های مختلف نوشتن، تجزیه می‌کند. یکی از آخرین لیست‌های واری که توسط کیم (۲۰۱۵) ارائه شده است، شامل ۳۵ نشانگر است که برای سنجش پنج مهارت نوشتن در زبان انگلیسی طراحی شده است.

سیاهه واری کیم (۲۰۱۵) برای سنجش نوشتن در آزمون انگلیسی به عنوان زبان خارجی (نافل)<sup>۵</sup> طراحی شده است. سیاهه مذکور برای هر نوشته، ۳۵ نشانگر را بررسی می‌کند که نشانگرهای مذکور با پنج خصیصه نوشتن در ارتباط است. خصیصه‌های مذکور عبارت هستند از تولید محتوا، سازماندهی متن، دستور زبان، واژگان و عرف نوشتن. تعاریفی که کیم (۲۰۱۵) برای خصیصه‌های مذکور ارائه کرده است عبارت‌اند از (۱) تولید محتوا: این خصیصه مشخص می‌کند که نویسنده تا چه اندازه موضوع تکلیف را فهمیده و می‌تواند درباره آن توضیح دهد. نویسندگانی که در این مهارت تسلط دارند می‌توانند با درک خوبی که از موضوع دارند، با ارائه شواهد و مثال‌های کافی، موضوع را تبیین کنند. برای تسلط در این خصیصه لازم است فرد بتواند به سؤال پاسخ صحیح بدهد، ایده پرداز می‌ماند، از ایده‌های خود با استفاده از شواهد و مثال‌ها پشتیبانی کند، به ویژگی‌های جامعه هدف (خوانندگان) توجه کافی داشته باشد. (۲) سازماندهی: شیوه‌ای که نویسنده بر اساس آن، ایده‌های خود را سازماندهی و بیان می‌کند. نویسنده‌ای که در این زمینه تسلط دارد، به طور کلی می‌تواند عناصر متن خود را

1. Engelhard
2. Banerjee and Wall
3. Kim
4. Struthers, Lapadat and MacMillan
5. Test of English as a Foreign Language (TOEFL)



به خوبی در داخل و بین پاراگراف‌ها قرار داده و با استفاده از ابزارهای یکپارچه‌سازی و انتقال‌دهنده‌های مناسب مطالب را به صورت منسجم ارائه دهد. (۳) دستور زبان: متن‌هایی که با کیفیت نوشته شده‌اند، قواعد دستور زبان را نیز به خوبی رعایت کرده‌اند؛ به‌خصوص اینکه در بعضی موارد، خطاهای مرتبط با دستور زبان می‌تواند مانع درک صحیح مطلب نوشته شده شود. برخی از نشانه‌های تسلط افراد بر دستور زبان عبارت‌اند از: به‌کارگیری صحیح زمان و شخص افعال، استفاده از حروف اضافه و ضمایر به صورت مناسب، علامت‌گذاری صحیح در متن، به‌کارگیری شکل صحیح کلمات (اسم، فعل و حرف). (۴) واژگان: این خصیصه، توان نویسنده در استفاده پیشرفته و پیچیده از مجموعه واژگان را نشان می‌دهد؛ نویسنده‌ای که در این زمینه قوی است، با دانستن شیوه استفاده از کلمات، از طیف گسترده‌ای از کلمات پیچیده استفاده می‌کند و می‌داند که چگونه کلمات را ترکیب کند. برای ارزیابی این خصیصه لازم است به نشانگرهای مانند استفاده فرد از مجموعه واژگان مناسب به‌منظور نیل هدف در نوشته، املا صحیح کلمات، تنوع واژگان ساده و پیچیده و کاربرد صحیح کلمه در جمله توجه کرد. (۵) عرف نوشتن: توانایی استفاده از قواعد نوشتن در متن‌های مختلف مانند نشانه‌گذاری صحیح، تورفتگی در ابتدای پاراگراف، لحن مناسب برای بیان هدف و ...

کیم (۲۰۱۵) نشان داد با استفاده از سیاهه مذکور می‌تواند پنج خصیصه فوق را با رویکرد شناختی تشخیصی بسنجد. همچنین تحقیق مذکور نشان داد ماتریس کیو ارائه شده می‌تواند ارتباط بین نشان‌گرهای سیاهه و آرسی مذکور و خصیصه‌های پنج‌گانه را به خوبی نشان دهد. شی<sup>۱</sup> (۲۰۱۷) نشان داد با انجام ویرایش‌های اندکی در نشان‌گرها و ساختار ماتریس کیو امکان استفاده از سیاهه و آرسی کیم (۲۰۱۵) برای سنجش نوشتن علمی در دانشگاه‌های هنگ کنگ وجود دارد.

همان‌گونه که ذکر شد در چند تحقیق، ضعف جدی دانش‌آموزان ایرانی در حوزه نوشتن مشاهده شده و همین وضعیت باعث شده است که مهارت‌های نوشتن در سایر مقاطع تحصیلی نیز تحت تأثیر قرار گیرد، به صورتی که موسوی و همکاران (۱۳۹۵) علت عمده عدم مشارکت اساتید و دانشجویان ایرانی در تولید علمی داخلی و بین‌المللی را ضعف در مهارت‌های نگارشی آن‌ها دانسته‌اند. از سوی دیگر ظرفیت سنجش شناختی تشخیصی برای تشخیص نقاط



ضعف و قوت دانش آموزان این سؤال را به وجود می آورد که آیا می توان از این ظرفیت برای بررسی خصیصه های نوشتن دانش آموزان و کشف نقاط ضعف آن استفاده کرد؟ از سوی دیگر کیم (۲۰۱۵) نشان داد نشان گرهای سیاهه واری و ماتریس کیوی پیشنهادی قادر است به صورت دقیق و قابل اعتمادی افراد مسلط و غیر مسلط را از هم تشخیص دهد. بنابراین، با توجه به اینکه اغلب نشان گرهای موجود در سیاهه واری مذکور مربوط به مهارت های مختلف نوشتن به صورت عمومی است و این امکان وجود دارد که همان نشان گرها برای سنجش نوشتن در زبان های دیگر نیز قابلیت استفاده داشته باشد، این تحقیق قصد دارد امکان سنجش شناختی تشخیصی نوشتن در دانش آموزان ابتدایی ایرانی را با استفاده از سیاهه واری مذکور، بررسی کند. در این راستا سعی شده است با ویرایش سیاهه واری و بررسی دوباره اعتبار ماتریس کیو، نیم رخ خصیصه های دانش آموزان استخراج شود؛ بنابراین، تحقیق حاضر به دنبال پاسخ گویی به سؤال های زیر است. (۱) کدام یک از نشانگرهای سیاهه واری کیم (۲۰۱۵) برای سنجش خصیصه های نوشتن دانش آموزان ایرانی کاربرد دارد؟ (۲) برای استفاده از سیاهه واری کیم (۲۰۱۵) در ماتریس کیوی ارائه شده در تحقیق مذکور چه تغییراتی باید اعمال کرد؟

اجرای این تحقیق در چهار مرحله انجام شد. در مرحله اول با استفاده از نظر شش نفر از معلمان ابتدایی و متخصصان زبان و ادبیات فارسی، سیاهه واری کیم (۲۰۱۵) مورد بازنگری قرار گرفت نشانگرهای سیاهه واری مذکور، متناسب با اقتضات زبان فارسی و دانش کسب شده دانش آموزان تا پایه ششم ابتدایی، به روز رسانی شد. در مرحله دوم، ارتباط بین سیاهه واری نهایی، با خصیصه های ماتریس کیو، توسط همان متخصصین بازنگری شد و ماتریس کیو نهایی تدوین شد. در مرحله سوم سؤال های آزمون با نظر معلمان ابتدایی تدوین شد. در مرحله بعد، آزمون نهایی، اجرا و پارامترهای مدل برآورد شد.

