

مقاله پژوهشی

کاربرد مدل تشخیصی شناختی در تعیین مهارت‌های شناختی زیربنایی عملکرد
در آزمون ماتریس‌های پیشرونده ریون

Application of Cognitive Diagnostic Model in Determining the Underlying
Cognitive Skills of Performance in Ryon Progressive Matrix Test

سعیده حسینی^۱، جلیل یونسی^{۲*}، اصغر مینائی^۳، ریحانه رحیمی^۴

پذیرش مقاله: ۱۳۹۹/۱۱/۰۶

دریافت مقاله: ۱۳۹۹/۰۲/۲۲

چکیده

هدف: هدف پژوهش حاضر تعیین مهارت‌های شناختی زیربنایی لازم برای پاسخگویی به سؤال‌های آزمون هوش ریون و شناسایی نقاط قوت و ضعف دانش‌آموزان پایه نهم در این مهارت با استفاده از مدل‌های سنجش تشخیصی شناختی بود. **روش:** پژوهش حاضر به لحاظ نوع داده‌های جمع‌آوری‌شده، جزء پژوهش‌های توصیفی و به لحاظ هدف جزء پژوهش‌های کاربردی محسوب می‌شود. جامعه آماری این پژوهش کلیه دانش‌آموزان پایه نهم مدارس دولتی شهر تهران در سال تحصیلی ۹۵-۱۳۹۴ بود که از میان مناطق نوزده‌گانه آموزش و پرورش ۷۰۰ نفر دانش‌آموز دختر و پسر از طریق روش نمونه‌گیری تصادفی خوشه‌ای انتخاب شدند. نخست با کمک نظر متخصصین حوزه هوش یک ماتریس Q که شامل روابط بین ۸ مهارت زیربنایی و ۶۰ سؤال آزمون ریون بود تشکیل گردید و با استفاده از مدل غیرجبرانی DINA با کمک بسته CDM در محیط نرم‌افزار R مورد تجزیه و تحلیل قرار گرفت. سپس به منظور تعیین احتمال تسلط آزمودنی -ها در هر یک از مهارت‌ها با تعیین نقطه برش ۰/۷ برای حد تسلط، با استفاده از روش پسینی مورد انتظار در چهارچوب مدل DINA، افراد مورد بررسی در هر یک از مهارت‌ها به دو گروه مسلط و غیر مسلط تقسیم شدند. **یافته‌ها:** نتایج سنجش تشخیصی شناختی نشان داد که بیشترین احتمال تسلط مربوط به مهارت‌های ثبات و پیشرفت بود و در مجموع وضعیت آزمودنی‌ها در همه مهارت‌ها به جز دو مهارت تسلط اندیشه و استدلال کمی مطلوب ارزیابی شد. **نتیجه‌گیری:** با توجه به وضعیت نامطلوب آزمودنی‌ها در دو مهارت استدلال کمی و تسلط اندیشه نسبت به سایر مهارت‌ها، می‌توان نتیجه گرفت که دانش‌آموزان در تولید ایده سریع و سیالی اندیشه و کشف روابط ریاضی ضعیف هستند؛ بنابراین پیشنهاد می‌شود در مدارس به پرورش خلاقیت دانش‌آموزان توجه بیشتری شده و تکالیفی که نیاز به ایده‌پردازی دارد به عهده آن‌ها گذاشته شود.

کلید واژه‌ها: پسینی مورد انتظار، سنجش تشخیصی شناختی، ماتریس Q، مدل DINA.

۱. کارشناس ارشد سنجش و اندازه‌گیری

۲. دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

۳. دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

۴. دانشجوی دکتری سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، تهران، ایران

۱. مقدمه

به لحاظ آموزشی، دلیل اهمیت آگاهی از درک مفهوم هوش به‌عنوان یک توانایی کلی ذهن یا مجموعه‌ای از توانایی‌های مختلف این است که معلمان می‌توانند با بررسی کلی توانایی‌های ذهنی شاگردان خود، با هریک از آن‌ها متناسب با جایگاه هوش کلی‌شان برخورد کرده و با همه شاگردان خود براساس یک برداشت کلی و کلیشه‌ای رفتار نکنند. به‌عبارت‌دیگر با بذل دقت بیشتری می‌توانند از توانمندی‌های ذهنی هر فرد درک دقیق‌تری داشته و با تأکید بیشتر بر توانایی‌های قوی‌تر او به شکوفایی استعدادهایش کمک کنند (امیرتیموری، ۱۳۸۲). از همین روی مفهوم هوش در سنجش همواره حائز اهمیت بوده، به‌گونه‌ای که به نظر استرنبرگ^۱ هوش چهارچوب بنیادی سازمان دهنده همه انواع شناخت انسانی و مفهوم آن دربرگیرنده وحدت کارکردی بنیادی شناخت نیز هست (استرنبرگ؛ ترجمه خرازی و حجازی، ۱۳۸۷). باوجود اهمیت جایگاه هوش، همواره شناخت و اندازه‌گیری آن با محدودیت‌ها، ابهامات و انتقادهایی روبرو بوده است.

هوش یک مفهوم انتزاعی و غیرقابل لمس بوده که از طریق نمودهای رفتاری و پاسخ‌های عینی افراد قابل استنباط است. در حقیقت برای شناخت هوش باید آثار آن را شناخت، اما نبود توافق بین نظریه‌پردازان در این مورد که آثار هوش چه مواردی هستند مناقشه‌ها و اختلاف‌نظرهای فراوانی در زمینه تعریف، شناخت و اندازه‌گیری آن ایجاد کرده است و در نهایت منجر به پیدایش رویکردها و نظریات مختلف و گاه متضادی در این حوزه شده است. مارنات (ترجمه شریفی و نیکخو، ۱۳۹۱) با مطالعه و بررسی نظریه‌های این حوزه ۴ رویکرد کلی شامل رویکرد روان‌سنجی، رویکرد عصبی-زیستی، نظریه‌های رشدی و پردازش اطلاعات را پیشنهاد کرده است.