### ابزار گردآوری اطلاعات (سیاهه واری)

ابزار گردآوری مورد استفاده در این تحقیق سیاهه واری است که کیم (۲۰۱۵) برای سنجش مهارت های نوشتن در زبان انگلیسی طراحی کرده است. در این تحقیق با استفاده از نظرات شش نفر از معلمان ابتدایی و متخصصان زبان و ادبیات



فارسی، فهرست نشانگرهای مذکور بازنگری شد و برخی از نشانگرها که مختص ساختار زبان و ادبیات انگلیسی بود و در زبان فارسی کاربردی ندارد (مانند استفاده از حروف بزرگ در ابتدای جمله و اسامی خاص، یا استفاده از حرف تعریف و ...) از فهرست مذکور حذف شد. علاوه بر این برخی از نشانگرها با دانش و توانمندی‌های مورد انتظار دانش‌آموزان ابتدایی مطابقت نداشتند که موارد مذکور نیز از فهرست مذکور حذف شدند. در نهایت نشانگرهای نهایی برای ارزیابی نوشته‌های آزمودنی‌ها به شرح زیر مشخص شد.

- ۱- ارائه پاسخ صحیح به سؤال؛ ۲- بیان روشن و واضح در نوشته؛ ۳- مختصر و مفید بودن و عدم وجود حشو در نوشته؛ ۴- سازماندهی ایده‌ها و مطالب در متن؛ ۵- عدم وجود خطاهای قواعدی در متن؛ ۶- وجود جملات متعدد در متن؛ ۷- وجود جملات مرکب در متن
- ۸- عدم تأثیرگذاری منفی اشتباهات دستوری و زبان‌شناختی بر درک مطلب؛ ۹- به کار بردن زمان مناسب برای فعل‌ها؛ ۱۰- به کار بردن شخص مناسب برای فعل‌ها؛ ۱۱- به کار بردن حروف اضافه مناسب؛ ۱۲- به کار بردن ضمیر به صورت مناسب؛ ۱۳- استفاده از واژگان پیچیده؛ ۱۴- استفاده از واژگان مناسب برای بیان هدف؛ ۱۵- استفاده از جمله‌بندی مناسب؛ ۱۶- به کارگیری نوع کلمات (اسم، فعل و حرف) به صورت صحیح؛ ۱۷- املا صحیح کلمات؛ ۱۸- استفاده از علامت‌گذاری مناسب و ۱۹- وجود لحن مناسب در کل نوشته.

### تکلیف‌های آزمون

سؤالات آزمون نوشتن به منظور سنجش خصیصه‌های مذکور یا در نظر گرفتن موارد ذیل طراحی شدند.

الف- سؤالات نوشتن به گونه‌ای باشد که دانش‌آموز برای پاسخ‌گویی به آن لازم باشد، یک متن یکپارچه و هدفمند بنویسد. ب- سؤال‌های نوشتن باید متناسب با توانمندی دانش‌آموزان در پاسخ‌گویی به سؤال بر اساس دانش کسب شده دانش‌آموز تا پایه ششم ابتدایی باشد. ج- برخی از سؤالات نوشتن مبتنی بر متونی باشد که از قبل نوشته شده است و از آزمودنی خواسته شود بر اساس متن مذکور پاسخ مکتوبی را ارائه دهد.

بر این اساس سه نوع تکلیف برای بخش نوشتن طراحی شد. در نوع اول از دانش‌آموز خواسته شده است درباره موضوع از قبل تعیین شده‌ای چند بند بنویسد. در نوع دوم از





آزمودنی خواسته شده است با خواندن یک متن که یک داستان کوتاه بود، خلاصه‌ای از آن را تهیه کند و در نوع سوم از آزمودنی خواسته شده است متن‌های کوتاه در حد یک یا دو جمله بنویسد. به منظور کسب اطمینان نسبت به متناسب بودن سؤالات با میزان توانمندی‌های دانش‌آموزان، طراحی سؤالات نوشتن با مشورت دو نفر از معلمان باتجربه ابتدایی و یکی از متخصصان رشته زبان و ادبیات فارسی انجام شد. بر این اساس مقرر شد از دانش‌آموزان خواسته شود درباره آداب دوستی چند بند بنویسند. از این پس به این تکلیف، تکلیف بندنویسی می‌گوییم. همچنین مقرر شد از دانش‌آموزان خواسته شود پس از مطالعه یک متن، داستان را به صورت خلاصه بازنویسی کند. از این پس به این تکلیف، تکلیف خلاصه‌نویسی می‌گوییم. معلمان ابتدایی تأکید زیادی داشتند که دانش‌آموزان باید بتوانند شعرها یا متن‌های کهن را نیز به فارسی روان بازنویسی کنند. بنابراین، مقرر شد در یک سؤال، با ارائه یک بیت شعر، از دانش‌آموز خواسته شود که بیت مذکور را به نثر روان بازنویسی کند. از این پس به این تکلیف، تکلیف بازنویسی می‌گوییم. سؤال‌های آزمون پیوست شده است.

پاسخ‌های نوشتاری هر یک از دانش‌آموزان توسط دو نفر به صورت مستقل نمره‌گذاری شد. هر یک از دو ارزیاب در یک سیاهه و آرسی مشخص می‌کردند که هر نوشته واجد کدام یک از نشان‌گرهای فهرست فوق است. در صورت وجود تناقض در نحوه نمره‌دهی دو ارزیاب به هر یک از پاسخ‌ها، ضمن تبادل نظر دو ارزیاب و در صورت نیاز استفاده از یک داور دیگر که فارغ‌التحصیل دکتری رشته ادبیات فارسی بود، تناقضات برطرف می‌شد.

راهکار ۲-۱۹ سند تحول بنیادین آموزش و پرورش بر طراحی و اجرای ارزشیابی بر اساس استانداردهای ملی برای گذر از دوره تحصیلی تأکید کرده است. از سوی دیگر نتیجه آموخته‌های دانش‌آموزان در مقطع ابتدایی، در توانمندی‌های آن‌ها در پایه ششم منعکس می‌شود و سنجش مهارت نوشتن دانش‌آموزان ابتدایی و کشف نقاط ضعف و قوت خصیصه‌های نوشتن ایشان می‌تواند اطلاعات مفیدی را به منظور رفع اشکالات در فرایند آموزش در مقطع ابتدایی در اختیار معلمان و برنامه‌ریزان آموزشی قرار دهد؛ بنابراین، جامعه آماری این تحقیق را کلیه دانش‌آموزان ابتدایی تشکیل می‌دهند که در سال تحصیلی ۹۷-۱۳۹۶ در پایه ششم ابتدایی شهر قم در حال تحصیل بودند. بر اساس آمار اداره کل آموزش و پرورش قم، در سال تحصیلی ۹۷-۱۳۹۶ در مجموع در همه مدارس ابتدایی چهار ناحیه شهر



قم (۵۴۲) مدرسه ابتدایی)، ۱۹۳۴۹ دانش آموز در پایه ششم در حال تحصیل بودند. درباره اندازه نمونه مورد نیاز به منظور دست‌یابی به شاخص‌های برازش مناسب در انتخاب مدل شناختی تشخیصی و ماتریس کیوی مناسب اطلاعات زیادی وجود ندارد (لی و لی<sup>۱</sup>، ۲۰۱۶). برخی تحقیقات اثر اندازه نمونه بر برآورد پارامترها در مدل شناختی تشخیصی لگاریتم خطی را بررسی کرده و به این نتیجه رسیدند که نمونه‌های با اندازه ۵۰۰ (چوی، تمپلین، کوهن و اتوود<sup>۲</sup>، ۲۰۱۰، به نقل از لی و لی، ۲۰۱۶) یا ۱۰۰۰ (کونینا-هبنیچت، راب و ویلهلم<sup>۳</sup>، ۲۰۱۲) برای برآورد مقادیر ثابت و اثرات اصلی کافی است، اما برآورد اثرات مرتبه دوم و مراتب بالاتر حتی با اندازه نمونه ۴۰۰۰ (چوی و همکاران، ۲۰۱۰، به نقل از لی و لی، ۲۰۱۶) یا ۱۰۰۰۰ (کونینا-هبنیچت و همکاران، ۲۰۱۲) نیز گاهی با مشکلات جدی مواجه می‌شود. چسوی و همکاران (۲۰۱۰)، نقل در لی و لی، ۲۰۱۶) عنوان می‌کنند در شرایطی که ماتریس کیوی خوب تعیین شده باشد، شاخص‌های برازش نسبی با اندازه نمونه ۲۰۰ نیز می‌تواند به انتخاب مدل مناسب منجر شود؛ بنابراین، در این تحقیق اندازه نمونه ۸۰۰ دانش آموز پایه ششم در نظر گرفته شد. از آنجایی که معمولاً در این گونه تحقیقات، به دلایل مختلف، برخی از افراد به آزمون به صورت کامل پاسخ نمی‌دهند یا با بی‌دقتی پاسخ می‌دهند، حجم نمونه کمی بیش‌تر در نظر گرفته شده و ۸۵۰ دانش آموز آزموده شد. در نهایت پس از بررسی و کنار گذاشتن مواردی که به دلیل نقص در پاسخ‌گویی، قابل استفاده نبودند، پاسخ‌های ۸۱۶ نفر برای تحلیل آماده شد.

برای انتخاب نمونه دانش آموزان ابتدا به صورت متناسب با تعداد دانش آموزان، تعداد نمونه بین چهار ناحیه آموزش و پرورش شهر قم توزیع شد و سهم هر ناحیه مشخص شد. در مرحله بعد با تقسیم تعداد نمونه مربوط به هر ناحیه بر میانگین تعداد دانش آموزان مدارس آن ناحیه، تعداد احتمالی مدارس مورد نیاز مشخص شد. در مرحله بعد به صورت تصادفی مدارس مورد نیاز برای اجرای طرح انتخاب شد. در نهایت بنا بر مراجعه به مدارس پس از هماهنگی با مدیر مدرسه، همه دانش آموزان پایه ششم مدرسه به سؤالات آزمون پاسخ دادند. از آنجا که در تحقیق کیم (۲۰۱۵) تحلیل‌های کمی با استفاده از مدل بازپارامتری شده یک‌شکل

1. Lei and Li
2. Choi, Templin, Cohen and Atwood
3. Kunina-Habenicht, Rupp and Wilhelm



محدود شده<sup>۱</sup> انجام شده است؛ در این تحقیق نیز به منظور انجام تحلیل‌های کمی، بررسی اعتبار ماتریس کیو و برآورد پارامترهای سوالات از همان مدل استفاده شده تا نتایج این تحقیق قابلیت مقایسه با نتایج تحقیق مذکور را نیز داشته باشد. تحلیل‌ها با استفاده از بسته CDM نرم‌افزار R انجام شد.

### یافته‌ها

این تحقیق به دنبال پاسخ‌گویی به دو سؤال بود. سؤال‌های مذکور عبارت‌اند از ۱) کدام یک از نشانگرهای سیاهه واریس کیم (۲۰۱۵) برای سنجش خصیصه‌های نوشتن دانش‌آموزان ایرانی کاربرد دارد؟ (۲) برای استفاده از سیاهه واریس کیم (۲۰۱۵) چه تغییراتی باید در ماتریس کیوی ارائه شده در تحقیق مذکور اعمال کرد؟ از آنجا که بررسی مناسب بودن نشان‌گرهای سیاهه واریس، مستلزم بررسی ماتریس کیو است، پاسخ به هر دو سؤال به صورت همزمان حاصل می‌شود. بنابراین، هر دو سؤال تحقیق به صورت همزمان بررسی می‌شود. نتایج حاصل از بررسی برازش مدل نشان‌دهنده تفاوت سه تکلیف مختلف در برازش مدل بود. در تکلیف‌های بازنویسی و خلاصه‌نویسی نشان‌گرهای مختصر و مفید بودن و عدم وجود حسو در نوشته، سازماندهی ایده‌ها و مطالب در متن، وجود جملات متعدد در متن، وجود جملات مرکب در متن، عدم تأثیرگذاری منفی اشتباهات دستوری و زبان‌شناختی بر درک مطلب، به کار بردن ضمائر به صورت مناسب، استفاده از واژگان پیچیده و وجود لحن مناسب در کل نوشته، برازش مناسبی نداشتند. همچنین نشانگرهای عدم تأثیرگذاری منفی اشتباهات دستوری و زبان‌شناختی بر درک مطلب و استفاده از واژگان پیچیده، در همه سؤال‌ها از برازش مناسبی برخوردار نبودند. بررسی پارامترهای مدل برای نشان‌گرهای مذکور نشان‌دهنده مقادیر  $\pi^*$  کوچک (کمتر از ۰/۶) و مقادیر  $\tau^*$  بزرگ بود. عبارات دیگر برای نشان‌گرهای مذکور،