بحث چالشی دیگری که در زمینه هوش وجود دارد، بحث بر سر غیراخلاقی بودن استفاده از آزمون‌های هوش و استناد به یک نمره به‌عنوان نمره هوشیهر است. این انتقادهای دو جهت قابل‌بررسی است. اولین انتقاد وارد شده در این زمینه که توسط امبرتسون^۲ (۱۹۸۶) مطرح شده است، بی‌توجهی آزمون‌های هوش به فرایند موجود در مسئله‌گشایی و استفاده از یک عدد منفرد به نام هوشیهر^۳ برای نشان دادن نمره فرد در آزمون است. انتقاد دیگری که مطرح می‌شود سوگیری آزمون‌های موجود نسبت به زمینه‌های فرهنگی خاص است. در حقیقت آزمون‌ها تحت تأثیر فرهنگ سازنده آن قرار گرفته و منجر به عملکرد بهتر هم نژادهای سازنده آزمون می‌شود. رویکرد پردازش اطلاعات با تأکید بر روان‌شناسی شناختی تا حدی توانسته است به این محدودیت‌ها پاسخگو باشد. در این رویکرد به‌جای محتوا بر کنش‌های فرایند ذهنی تمرکز شده است. یکی از نظریه‌های مشهور این رویکرد نظریه سه‌بخشی استرنبرگ (۱۹۸۵) است. در این نظریه بر تجربه‌های درونی شخص و بافت اجتماعی تجربه تمرکز و هوش به‌عنوان یک پدیده اجتماعی-فرهنگی معرفی شده است

1. Sternberg
2. Embretson
3. Intelligence Coefficient

(مارنات؛ ترجمه شریفی و نیکخو، ۱۳۹۱). علاوه بر رویکرد پردازش اطلاعات، رویکرد سنجش تشخیصی شناختی نیز فرایند رسیدن به پاسخ مسئله را مدنظر قرار داده است.

رویکرد سنجش تشخیصی شناختی برخلاف روش‌های مرسوم اندازه‌گیری آموزشی، به دلیل تلفیق نظریه‌های شناختی با نظریه‌های یادگیری و ارائه نیم‌رخ‌های صلاحیت‌های شناختی دانش‌آموزان، جایگاه ویژه‌ای پیدا کرده است. این مدل‌ها به‌جای یک نمره کلی با بررسی فرایندهای شناختی موردنیاز برای تکمیل موفقیت‌آمیز تکالیف، اطلاعات خاص و دقیقی را درباره مهارت یا صفتی که دانش‌آموزان برای پاسخ دادن به سؤال نیاز دارند فراهم می‌آورند (مک‌گلوهن^۱، ۲۰۰۴). در حقیقت این رویکرد قصد دارد تا با فراهم آوردن اطلاعات موردنیاز معلمان برای اصلاح آموزش و یادگیری در کلاس درس، سودمندی و نقش سنجش را در یادگیری ارتقا بخشد. معلمان می‌توانند از اطلاعات تشخیصی تکوینی برای بازسازی و طراحی مجدد رویکردهای آموزشی، ارزشیابی منابع آموزشی و جبران ضعف‌های دانش‌آموزان استفاده کنند. این رویکرد می‌تواند با تشویق دانش‌آموزان به استفاده از سنجش به‌عنوان ابزار یادگیری سهم داشته و نقش آن‌ها را در یادگیری افزایش دهد (جانگ^۲، ۲۰۰۸). علاوه بر این اطلاعات به‌دست‌آمده از آزمون‌های تشخیصی شناختی به‌طور معناداری می‌تواند به‌وسیله ذی‌نفعان آموزشی تفسیر شود (راپ، تمپلین و هنسون^۳، ۲۰۱۰). نکته دیگری که لازم است به آن اشاره شود، این است که تهیه و تدوین ارزشیابی‌های آموزشی بر مبنای سنجش تشخیصی شناختی، مزایای زیادی دارد که شاید بتوان گفت مهم‌ترین آن‌ها، پیوند دادن نظریه‌های شناختی و یادگیری با آموزش است، زیرا نتایج این نوع سنجش، گزارشی از نیم‌رخ شناختی دانش‌آموزان در زمینه تکلیف‌های درسی فراهم می‌آورد تا بتوان با طراحی‌های مناسب، نقاط ضعف شناسایی شده دانش‌آموزان را برطرف نمود (گیلر، وانگ و ژو^۴، ۲۰۰۸). در واقع از نتایج سنجش تشخیصی شناختی می‌توان در جهت اصلاح و متناسب کردن محتوا و روش‌های آموزشی با نیازهای دانش‌آموزان و برطرف ساختن ضعف و کمبودهای آنان استفاده کرد (مینایی، ۱۳۹۱).

اهمیت مدل‌های تشخیصی شناختی در مقایسه با نظریه کلاسیک آزمون^۵ (CTT) و نظریه سؤال - پاسخ^۶ (IRT) از آنجا ناشی می‌شود که در این مدل‌ها پروفایلی برای دانش‌آموزان ارائه می‌شود که در آن مجموعه‌ای از مهارت‌ها یا صفت‌های دو ارزشی که فرد در آن به تسلط رسیده یا نرسیده است نمایش داده می‌شود. پروفایل‌های صفات گسسته، اطلاعاتی در مورد نیازهای گروهی از دانش‌آموزان در اختیار معلم قرار می‌دهد (برخلاف مدل‌های نظریه سؤال - پاسخ چندبعدی که پروفایلی از نمرات ارائه نمی‌دهد) (هنسن و داگلاس، ۲۰۰۵). انتظار می‌رود پروفایل‌های تشخیصی تولیدشده از

1. McGlohen
2. Jang
3. Rupp, Templin & Henson
4. Gierl, Wang & Zhou
5. Classical Test Theory
6. Item Response Theory

سنجش تشخیصی شناختی در ارائه اطلاعات تشخیصی آموزشی مفید به‌عنوان بازخورد برای معلمان و دانش‌آموزان و کمک بیشتر به افزایش آگاهی فراشناختی دانش‌آموزان از پیشرفت‌های یادگیری خودشان به‌کار گرفته شود (باتلر و وین^۱، ۲۰۱۵؛ به نقل از جانگ، دان لوب، پارک و واندربوم^۲، ۲۰۱۵). مدل‌های تشخیصی شناختی، خانواده‌ای از مدل‌های روانسنجی هستند که به‌جای نمره، یا احتمالاً علاوه بر نمرات کلی آزمون، پروفایل‌هایی برای کمک به مربیان آموزشی در ارزیابی تسلط یا عدم تسلط دانش‌آموزان در مهارت‌ها و یا ویژگی‌های موردنیاز برای حل مسائل در یک آزمون ارائه می‌کنند (هو و دی‌لاتوره^۳، ۲۰۱۴). در واقع مدل‌های تشخیصی شناختی تلاش می‌کنند تا مهارت‌های مکنون یا ویژگی‌هایی را که آزمودنی باید برای درستی پاسخ‌های آزمون دارا باشد کشف کنند (دی کارلو^۴، ۲۰۱۰).