#### 1. Reduced Reparametrized Unified Model (R-RUM)

۲. پارامتر خط پایه ( $\pi^*$ ) عبارت است از احتمال ارائه پاسخ صحیح به یک سؤال با این فرض که آزمودنی در همه مهارت‌های مورد نیاز به حد تسلط رسیده است. مقادیر کم‌تر از ۰/۶ برای این پارامتر به این معنی است که سؤال برای اندازه‌گیری خصیصه‌ها اختصاص یافته، مشکل است و احتمالاً لازم است خصیصه‌های سخت‌تری را به این سؤال اختصاص یابد. پارامتر جریمه ( $\tau^*$ ) مشخص می‌کند که تا چه میزان تسلط در یک خصیصه بر احتمال ارائه پاسخ صحیح به سؤال اثر دارد. مقادیر کم این پارامتر نشان‌دهنده قدرت بالای سؤال برای تشخیص تسلط بر خصیصه است. مقادیر بالاتر از ۰/۹ برای این پارامتر نشان می‌دهد سؤال مذکور قادر به تشخیص خصیصه مربوطه نیست.

نشانگرها برای مهارت‌های مشخص شده سخت بوده و نشان‌گرها نیز توان تشخیص افراد مسلط از افراد غیرمسلط را نداشتند. در بررسی بیشتر مشخص شد موارد مذکور مربوط به تکالیفی است که در آن‌ها پاسخ‌گو ملزم به ارائه یک جمله یا یک متن کوتاه (حداکثر سه جمله) است؛ بنابراین، نشانگرهای مذکور در سؤال‌ها حذف شد و یک بار دیگر مدل بررسی شد. در جدول‌های ۱ تا ۳ تخمین پارامترهای مدل برای هر سه تکلیف (بعد از حذف نشان‌گرهای نامناسب) ارائه شده است. در هر سه جدول علامت × نشان می‌دهد نشان‌گر مذکور پس از بررسی پارامترهای مدل، حذف شده‌اند.

جدول ۱: برآورد پارامترهای نشان‌گرها برای تکلیف بندنویسی

نشان‌گرها	$\pi^*$	$\Gamma^*_{con}$	$\Gamma^*_{org}$	$\Gamma^*_{gra}$	$\Gamma^*_{voc}$	$\Gamma^*_{mch}$
۱	۰/۹۱۵	۰/۷۱۰				
۲	۰/۹۴۵	۰/۴۸۰	۰/۶۲۵			
۳	۰/۹۰۵	۰/۴۷۰	۰/۵۲۵	۰/۶۵۵		
۴	۰/۹۱۰		۰/۵۱۰			
۵	۰/۸۴۰		۰/۵۵۰			
۶	۰/۷۹۰		۰/۸۲۰			۰/۵۵۵
۷	۰/۵۷۰		۰/۴۸۰			
۸	×		×			
۹	۰/۷۴۰		۰/۴۳۰			
۱۰	۰/۷۰۵		۰/۵۱۵			
۱۱	۰/۸۳۰		۰/۴۱۰			
۱۲	۰/۶۷۰		۰/۴۸۰			
۱۳	×		×			
۱۴	۰/۷۷۰		۰/۳۹۵			
۱۵	۰/۹۷۰		۰/۶۳۵			
۱۶	۰/۸۴۵		۰/۳۵۰			
۱۷	۰/۸۳۰		۰/۶۶۰			۰/۵۵۰
۱۸		۰/۹۴۰				۰/۳۶۵
۱۹		۰/۷۱۵				۰/۷۶۰



جدول ۲: برآورد پارامترهای نشان‌گرها برای تکلیف خلاصه‌نویسی

نشان‌گرها	$\pi^*$	$\Gamma_{con}^*$	$\Gamma_{org}^*$	$\Gamma_{gra}^*$	$\Gamma_{voc}^*$	$\Gamma_{mch}^*$
۱	۰/۸۳۳	۰/۵۹۶				
۲	۰/۹۸۳	۰/۵۷۶	۰/۷۱۳			
۳	x	x	x			
۴	x		x			
۵	۰/۷۶۹			۰/۵۲۸		
۶	x			x		
۷	x			x		
۸	x			x		
۹	۰/۸۱۹			۰/۴۵۱		
۱۰	۰/۹۴۱			۰/۴۸۴		
۱۱	۰/۷۸۲			۰/۴۴۳		
۱۲	x			x		
۱۳	x			x		
۱۴	۰/۶۳۸			۰/۳۵۵		
۱۵	۰/۶۳۶			۰/۷۶۲		
۱۶	۰/۸۱۱			۰/۳۰۵		
۱۷	۰/۷۹۷			۰/۶۹۹		۰/۶۵۵
۱۸	۰/۸۰۴					۰/۳۰۳
۱۹	x					x

پژوهشگاه علوم انسانی و مطالعات فرهنگی  
 پرتال جامع علوم انسانی

جدول ۳: برآورد پارامترهای نشان‌گرها برای تکلیف بازنویسی

نشان‌گرها	$\pi^*$	$\Gamma_{con}^*$	$\Gamma_{org}^*$	$\Gamma_{gra}^*$	$\Gamma_{voc}^*$	$\Gamma_{mch}^*$
۱	۰/۵۹۹	۰/۶۳۹				
۲	۰/۸۷۳	۰/۵۶۲	۰/۶۴۴			
۳	×	×	×			
۴	×		×			
۵	۰/۶۷۷			۰/۵۱۷		
۶	×			×		
۷	×			×		
۸	×			×		
۹	۰/۶۶۴			۰/۴۶۰		
۱۰	۰/۶۸۳			۰/۴۳۳		
۱۱	۰/۶۸۵			۰/۴۲۶		
۱۲	×			×		
۱۳	×			×		
۱۴	۰/۹۵۷			×	۰/۳۲۴	
۱۵	۰/۷۹۴			۰/۷۳۰		
۱۶	۰/۹۳۰			۰/۲۸۰		
۱۷	۰/۷۸۰			۰/۶۶۷		۰/۶۵۵
۱۸	۰/۸۵۱					۰/۳۴۳
۱۹	×					×

درباره نشان‌گر شماره ۱ هر تکلیف بازنویسی، مقدار  $\pi^*$  کوچک‌تر از ۰/۶ بوده و در نتیجه این نشان‌گر یک نشان‌گر سخت به حساب می‌آید. به عبارت دیگر داشتن حصیصه تولید محتوا» برای پاسخ‌گویی به این سؤال کافی نیست؛ اما برای این نشان‌گر مقدار  $\Gamma$  متناسب است و در نتیجه این نشان‌گر توان تشخیص افراد غیر مسلط از افراد مسلط را دارد. بنابراین، نشان‌گر مذکور حذف نشده است. همه نشان‌گرهای حذف شده دارای مقادیر  $\pi^*$  کوچک و مقادیر  $\Gamma$  بزرگ بودند.