با عنایت به مطالب ذکرشده هدف از پژوهش پیش‌رو شناسایی مهارت‌های زیربنایی آزمون ماتریس‌های پیش‌رونده ریون و بررسی میزان تسلط و عملکرد دانش‌آموزان در هریک از این مهارت‌های زیربنایی است. با توجه به اینکه آزمون هوش در طبقه آزمون‌های شناختی قرار گرفته و هر فرد برای یافتن پاسخ صحیح نیازمند استفاده از راهبردهای شناختی خاصی هست از رویکرد سنجش تشخیصی شناختی برای نیل به این هدف استفاده شده است. نتایج حاصل از سنجش تشخیصی شناختی به‌منظور فراهم آوردن بازخورد، یک پروفایلی از نمرات با اطلاعات خاصی در مورد نقاط قوت و ضعف شناختی دانش‌آموزان ارائه می‌دهد که این بازخورد تشخیصی شناختی، هدایت معلمان، والدین و دانش‌آموزان را در فرایندهای آموزشی و یادگیری امکان‌پذیر می‌سازد؛ بنابراین سؤال موردبررسی این است که مهارت‌ها لازم برای رسیدن به موفقیت در آزمون ماتریس‌های پیش‌رونده ریون کدام است؟ و دانش‌آموزان در کدام یک از این مهارت‌ها تسلط دارند؟

۲. روش پژوهش

پژوهش حاضر به لحاظ نوع داده‌های جمع‌آوری‌شده، جزء پژوهش‌های توصیفی است که در چارچوب مدل‌سازی تشخیصی شناختی و با رویکرد بازبرزش (ریتروفیت) انجام شده است. در رویکرد بازبرزش از داده‌های آزمون‌های غیرتشخیصی با هدف تشخیص استفاده می‌شود؛ در حقیقت هدف اولیه طراحی آزمون مورد استفاده سنجش تشخیصی شناختی نبوده است. همچنین به‌دلیل تعیین مهارت‌های شناختی پایه و تعیین میزان تسلط دانش‌آموز در هریک از این مهارت‌ها و مشخص نمودن نقاط قوت و ضعف آن‌ها به لحاظ هدف جزء پژوهش‌های کاربردی محسوب می‌شود.

-
1. Butler & Winn
 2. Jang, Dunlop, Park, & van der Boom
 3. Huo & de la Torre
 4. DeCarlo

۱-۲. جامعه آماری و نمونه

جامعه آماری این پژوهش، کلیه دانش‌آموزان پایه نهم مدارس دولتی شهر تهران در سال تحصیلی ۱۳۹۴-۹۵، براساس آمار آموزش و پرورش در سال تحصیلی ۹۵-۱۳۹۴، به‌طورکلی ۷۴۷۳۱ دانش‌آموز در پایه نهم مشغول به تحصیل بودند که از این تعداد ۳۹۰۱۰ نفر (۵۲ درصد) دختر و ۳۵۷۲۱ نفر (۴۷ درصد) پسر هستند. اطلاعات مربوط به دانش‌آموزان پایه نهم شهر تهران در جدول ۱ آمده است.

جدول ۱: تعداد کل مدارس، کلاس و دانش‌آموزان پایه نهم شهر تهران در سال تحصیلی ۹۵-۱۳۹۴

تعداد کل	جامعه آماری
۶۱۴	مدارس (متوسطه اول)
۲۴۳۲	کلاس (پایه نهم)
۳۹۰۱۰	دانش‌آموزان دختر (پایه نهم)
۳۵۷۲۱	دانش‌آموزان پسر (پایه نهم)
۷۴۷۳۱	کل دانش‌آموزان (پایه نهم)

برای انتخاب نمونه از روش نمونه‌گیری تصادفی خوشه‌ای استفاده شد. به این صورت که ابتدا به سازمان آموزش و پرورش کل تهران مراجعه و پس از اخذ تأییدیه اجرای پژوهش توسط این سازمان، شهر تهران به پنج ناحیه شمال، جنوب، مرکز، شرق و غرب تقسیم و از هر ناحیه یک منطقه به‌صورت تصادفی انتخاب شد. سپس به سازمان آموزش و پرورش هر یک از مناطق مراجعه و مجوز اجرا و معرفی‌نامه جهت ورود به مدارس دولتی در مقطع متوسطه اول این مناطق اخذ گردید. از میان مدارس هر منطقه یک مدرسه دخترانه و یک مدرسه پسرانه به شیوه تصادفی انتخاب و از بین دانش‌آموزان پایه نهم هر مدرسه حدود ۷۰ نفر به‌عنوان نمونه پژوهش انتخاب شدند. از آنجایی که مدل‌های تشخیصی شناختی مستلزم حجم نمونه بزرگی است تعداد ۷۰۰ دانش‌آموز به‌عنوان نمونه انتخاب و مورد آزمون قرار گرفتند. از این ۷۰۰ نفر ۳۵۰ نفر دختر و ۳۵۰ نفر پسر بودند.

۲-۲. ابزار اندازه‌گیری

ابزار موردبررسی در این مطالعه، آزمون ماتریس‌های پیشرونده ریون فرم بزرگسالان بود. آزمون ماتریس‌های پیشرونده ریون یکی از آزمون‌های هوش غیرکلامی است که توسط ریون در انگلستان برای اندازه‌گیری عامل عمومی (۷) اسپیرمن ساخته شده است. فرم اول آزمون ماتریس‌های پیشرونده ریون در سال ۱۹۳۷ و فرم دوم آن در سال ۱۹۴۷ توسط پن رز و ریون تهیه شده است. بعد از تهیه این آزمون پژوهش‌های زیادی درباره آن به‌عمل آمده و بارها مورد تجدیدنظر قرار گرفته است. فرم‌های تجدیدنظر شده آزمون ریون، برای اندازه‌گیری هوش افراد در همه‌ی سطوح توانایی (از کودکان ۵ ساله تا بزرگسالان سرآمد) به‌کار می‌رود. در این پژوهش فرم ۱۹۳۷ آزمون ریون که دارای ۶۰ سؤال تصویری سیاه‌وسفید است و به ۵ سری ۱۲ تایی با درجه دشواری فزاینده تقسیم می

شود، اجرا گردید. زمان اختصاص یافته برای پاسخگویی به سؤال‌های این آزمون که ۶ یا ۸ گزینه‌ای هستند ۴۵ دقیقه است. محتوای سؤال‌های آزمون ریون، مستلزم فرایند استدلال ذهنی برای کشف اصول و منطق حاکم بر روابط اجزای الگوهای ماتریس‌ها است.