برای بررسی برازش ماتریس کیو، از برخی ملاک‌های مطلق برای بررسی برازش مدل استفاده شد. ملاک  $MADcor^1$  میانگین قدر مطلق تفاوت بین مقادیر مشاهده شده و برآورد شده همبستگی‌های بین جفت سؤالات، در بین همه جفت‌های ممکن و  $MADres$  میانگین قدر مطلق باقیمانده

1. Mean absolute difference for the item-pair correlations



کوواریانس بین همه جفت سؤالات دو شاخص برازش مطلق اصلی است (لی، هانتز و لی، ۲۰۱۶). این دو شاخص در سطح جفت سؤالات نیز محاسبه می‌شود و شاخصی برای برازش مناسب سؤال با مدل و ماتریس کیو است. به منظور ایجاد امکان آزمون مقایسه MADcor با صفر، از تبدیل فیشر استفاده شد که تبدیل مذکور با نشانه‌گذاری  $z(\rho)$  مشخص شده است (لی، هانتز<sup>۱</sup> و لی، ۲۰۱۶). در نرم‌افزار R این آماره در سطح جفت سؤالات و در سطح کل آزمون محاسبه و آزمون می‌شود. همچنین برای بررسی برازش سؤال‌ها با مدل، از آماره  $\chi^2$  تفاوت نسبت مشاهده شده و برآورد شده برای هر سؤال و آماره RMSEA استفاده شد.

در تحلیل نهایی نشان‌گرهای هر سه تکلیف، مقدار RMSEA برای همه نشانگرها کم‌تر از ۰/۱ شده و برای کل مدل نیز مقدار  $z(\rho)$  کوچک و غیرمعنادار شد.

شاخص‌های برازش مدل (جدول ۴) نشان داد، هر چند ماتریس کیو در این تحقیق دارای نشان‌گرهای کم‌تری نسبت به ماتریس کیو تحقیق کیم (۲۰۱۵) است، اما این کم‌تر شدن تعداد نشان‌گرها تأثیر زیادی بر برازش مدل نداشته است، به صورتی که شاخص‌های برازش مدل همچنان نشان‌دهنده تفاوت ناچیز بین مقادیر پیش‌بینی شده و مشاهده شده است که برازش مدل را تأیید می‌کند.

جدول ۴: شاخص‌های برازش مدل به تفکیک سه تکلیف نوشتن

	تکلیف بند نویسی		تکلیف خلاصه نویسی		تکلیف یاز نویسی	
	برآورد	انحراف	برآورد	انحراف	برآورد	انحراف
MADcor	۰/۰۳۵	۰/۰۲۵	۰/۰۴۸	۰/۰۳۶	۰/۰۱۴	۰/۰۶۳
SRMSR	۰/۰۴۸	۰/۰۴۲	۰/۰۳۱	۰/۰۵۵	۰/۰۶۳	۰/۰۷۹
MADQ3	۰/۰۲۳	۰/۰۸۱	۰/۰۳۹	۰/۰۶۷	۰/۰۴۷	۰/۰۴۵
MADaQ3	۰/۰۴۱	۰/۰۴۳	۰/۰۵۲	۰/۰۴۹	۰/۰۲۳	۰/۰۸۰
MADRESIDCOV*100	۰/۵۹	۰/۰۴۴	۰/۴۳	۰/۰۵۸	۰/۵۵	۰/۰۱۲

برای بررسی همسانی<sup>۲</sup> و دقت<sup>۳</sup> طبقه‌بندی توسط مدل شناختی تشخیصی، از شاخص‌هایی

1. Hunter
2. consistency
3. accuracy

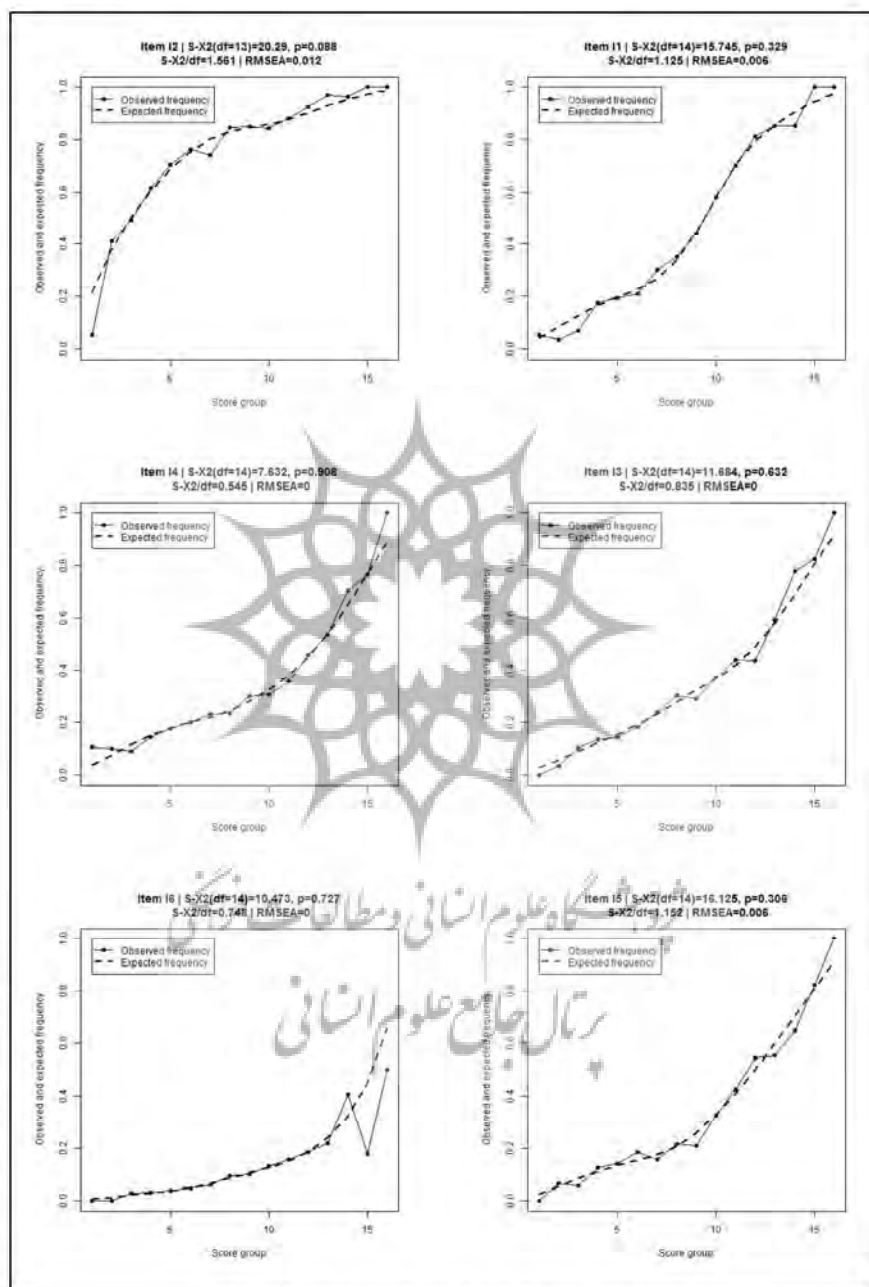
استفاده شد که کوی، گیرل و چانگ<sup>۱</sup> (۲۰۱۲) پیشنهاد داده‌اند. کوی و همکاران (۲۰۱۲) شاخص همسانی طبقه‌بندی را به صورت احتمال همسانی در طبقه‌بندی یک نمونه تصادفی از دانش‌آموزان در دو اجرا یا دو فرم موازی آزمون تعریف کردند. همچنین احتمال دقت در طبقه‌بندی یک نمونه تصادفی از دانش‌آموزان بر اساس پاسخ‌های آن‌ها به سؤالات حاصل از یک‌بار اجرای آزمون، شاخص دقت طبقه‌بندی نام دارد. مدلی با درجه همسانی و دقت طبقه‌بندی پایین می‌تواند به تفسیر نادرست درباره نیم‌رخ خصیصه‌های دانش‌آموزان و ایجاد خطا در تصمیم‌گیری‌های آموزشی منجر شود. چنین خطایی علاوه بر اتلاف زمان و تلاش‌های معلمان و دانش‌آموزان، ممکن است آینده شغلی و تحصیلی دانش‌آموزان را تحت تأثیر قرار دهد (کوی و همکاران، ۲۰۱۲). در جدول ۵ شاخص‌های همسانی و دقت طبقه‌بندی توسط مدل برای سه تکلیف آزمون نوشتن ارائه شده است.