۳. یافته‌های پژوهش

در هر پژوهش تشخیصی شناختی نیاز به یک مدل شناختی و یک مدل آماری است. مدل شناختی در واقع همان ماتریس Q است که رابطه بین هر مهارت و سؤال‌های آزمون را مشخص می‌کند. تدوین ماتریس Q مستلزم انجام یک پژوهش کیفی با هدف تعیین مهارت‌های زیربنایی موردنیاز برای پاسخگویی به سؤال‌های آزمون است. در پژوهش حاضر به منظور تعیین مهارت‌ها ابتدا پیشینه موجود در زمینه نظریه‌های مرتبط با هوش مطالعه و تعدادی مهارت شناختی استخراج گردید. در نهایت با تأیید چهار متخصص حیطة هوش، ۵ مهارت تعیین شده توسط تمپلین به علاوه ۴ مهارت دیگر به عنوان مهارت‌های زیربنایی لازم برای موفقیت در آزمون ریون منظور گردید. ۹ مهارت تعیین شده همراه با تعریف هر یک از آن‌ها در جدول ۲ ارائه شده است.

جدول ۲: مهارت‌های پایه و تعریف هر یک از آن‌ها

ردیف	مهارت	تعریف
۱	ثبات ^۱	ثابت ماندن شکل تصویر در سطر یا ستون
۲	پیشرفت ^۲	تغییر ویژگی‌ها برحسب درجه (اندازه) در سطر یا ستون
۳	افزایش یا کاهش تصویر ^۳	ویژگی‌های اول و دوم عناصر برای ایجاد عنصر سوم اضافه یا کم می‌شوند.
۴	توزیع ۳ تایی ^۴	سه عنصر متفاوت به طور مساوی در سطرها یا ستون‌ها توزیع شده‌اند.
۵	توزیع ۲ تایی ^۵	دو عنصر یکسان در سطر یا ستون که با عنصر سوم هیچ وجه اشتراکی ندارند.
۶	استقرا ^۶	توانایی مشاهده پدیده‌ها و کشف اصول یا قوانین اساسی که رفتار را تعیین می‌کنند.
۷	استدلال متوالی عمومی ^۷	توانایی استدلال منطقی با استفاده از کان‌های شناخته شده و اصول است. این توانایی نیز به عنوان استدلال قیاسی شناخته شده و قانون برنامه است. در حالی که القای توانایی استفاده از واقعیت‌های شناخته شده به کشف اصول جدید منجر می‌شود، استدلال‌های ترتیبی کلی، توانایی استفاده از اصول شناخته شده در یک یا بیشتر مراحل منطقی برای کشف حقایق

1. Identity
2. Progression
3. Figure addition / subtraction
4. Distribution of three
5. Distribution of two
6. Induction
7. General sequential reasoning

ردیف	مهارت	تعریف
۸	استدلال کمی ^۱	جدید و یا حل مشکلات است. توانایی استدلال، یا استقرا و یا کسر، با اعداد، روابط ریاضی و عملیات است. این آزمون اندازه‌گیری استدلال کمی به دانش پیشرفته از ریاضیات نیازی نیست. محاسبات در این آزمون به‌طور معمول بسیار ساده است.
۹	تسلط (روانی) اندیشه ^۲	توانایی سریع تولید یک سری از ایده‌ها، کلمات و یا عبارات مربوط به یک شرایط خاص یا شیء است. در اینجا کمیت، نه کیفیت یا اصالت پاسخ، تأکید شده است.

۹ مهارت شناسایی شده همراه با سؤال‌های آزمون در اختیار چهار متخصص باسابقه تدریس حداقل ۵ سال قرار گرفت و از آن‌ها درخواست شد که رابطه هر سؤال با هر مهارت را مشخص کنند؛ به این صورت که تعیین شود برای پاسخ به هریک از این سؤال‌ها وجود کدام‌یک از این ۹ مهارت ضروری است. بدین ترتیب اطلاعات چهار ماتریس Q تنظیم شده توسط متخصصین با یکدیگر ترکیب و ماتریس Q اولیه تهیه شد. بررسی‌های اولیه ماتریس Q نشان داد که از ۹ مهارت موردنظر، مهارت توزیع دوتایی به کمتر از سه سؤال مربوط می‌شود به همین خاطر این مهارت از مجموعه مهارت‌ها حذف و ماتریس Q با ۸ مهارت تشکیل گردید. در پایان به‌منظور بررسی و رفع ابهامات و تأیید نهایی، این ماتریس در اختیار هر چهار متخصص قرار گرفت و با تأیید ایشان ماتریس Q نهایی (جدول ۳) تشکیل شد.

جدول ۳: رابطه سؤال‌ها و مهارت‌ها

سؤال	ثبات	پیشرفت	افزایش یا کاهش	توزیع ۳ تایی	استقرا	استدلال متوالی	استدلال کمی	تسلط (روانی) اندیشه
۱	۱	۰	۰	۰	۰	۰	۰	۰
۲	۱	۰	۰	۰	۰	۰	۰	۰
۳	۱	۰	۰	۰	۰	۰	۰	۰
۴	۱	۰	۰	۰	۰	۰	۰	۰
۵	۱	۰	۰	۰	۰	۰	۰	۰
۶	۱	۰	۰	۰	۰	۰	۰	۰
۷	۱	۰	۰	۰	۰	۰	۰	۰
۸	۱	۰	۱	۰	۰	۰	۰	۰
۹	۱	۱	۰	۰	۰	۰	۰	۰
۱۰	۱	۱	۰	۰	۰	۰	۰	۰
۱۱	۱	۰	۰	۰	۰	۱	۰	۰

1. Quantitative reasoning

2. Ideational fluency

سؤال	ثبات	پیشرفت	افزایش یا کاهش	توزیع ۳ تایی	استقرا	استدلال متوالی	استدلال کمی	تسلط (روانی) اندیشه
۱۲	۰	۱	۰	۰	۰	۰	۰	۰
۱۳	۱	۰	۰	۰	۰	۰	۰	۰
۱۴	۱	۰	۰	۰	۰	۰	۰	۰
۱۵	۱	۰	۰	۰	۰	۰	۰	۰
۱۶	۱	۰	۰	۰	۱	۰	۰	۰
۱۷	۱	۰	۰	۰	۱	۰	۰	۰
۱۸	۱	۰	۱	۰	۱	۰	۰	۰
۱۹	۱	۰	۰	۰	۱	۰	۰	۰
۲۰	۱	۰	۰	۰	۱	۰	۰	۰
۲۱	۱	۰	۰	۰	۱	۰	۰	۰
۲۲	۱	۰	۰	۰	۱	۰	۰	۰
۲۳	۱	۰	۱	۰	۱	۰	۰	۰
۲۴	۱	۰	۱	۰	۱	۰	۰	۰
۲۵	۱	۱	۱	۱	۱	۱	۰	۰
۲۶	۱	۱	۰	۱	۱	۱	۰	۰
۲۷	۱	۰	۱	۰	۱	۱	۱	۰
۲۸	۱	۰	۱	۰	۱	۱	۱	۰
۲۹	۱	۰	۱	۰	۱	۱	۰	۰
۳۰	۱	۰	۱	۰	۱	۱	۰	۰
۳۱	۰	۰	۰	۰	۱	۱	۰	۰
۳۲	۰	۰	۱	۰	۱	۱	۰	۰
۳۳	۰	۰	۰	۱	۱	۱	۰	۱
۳۴	۰	۰	۰	۰	۱	۱	۰	۱
۳۵	۰	۰	۱	۰	۱	۱	۰	۱
۳۶	۱	۰	۱	۰	۱	۱	۰	۱
۳۷	۱	۰	۰	۱	۰	۰	۰	۰
۳۸	۱	۰	۰	۱	۰	۰	۰	۰
۳۹	۱	۰	۱	۱	۰	۰	۰	۰
۴۰	۱	۰	۰	۱	۱	۱	۰	۰
۴۱	۰	۰	۰	۱	۱	۱	۰	۰
۴۲	۱	۰	۱	۱	۱	۱	۰	۰
۴۳	۱	۰	۰	۱	۱	۱	۰	۰
۴۴	۱	۰	۰	۱	۱	۱	۰	۰
۴۵	۱	۰	۰	۱	۱	۱	۰	۰
۴۶	۱	۰	۰	۱	۱	۱	۰	۰
۴۷	۰	۰	۰	۰	۱	۱	۰	۰
۴۸	۰	۰	۰	۱	۱	۰	۱	۰

سؤال	ثبات	پیشرفت	افزایش یا کاهش	توزیع ۳ تایی	استفرا	استدلال متوالی	استدلال کمی	تسلط (روانی) اندیشه
۴۹	۰	۰	۱	۰	۱	۰	۰	۱
۵۰	۱	۰	۰	۱	۱	۱	۰	۱
۵۱	۰	۰	۱	۱	۱	۱	۰	۰
۵۲	۱	۰	۱	۱	۱	۱	۰	۱
۵۳	۱	۰	۱	۱	۱	۱	۰	۱
۵۴	۰	۰	۱	۱	۱	۱	۰	۱
۵۵	۱	۰	۰	۰	۱	۱	۰	۱
۵۶	۱	۰	۱	۰	۱	۱	۰	۱
۵۷	۰	۰	۱	۱	۱	۱	۰	۱
۵۸	۱	۰	۱	۱	۱	۱	۰	۱
۵۹	۰	۰	۱	۰	۱	۱	۰	۱
۶۰	۰	۰	۱	۰	۱	۱	۱	۱

بعد از تدوین مدل شناختی در قالب ماتریس Q، با کمک نرم افزار R و با انتخاب مدل تشخیصی شناختی DINA که جزء مدل های غیرجبرانی است، تجزیه و تحلیل کمی داده ها انجام شد. در مدل های غیرجبرانی شرط پاسخ صحیح به سؤال تسلط یافتن بر تمام مهارت های مورد نیاز آن سؤال است. مدل DINA چهار پارامتر حدس، لغزش، تمیز و RMSEA را ارائه می دهد. نتایج حاصل از اجرای مدل DINA در جدول زیر گزارش شده است.

جدول ۴: ویژگی های روانسنجی سؤال های آزمون براساس مدل تشخیصی شناختی DINA

سؤال	حدس	لغزش	ضریب تمیز	RMSEA
۱	۱	۰/۰۰۰۱	۰/۰۰۰۱	۰/۰۰۰۱
۲	۱	۰/۰۰۰۱	۰/۰۰۰۱	۰/۰۰۰۱
۳	۰/۸۶۰	۰/۰۰۰۱	۰/۱۴۰	۰/۰۰۲
۴	۰/۸۶۰	۰/۰۰۱	۰/۱۴۰	۰/۰۰۲
۵	۰/۸۸۲	۰/۰۰۰۱	۰/۱۱۸	۰/۰۰۳
۶	۰/۹۲۷	۰/۰۰۱	۰/۰۷۱	۰/۰۰۲
۷	۰/۶۳۵	۰/۰۰۴	۰/۳۶۱	۰/۰۰۶
۸	۰/۶۷۰	۰/۰۲۱	۰/۳۰۹	۰/۰۳۱
۹	۰/۹۰۹	۰/۰۰۰۱	۰/۰۹۱	۰/۰۱۹
۱۰	۰/۷۰۹	۰/۰۰۰۱	۰/۲۹۱	۰/۰۴۰
۱۱	۰/۶۳۷	۰/۱۲۴	۰/۲۳۹	۰/۰۶۱
۱۲	۰/۱۸۰	۰/۲۴۰	۰/۵۸۰	۰/۰۵۷
۱۳	۰/۷۹۰	۰/۰۰۱	۰/۲۰۹	۰/۰۰۳
۱۴	۰/۷۹۰	۰/۰۰۰۱	۰/۲۱۰	۰/۰۰۳