جدول ۵: شاخص‌های همسانی و دقت طبقه‌بندی مدل سه تکلیف آزمون نوشتن

تکلیف بنام نویسی	تکلیف خلاصه نویسی		تکلیف باز نویسی		خصیصه‌ها
	همسانی	دقت	همسانی	دقت	
تولید محتوا	۰/۸۶۸	۰/۸۰۴	۰/۹۲۱	۰/۷۱۹	۰/۶۹۵
سازماندهی	۰/۸۸۴	۰/۹۰۹	۰/۸۹۳	۰/۸۲۸	۰/۸۴۱
دستور زبان	۰/۹۵۵	۰/۸۲۳	۰/۸۰۹	۰/۷۶۱	۰/۹۷۷
دانش واژگان	۰/۹۷۷	۰/۸۴۴	۰/۸۹۴	۰/۹۷۹	۰/۸۳۰
عرف نوشتاری	۰/۸۸۵	۰/۷۰۸	۰/۷۹۳	۰/۸۹۳	۰/۹۰۲

شاخص‌های همسانی و دقت طبقه‌بندی که در جدول ۵ ارائه شده است، نشان‌دهنده همسانی و دقت مناسب طبقه‌بندی آزمودنی‌ها توسط مدل شناختی تشخیص مورد استفاده در سه تکلیف آزمون نوشتن است. به عنوان یکی دیگر از شاخص‌های برآزش مدل، از نمودار فراوانی‌های مشاهده شده و برآورد شده استفاده شده است. تفاوت کوچک این دو مقدار نشان‌دهنده برآزش خوب مدل است. به دلیل محدودیت تعداد صفحات، در شکل ۱ نمودار مذکور فقط برای ۶ نشان‌گر نشان داده شده است.





نمودار ۱: نمودار فراوانی‌های مشاهده شده و برآورد شده

در مراحل تدوین آزمون، محتوای آزمون، بر اساس اهداف آموزشی و با در نظر گرفتن محتوای درس فارسی و با نظر معلمان ابتدایی و متخصصان حوزه زبان و ادبیات فارسی تدوین شده است. بر این اساس می‌توان گفت که این آزمون از روایی محتوایی لازم برخوردار است. شاخص‌های کمی برازش سؤال و مدل نیز مؤید روایی مناسب آزمون بود. در این مرحله با حذف نشان‌گرهای نامناسب، برازش سؤال‌های باقیمانده وضعیت مطلوبی را نشان داد. علاوه بر این برای اطمینان از روایی آزمون از روش مقایسه گروهی استفاده شد که لیتون و گیرل (۲۰۰۷) برای بررسی روایی آزمون‌های شناختی تشخیصی، پیشنهاد داده‌اند.

به منظور بررسی روایی آزمون با استفاده از روش مقایسه گروهی، نیم‌رخ خصیصه‌های برخی از دانش‌آموزان با همکاری معلم آن‌ها بررسی شد. اجرای این روش با محدودیت مواجه بود، زیرا تعداد معلمان آشنا با این خصیصه‌ها و حاضر به همکاری کم بود. از سویی معلم‌ها نیز با همه دانش‌آموزان به صورت کامل آشنایی نداشتند به صورتی که در بیان وضعیت تسلط دانش‌آموزان در مورد چند دانش‌آموز به خوبی اظهار نظر می‌کردند، اما در مورد سایر دانش‌آموزان با تردید اظهار نظر می‌کردند. با وجود این، نیم‌رخ خصیصه‌های دانش‌آموزان با اطلاعاتی مقایسه شد که معلم بر اثر تجربه کسب کرده بود. در مواردی که نیم‌رخ خصیصه‌ها به صورت واضح وجود یا عدم وجود خصیصه خاصی را در دانش‌آموز نشان می‌داد، ارزیابی معلم نیز با این یافته انطباق داشت.

### بحث و نتیجه‌گیری

این تحقیق قصد داشت امکان به‌کارگیری سیاهه واریسی را، که کیم (۲۰۱۵) پیشنهاد داده بود، برای سنجش نوشتن دانش‌آموزان پایه ششم ابتدایی بررسی کند. برای انجام این کار ابتدا نشان‌گرهای سیاهه واریسی مذکور مورد تجدید نظر قرار گرفت. در این مرحله نشان‌گرهایی حذف شدند که با ساختار زبان فارسی هم‌خوانی نداشت. از آنجا که کیم (۲۰۱۵) سیاهه واریسی خود را برای سنجش نوشتن برای دانشجویان طراحی کرده بود، برخی از نشان‌گرها با اقتضائات نوشتن در پایه ششم ابتدایی هم‌خوانی نداشت؛ بنابراین، نشان‌گرهای مذکور نیز حذف شدند. سیاهه واریسی با ۱۹ نشان‌گر در سه تکلیف نوشتن در نمونه‌ای از ۸۱۶ دانش‌آموز بررسی شد. سه تکلیف بر اساس انتظارات معلمان ابتدایی از توانایی نوشتن دانش‌آموزان