سؤال	حدس	لغزش	ضریب تمیز	RMSEA
۱۵	۰/۸۲۳	۰/۰۰۵	۰/۱۷۲	۰/۰۱۸
۱۶	۰/۹۰۹	۰/۰۰۷	۰/۰۸۴	۰/۰۲۵
۱۷	۰/۸۵۶	۰/۰۱۴	۰/۱۳۱	۰/۰۴۷
۱۸	۰/۷۶۴	۰/۰۰۷	۰/۲۲۹	۰/۰۴۳
۱۹	۰/۶۵۱	۰/۰۸۴	۰/۲۶۴	۰/۰۴۸
۲۰	۰/۶۴۷	۰/۰۳۲	۰/۳۲۱	۰/۰۴۱
۲۱	۰/۶۰۷	۰/۰۰۵	۰/۳۸۸	۰/۰۳۶
۲۲	۰/۸۴۴	۰/۰۰۱	۰/۱۵۵	۰/۰۴۴
۲۳	۰/۷۱۰	۰/۰۰۶	۰/۲۸۴	۰/۰۵۹
۲۴	۰/۴۸۲	۰/۰۲۹	۰/۴۸۹	۰/۰۵۷
۲۵	۰/۹۷۵	۰/۰۰۰۱	۰/۰۲۶	۰/۰۲۴
۲۶	۰/۹۲۹	۰/۰۰۲	۰/۰۷۰	۰/۰۴۵
۲۷	۰/۹۷۴	۰/۰۰۲	۰/۰۲۴	۰/۰۳۲
۲۸	۰/۸۴۷	۰/۰۱۷	۰/۱۳۶	۰/۰۵۶
۲۹	۰/۸۵۰	۰/۰۰۷	۰/۱۴۳	۰/۰۳۷
۳۰	۰/۷۲۴	۰/۰۲۱	۰/۲۵۵	۰/۰۴۴
۳۱	۰/۸۲۴	۰/۰۰۸	۰/۱۶۸	۰/۰۳۲
۳۲	۰/۴۷۳	۰/۱۷۶	۰/۳۵۰	۰/۱۲۰
۳۳	۰/۸۱۲	۰/۰۷۳	۰/۱۱۵	۰/۰۷۵
۳۴	۰/۴۹۷	۰/۱۶۶	۰/۳۳۷	۰/۱۲۰
۳۵	۰/۵۵۳	۰/۱۶۱	۰/۲۸۶	۰/۰۷۵
۳۶	۰/۰۷۷	۰/۶۴۶	۰/۲۷۷	۰/۰۶۵
۳۷	۰/۹۲۹	۰/۰۰۴	۰/۰۶۷	۰/۰۳۹
۳۸	۰/۸۷۶	۰/۰۰۶	۰/۱۱۸	۰/۰۳۰
۳۹	۰/۸۹۹	۰/۰۰۳	۰/۰۹۸	۰/۰۴۷
۴۰	۰/۸۰۲	۰/۰۱۸	۰/۱۸۱	۰/۰۵۸
۴۱	۰/۹۳۹	۰/۰۱۲	۰/۰۴۹	۰/۰۲۴
۴۲	۰/۸۴۹	۰/۰۲۹	۰/۱۲۲	۰/۰۶۰
۴۳	۰/۶۰۲	۰/۰۵۳	۰/۳۴۵	۰/۰۵۷
۴۴	۰/۶۰۷	۰/۰۱۴	۰/۳۷۸	۰/۰۸۰
۴۵	۰/۳۷۰	۰/۰۹۷	۰/۵۳۳	۰/۰۶۴
۴۶	۰/۵۱۸	۰/۰۹۷	۰/۳۸۴	۰/۰۷۰
۴۷	۰/۲۴۹	۰/۶۵۲	۰/۰۹۹	۰/۱۴۰
۴۸	۰/۰۶۴	۰/۶۶۶	۰/۲۷۰	۰/۰۴۰
۴۹	۰/۶۷۰	۰/۰۲۱	۰/۳۰۹	۰/۰۶۴
۵۰	۰/۶۰۰	۰/۰۱۷	۰/۳۸۳	۰/۰۷۵
۵۱	۰/۴۷۸	۰/۱۴۰	۰/۳۸۱	۰/۱۵۱
۵۲	۰/۴۴۳	۰/۰۴۲	۰/۵۱۵	۰/۰۸۵
۵۳	۰/۴۳۱	۰/۰۴۰	۰/۵۲۸	۰/۰۵۸
۵۴	۰/۴۳۰	۰/۱۲۴	۰/۴۴۶	۰/۰۸۸
۵۵	۰/۲۳۴	۰/۳۲۵	۰/۴۴۱	۰/۰۷۳

سؤال	حدس	لغزش	ضریب تمیز	RMSEA
۵۶	۰/۲۱۸	۰/۲۴۱	۰/۵۴۰	۰/۰۵۲
۵۷	۰/۲۲۴	۰/۴۶۴	۰/۳۱۲	۰/۰۳۸
۵۸	۰/۰۷۴	۰/۵۰۶	۰/۴۲۰	۰/۰۳۴
۵۹	۰/۰۵۷	۰/۸۶۴	۰/۰۷۹	۰/۰۴۱
۶۰	۰/۰۷۸	۰/۶۲۶	۰/۲۹۶	۰/۰۳۹

با توجه به نتایج مندرج در جدول بالا سؤال‌های ۳۶، ۴۸، ۵۸، ۵۹ و ۶۰ با مقادیر ۰/۰۷۷، ۰/۰۶۴، ۰/۰۷۴، ۰/۰۵۷ و ۰/۰۷۸ کمترین ضریب حدس را دارند و بالاترین ضریب حدس^۱ مربوط به سؤال‌های ۲۷، ۲۵، ۴۱، ۲۶، ۶، ۹ و ۱۶ با مقادیر ۰/۹۷۹، ۰/۹۷۵، ۰/۹۳۹، ۰/۹۲۹، ۰/۹۲۷ و ۰/۹۰۹ است. همچنین تمامی سؤال‌ها به‌غیر از سؤال‌های ۵۹، ۴۸، ۴۷، ۳۶، ۶۰، ۵۸ و ۵۷ با مقادیر ۰/۸۶۴، ۰/۶۶۶، ۰/۶۵۲، ۰/۶۴۶، ۰/۶۲۶، ۰/۵۰۶ و ۰/۴۶۴ دارای ضریب لغزش پایینی هستند. نتایج جدول حاکی از آن است همه سؤال‌ها دارای شاخص تمیز مثبت هستند و سؤال‌های شماره ۵۲، ۵۳، ۴۵، ۵۶ و ۱۲ به ترتیب با مقادیر ۰/۵۱۵، ۰/۵۲۸، ۰/۵۳۳، ۰/۵۴۰ و ۰/۵۸۰ شاخص تمیز بالاتری نسبت به بقیه سؤال‌ها دارند و ۱۷ سؤال دارای شاخص تمیزی بین ۰/۳ تا ۰/۵ هستند و بقیه مقادیری پایین‌تر از ۰/۳ دارند.

شاخص RMSEA برای بررسی برازش سؤال‌ها با مدل به کار می‌رود. سؤال‌های با $RMSEA > 0/1$ نشان‌دهنده برازش ضعیف، سؤال‌های با $0/1 < RMSEA < 0/05$ نشان‌دهنده برازش متوسط و سؤال‌های با $RMSEA < 0/05$ نشان‌دهنده برازش خوب با مدل هستند (کانینا-هینت، راپ و ویلهلم^۲، ۲۰۰۹)؛ بر طبق این معیار سؤال‌های ۳۲، ۳۴، ۴۷ و ۵۱ برازش ضعیفی با مدل DINA نشان دادند. در جدول زیر شاخص‌های دقت و ثبات طبقه‌بندی‌ها گزارش شده است. دقت طبقه‌بندی نشان می‌دهد به چه اندازه تک‌تک آزمودنی‌ها در کلاس‌های مهارتی درست طبقه‌بندی شده‌اند؛ در واقع دقت به روایی طبقه‌بندی‌ها اشاره دارد. ثبات طبقه‌بندی‌ها نیز پایایی طبقه‌بندی‌ها را نشان می‌دهد؛ یعنی اگر یک آزمودنی دوباره در همین آزمون یا در یک آزمون موازی با آن شرکت کند آیا باز در همین طبقه فعلی قرار خواهد گرفت.

۱. سؤال ۱ و ۲ به دلیل سادگی بیش از حد توسط همه آزمودنی‌ها پاسخ داده شده است، بنابراین شاخص حدس این دو سؤال برابر با یک برآورد شده است؛ لذا در اینجا منظور سؤالات دیگر به جز سؤال ۱ و ۲ است.

2. Kunina-Habenicht, Rupp & Wilhelm

جدول ۵: دقت و ثبات مهارت‌های هشت‌گانه در نمونه مورد مطالعه

مهارت	دقت	ثبات
ثبات	۰/۸۹۶	۰/۸۴۵
پیشرفت	۰/۷۱۹	۰/۹۳۶
افزایش یا کاهش تصویر	۰/۹۳۲	۰/۹۸۸
توزیع سه‌تایی	۰/۹۲۲	۱
استقرا	۰/۸۶۳	۰/۹۴۶
استدلال متوالی عمومی	۰/۹۰۱	۰/۹۹۸
استدلال کمی	۰/۹۲۹	۱
تسلط (روانی) اندیشه	۰/۹۸۰	۱

با توجه به جدول ۵ همه مهارت‌ها از شاخص دقت بالایی برخوردار هستند (بالتر از ۰/۷). نتایج شاخص ثبات مرتبط به مهارت‌ها نشان می‌دهد که شاخص ثبات سه مهارت توزیع سه‌تایی، استدلال کمی و تسلط (روانی) اندیشه برابر ۱ است؛ زیرا احتمال تسلط به این ۳ مهارت بسیار بالا است. شاخص دقت کلی برابر با ۰/۸۲۴ گزارش شده است که در مقایسه با میزان به‌دست‌آمده از مقدار شبیه‌سازی (۰/۷۸۰) بیشتر است. همچنین شاخص ثبات کلی برابر با ۰/۷۷۷ و بیشتر از مقدار شبیه‌سازی شده (۰/۶۷۶) است؛ بنابراین می‌توان نتیجه گرفت که طبقه‌بندی‌های صورت گرفته در این پژوهش از دقت و ثبات بالایی برخوردارند.

به‌منظور تعیین وضعیت تسلط آزمودنی‌ها بر مهارت‌های موجود، ابتدا با روش EAP (پسینی مورد انتظار^۱) احتمال تسلط هر فرد بر هریک از مهارت‌ها محاسبه شد. سپس با توجه به مطالعه پیشینه و نظر متخصصین نقطه برش ۰/۷ برای حد تسلط تعیین گردید و بر این اساس افراد گروه نمونه به دو دسته مسلط و غیر مسلط تقسیم‌شده و فراوانی و درصد فراوانی افراد در هریک از این دو گروه برای هریک از مهارت‌ها محاسبه شد که نتایج آن در جدول ۶ قابل مشاهده است.

جدول ۶: فراوانی و درصد فراوانی افراد مسلط و غیر مسلط در هر مهارت

مهارت	احتمال حاشیه‌ای		مسلط		غیرمسلط	
	تسلط		فراوانی	درصد	فراوانی	درصد
ثبات	۰/۹۷۹		۶۸۸	۹۸/۳	۱۲	۱/۷
پیشرفت	۰/۹۲۹		۶۴۷	۹۲/۴	۵۳	۷/۶
افزایش یا کاهش تصویر	۰/۹۰۰		۶۱۴	۸۷/۷	۸۶	۱۲/۳
توزیع سه‌تایی	۰/۸۳۰		۵۵۴	۷۹/۱	۱۴۶	۲۰/۹
استقرا	۰/۹۲۱		۶۲۱	۸۸/۷	۷۹	۱۱/۳
استدلال متوالی عمومی	۰/۹۲۵		۶۳۲	۹۰/۳	۶۸	۹/۷
استدلال کمی	۰/۷۳۵		۴۴۶	۶۳/۷	۲۵۴	۳۶/۳
تسلط (روانی) اندیشه	۰/۶۲۹		۳۷۷	۵۳/۹	۳۲۳	۴۶/۳

1. Eepected Aposteriori

نتایج مربوط به احتمال حاشیه‌ای نشان می‌دهد که احتمال تسلط آزمودنی‌ها در مهارت‌های ثبات (۰/۹۷۹)، پیشرفت (۰/۹۲۹)، استدلال متوالی عمومی (۰/۹۲۵)، استقرا (۰/۹۲۱) و افزایش یا کاهش تصویر (۰/۹۰۰)، بیش از ۰/۹ است. در مهارت ثبات تنها ۱۲ نفر به تسلط نرسیده‌اند. ضعیف‌ترین عملکرد افراد در مهارت تسلط اندیشه است و تقریباً نیمی از افراد (۴۶/۳٪) در این مهارت به حد تسلط نرسیده‌اند. همچنین در مهارت استدلال کمی ۳۶/۳ درصد افراد به تسلط نرسیده‌اند؛ بنابراین در مجموع وضعیت آزمودنی‌ها در همه مهارت‌ها به‌غیراز دو مهارت تسلط اندیشه و استدلال کمی مطلوب بوده است.