طراحی شد. نتایج نشان داد در صورتی که معلم قصد داشته باشد در قالب یک تکلیف نوشتن، که مستلزم نوشتن یک متن طولانی است، اطلاعاتی شناختی از خصیصه‌های نوشتن دانش‌آموزان کسب کند، ۱۷ نشان‌گر قابل استفاده است و اطلاعات شناختی مفیدی در اختیار معلم قرار خواهد گرفت. دو نشان‌گر «عدم تأثیرگذاری منفی اشتباهات دستوری و زبان‌شناختی بر درک مطلب» و «استفاده از واژگان پیچیده» برای سنجش نوشتن در پایه ششم ابتدایی دشوار بوده و توان تشخیص افراد مسلط از غیرمسلط را نداشتند. در صورتی که قرار باشد توانایی نوشتن دانش‌آموزان در قالب یک متن کوتاه مانند بازنویسی یک متن ادبی یا خلاصه‌نویسی یک متن دیگر، سنجیده شود، علاوه بر دو نشان‌گر مذکور نشان‌گرهای «مختصر و مفید بودن و عدم وجود حشو در نوشته»، «سازماندهی ایده‌ها و مطالب در متن»، «وجود جملات متعدد در متن»، «وجود جملات مرکب در متن»، «به کار بردن ضمایر به صورت مناسب» و «وجود لحن مناسب در کل نوشته» نیز غیرمفید خواهد بود؛ به عبارت دیگر نشان‌گرهای مذکور در متن‌های کوتاه امکان ظهور پیدا نمی‌کند. با وجود این، در قالب همان متن‌های کوتاه نیز می‌توان خصیصه‌های مرتبط با نوشتن را با دقت مناسبی سنجید؛ بنابراین، از سیاهه و آرسی نهایی این تحقیق می‌توان برای سنجش نوشتن در پایه ششم ابتدایی استفاده کرد.

این تحقیق نیز مانند هر تحقیق دیگر با محدودیت‌هایی مواجه بوده است. مهم‌ترین محدودیت این تحقیق این است که جامعه آماری این تحقیق را دانش‌آموزان شهر قم تشکیل داده‌اند. بنابراین، تعمیم نتایج این تحقیق به دانش‌آموزان سایر شهرها خصوصاً شهرهای دارای دانش‌آموزان دوزبانه، باید با احتیاط انجام شود. همچنین در مرحله نمره‌گذاری نوشته‌های دانش‌آموزان از دو نفر ارزیاب استفاده شد. هرچند قبل از شروع فرایند نمره‌گذاری با آموزش و تمرین، تلاش شد نمره‌گذاری بر اساس ملاک ثابت و مشخصی انجام شود، اما همچنان تأثیر سلیقه و گرایش‌های نمره‌گذاران را نمی‌توان نادیده گرفت. ناآشنایی معلمان ابتدایی با خصیصه‌های نوشتن و عدم علاقه برخی از ایشان به کسب اطلاعات لازم در این زمینه، انجام تحقیق را با محدودیت جدی مواجه کرد؛ و در نهایت باید گفت سنجش نوشتن با استفاده از روش مورد بررسی در این تحقیق، تنها به سنجش نوشتن به عنوان یک محصول توجه دارد و سنجش فرایند نوشتن را بررسی نمی‌کند. بنابراین، نمی‌توان از روش و نتایج این تحقیق برای سنجش فرایند نوشتن استفاده کرد.

بر این اساس پیشنهاد می‌شود علاوه بر استفاده از الگو و نتیجه این تحقیق به منظور برنامه‌ریزی برای سنجش‌های شناختی تشخیصی در حوزه نوشتن، برای استفاده از ظرفیت‌های سنجش شناختی تشخیصی در سایر حوزه‌های مرتبط با یادگیری نیز تحقیقات مستقلی انجام شود. فرایند اجرای بخش کیفی این تحقیق نشان داد بسیاری از معلمان و والدین با خصیصه‌های زبانی از جمله نوشتن آشنایی کافی ندارند. این درحالی است که طرح موضوع و توضیح در مورد اینکه نوشتن دارای ابعاد و خصیصه‌هایی است که افراد می‌توانند آن‌ها را فرا بگیرند و عملکرد خود را بهبود بخشند، مورد علاقه و استقبال زیاد معلمان و والدین قرار می‌گرفت. به صورت خاص معلمانی که در این تحقیق همکاری داشتند عنوان می‌کردند تصمیم دارند با این خصیصه‌ها بیشتر آشنا شوند و در آموزش‌های خود آن‌ها را به‌کار ببرند؛ بنابراین، پیشنهاد می‌شود آموزش این خصیصه‌ها در برنامه توسعه حرفه‌ای معلمان قرار گیرد. معلمان نیز در جلساتی که با اولیاء دانش‌آموزان دارند، خصیصه‌های مذکور را به والدین آموزش داده و از آن‌ها بخواهند در منزل خصیصه‌های مذکور را با دانش‌آموزان تمرین کنند.

پاسخ‌های برخی از دانش‌آموزان به سؤال‌های نوشتن نشان داد برخی از دانش‌آموزان نتوانسته‌اند متوجه شوند که در مورد چه موضوعی باید متن بنویسند؛ به عبارت دیگر در فرایند حل مسأله نتوانسته‌اند صورت مسأله را متوجه شوند؛ بنابراین، پیشنهاد می‌شود برای آموزش الگوهای حل مسأله در مقاطع مختلف تحصیلی و به صورت خاص در دوره ابتدایی برنامه‌ریزی‌های لازم صورت گیرد و این الگوها به دانش‌آموزان تدریس شود.

پیشنهاد می‌شود ضمن اضافه کردن تعداد محدودی سؤال به این آزمون، سوگیری‌های سؤالات بر اساس برخی از ویژگی‌های مهم مانند جنسیت، نوع مدرسه، لذت بردن از فعالیت آموزشی، لذت بردن از نوشتن و مطالعه، رشد مهارت‌های اجتماعی و ... بررسی شود.

با توجه به اهمیت نوشتن در همه مقاطع تحصیلی، پیشنهاد می‌شود تحقیق‌های مشابهی در سایر مقاطع تحصیلی انجام شود و ضمن ساخت آزمون مناسب، نقاط ضعف و قوت دانش‌آموزان و دانشجویان در سایر مقاطع تحصیلی نیز مشخص شود.





## منابع

- شریفی، حسن پاشا و دانش پژوه، زهرا (۱۳۸۳). سنجش ملی پیشرفت تحصیلی زبان فارسی. *تعلیم و تربیت*. (۷۹): ۷-۵۰.
- دانشگر، مریم (۱۳۹۶). مهارت‌های دانش‌آموختگان نظام متوسطه با محوریت درس زبان و ادبیات فارسی (مطالعه موردی: شهر تهران). *جستارهای زبانی*. ۸(۱): ۲۳۱-۲۵۶.
- دانشگر، مریم، عابدینی، سهیلا و محمدی ندا (۱۳۹۴). تحلیل آماری نمره‌های زبان و ادبیات فارسی در کنکور سراسری. *مطالعات ملی*. ۱۶(۲): ۷۷-۹۴.
- دفتر برنامه‌ریزی و تألیف کتب درسی (۱۳۸۶). *راهنمای برنامه درسی زبان و ادبیات فارسی*. تهران: سازمان پژوهش و برنامه‌ریزی آموزشی.
- موسوی، سیدایمان، کیانی، غلامرضا، اکبری، رامین و غفارثمر، رضا (۱۳۹۵). سیر تکاملی آگاهی از ژانر رشته تخصصی با آموزش مهارت‌های نگارش دانشگاهی: یک مطالعه موردی. *جستارهای زبانی*. ۷(۳): ۱۷۱-۱۹۶.
- Alderson, J. C. (2005). *Diagnosing foreign language proficiency: The interface between learning and assessment*. London: Continuum.
- Banerjee, J. and Wall, D. (2006). Assessing and reporting performances on pre-sessional EAP courses: Developing a final assessment checklist and investigating its validity. *Journal of English for Academic Purposes*, 5(1): 50-69.
- Cui, Y., Gierl, M. J. and Chang, H. H. (2012). Estimating classification consistency and accuracy for cognitive diagnostic assessment. *Journal of Educational Measurement*, 49(1): 19-38.
- Daneshgar, M. (2017). Literary and Lingual Skills Evaluation (Case Study: Tehran Secondary School Students). *Language Related Research*. 8(1): 231-256 (Text in Persian).
- Daneshgar, M., Abedini, S. and Mohammadi, N. (2015). The Statistical analysis of persian language & literature in the university entrance exam. *National Studies*. 16 (2): 77-94 (Text in Persian).
- Engelhard J.r. G. (2013). *Invariant measurement: Using Rasch models in the social, behavioral, and health sciences*. Routledge.
- Gorin, J. S. (2007). Test construction and diagnostic testing. In J. P. Leighton & M. J. Gierl (Eds.), *Cognitive diagnostic assessment for education: Theory and applications* (pp. 173-201). New York: Cambridge University Press.
- Jarvis, S., Grant, L., Bikowski, D. and Ferris, D. (2003). Exploring multiple profiles of highly rated learner compositions. *Journal of Second Language Writing*, 12(4): 377-403.
- Kim, A. Y. (2015). Exploring ways to provide diagnostic feedback with an ESL placement test: Cognitive diagnostic assessment of L2 reading ability. *Language*



- Testing*, 32: 227–258.
- Knoch, U. (2011). Rating scales for diagnostic assessment of writing: What should they look like and where should the criteria come from?. *Assessing Writing*, 16(2): 81-96.
- Kulikowich, J. M., Mason, L. H. and Brown, S. W. (2008). Evaluating fifth-and sixth-grade students' expository writing: task development, scoring, and psychometric issues. *Reading and Writing*, 21(1-2): 153-175.
- Kunina-Habenicht, O., Rupp, A. A. and Wilhelm, O. (2012). The Impact of Model Misspecification on Parameter Estimation and Item-Fit Assessment in Log-Linear Diagnostic Classification Models. *Journal of Educational Measurement*, 49(1): 59-81.
- Lei, P. W. and Li, H. (2016). Performance of fit indices in choosing correct cognitive diagnostic models and Q-matrices. *Applied Psychological Measurement*, 40(6): 405-417.
- Leighton, J. P. and Gierl, M. J. (2007). Defining and evaluating models of cognition used in educational measurement to make inferences about examinees' thinking processes. *Educational Measurement: Issues and Practices*, 26(2): 3–16.
- Li, H., Hunter, C. V. and Lei, P. W. (2016). The selection of cognitive diagnostic models for a reading comprehension test. *Language Testing*, 33(3): 391-409.
- Lin, J. and Wu, F. (2003). *Differential performance by gender in foreign language testing*. The Centre for Research in Applied Measurement and Evaluation, The University of Alberta.
- Mousavi, S. I., Kiany, G., Akbari, R. and Ghafarsamar R. (2016) The process of developing disciplinary genre awareness through training academic writing skills: A case study. *Language Related Research*. 7 (3): 171-196 (Text in Persian).
- Office of Planning and writing textbooks (2007). *Persian Language and Literature Secondary school curriculum guide*. Tehran: Organization for Educational Research and Planing (Text in Persian).
- Shareefee, H. P. and Daneshpajoo, Z. (2004). National assessment of achievement in Farsi language. *Quarterly Journal of Education*, 20 (3) (Text in Persian).
- Struthers, L., Lapadat, J. C. and MacMillan, P. D. (2013). Assessing cohesion in children's writing: Development of a checklist. *Assessing Writing*, 18(3): 187-201.
- Xie, Q. (2017). Diagnosing university students' academic writing in English: Is cognitive diagnostic modelling the way forward?. *Educational Psychology*, 37(1): 26-47.

---

### Using cognitive diagnostic assessment to assess writing in sixth grade elementary school students

---

Hassan Taghian \*<sup>1</sup>, Ebrahim Khodaei<sup>2</sup>, Abbas Bazargan<sup>3</sup>, Ali  
Moghadamzadeh<sup>4</sup> and Massoud Kabiri<sup>5</sup>

#### Abstract

Among the four language skills, writing is one of the most important communication skills that is planned for its education in all educational systems from the elementary school. Increasing advances in the field of assessment of education have provided the basis for obtaining the diagnostic information needed to identify students' strengths and weaknesses. This study intends to explore the possibility of using cognitive diagnostic assessment to assess the writing attributes of elementary school students. For this purpose, the check list, which was proposed by Kim (2015) for evaluating academic writing in English and the initial Q-matrix, was reviewed. 816 sixth grade elementary school students, answered three writing assignments. (Estimation of the model parameters were performed using Reduced Reparametrized Unified Model (R-RUM).) The relative and absolute indices indicate that the model is sufficiently fit. The result showed that 17 indicator can be used to assess the writing attitudes with appropriate accuracy in assignments requiring students to write a long text. But for assignments that require writing a short text, the 9 indicator can be used to assess the writing attitudes with appropriate accuracy.

**Keywords:** *writing, Cognitive Diagnostic Assessment, sixth Grade Elementary School, Writing Assessment, Reduced Reparametrized Unified Model*

- 
1. Corresponding Author: PhD Student, Education Assessment, University of Tehran, Tehran, Iran. taghian@gmail.com
  2. Associate Professor, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran. khodaie@ut.ac.ir
  3. Professor, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran. abazarga@hotmail.com
  4. Assistant Professor, Faculty of Psychology and Educational Sciences, University of Tehran, Tehran, Iran. amoghadamzadeh@ut.ac.ir
  5. Assistant Professor, Research Institute of the Ministry of Education, Tehran, Iran. maskabiri@yahoo.com
- DOI: 10.22051 / jontoe.2019.23583.2462  
[https://jontoe.alzahra.ac.ir/article\\_4367.html](https://jontoe.alzahra.ac.ir/article_4367.html)