بحث و نتیجه‌گیری

نیروی انسانی کارآمد و پاسخگو به نیازهای جامعه، یکی از شاخص‌های توسعه محسوب شده و پرورش این نیروی کارآمد در گرو یک نظام آموزشی پیشرفته و سازنده است. نظام آموزشی یکی از مهم‌ترین ارکان اجتماعی در اکثر جوامع و عهده‌دار رشد فرهنگی، اجتماعی، اقتصادی و همچنین شکوفایی استعدادهای بالقوه افراد است (آرمسترونک^۱، ۲۰۱۱). یکی از چالش‌برانگیزترین مسائل در هر نظام آموزشی، وجود تفاوت‌های فردی دانش‌آموزان است. به همین علت هوش^۲ به‌عنوان یکی از ویژگی‌های اساسی که موجب تفاوت‌های فردی بین انسان‌ها می‌شود، همواره مورد توجه بوده و اولین قدم جدی برای مطالعه و سنجش آن در آغاز قرن بیستم برداشته شده است. اما سنجشی که به معلم اجازه می‌دهد تا درک کند چگونه یک دانش‌آموز تکلیفی را انجام می‌دهد، سودمندتر از داشتن نمره‌ای است که صرفاً نمایشگر پاسخ‌های درست و نادرست دانش‌آموز باشد. معلم با اطلاع از اشتباهات مفهومی دانش‌آموز، بهتر می‌تواند برنامه‌ای آموزشی طراحی کند که علاوه بر محتوای مورد نظر، هم‌زمان به بررسی توانایی‌های شناختی دانش‌آموز توجه داشته باشد (اشمن و کانوی؛ ترجمه خرازی، ۱۳۸۴). در حال حاضر بیشتر از همیشه، تأکید بر اندازه‌گیری توانایی‌های دانش‌آموزان برای اطمینان از مناسب بودن گمارش نمره و کیفیت آموزش است (هنسن و داگلاس^۳، ۲۰۰۵). در سال‌های اخیر، نوع جدیدی از سنجش با عنوان سنجش تشخیصی شناختی به‌طور وسیعی در ارزیابی آموزشی به‌منظور ارائه پروفایل تشخیصی تسلط یا عدم تسلط برای آزمودنی‌ها معرفی شده است (چن، ۲۰۱۳).

از همین روی در پژوهش حاضر تلاش شد تا مهارت‌های شناختی زیربنایی آزمون ماتریس‌های پشرونده ریون شناسایی شود و با استفاده از مدل تشخیصی شناختی غیرجبرانی DINA میزان تسلط و عدم تسلط دانش‌آموزان پایه نهم در این مهارت‌ها بررسی گردد. با نظرسنجی از متخصصین

1. Armstrong
2. Intelligence
3. Henson & Douglas

۵ مهارت تعیین‌شده توسط تمپلین و ایوی^۱ (۲۰۰۶) که در آزمون ریون ۳۷ سؤالی شناسایی‌شده بود به‌اضافه ۴ مهارت دیگر، یعنی در مجموع ۹ مهارت به‌عنوان مهارت‌های پایه مشخص گردید. در مرحله بعد طرح کدگذاری و سؤال‌های آزمون در اختیار چهار متخصص مجرب در حیطه هوش قرار داده شد تا مشخص کنند که برای پاسخگویی به هر یک از سؤال‌های آزمون کدام‌یک از مهارت‌ها لازم است. در نهایت ماتریس Q با ۸ مهارت شامل ثبات، پیشرفت، افزایش یا کاهش تصویر، توزیع ۳ تایی، استقرا، استدلال متوالی عمومی، استدلال کمی و تسلط (روانی) اندیشه تشکیل شد. روسو^۲، تمپلین و هنسون (۲۰۰۷) توصیه می‌کند که مدل‌های غیرجبرانی برای آزمون‌های ریاضیات مناسب است؛ زیرا در این آزمون‌ها راه‌حل به یکسری از مراحل کوچک‌تر تقسیم می‌شود و اگر تمام مراحل موفقیت‌آمیز پشت سر گذاشته شود مسئله ریاضی موردنظر حل می‌شود. با تعمیم استدلال روسو به آزمون هوش در این پژوهش نیز از مدل DINA که مدلی غیرجبرانی است استفاده شد.

در برآورد ویژگی‌های سؤال‌های آزمون با استفاده از مدل غیر جبرانی DINA چهار پارامتر حدس، لغزش، تشخیص و RMSEA به دست آمد. نتایج نشان داد که سؤال‌های ۳۶، ۴۸، ۵۸، ۵۹ و ۶۰ کمترین ضریب حدس را دارند، این نتیجه با توجه به دشواری بالای سؤال‌های پایانی آزمون منطقی به‌نظر می‌رسد. همچنین همه سؤالات به‌غیر از سؤال‌های ۵۹، ۴۸، ۴۷، ۳۶، ۶۰، ۵۸ و ۵۷ دارای ضریب لغزش پایینی هستند. در واقع به دلیل دشواری بالای این سؤال‌ها، آزمون‌دهی‌های قوی نیز دچار خطا شده و از پاسخ دادن صحیح به این سؤال‌ها عاجز مانده‌اند. راب و همکاران (۲۰۱۰) معتقدند که هر چه مقادیر پارامترهای حدس و لغزش سؤالی کوچک‌تر باشد اطلاعات مفیدتری را ارائه می‌دهد. به‌عبارتی‌دیگر، مقادیر کوچک پارامترهای حدس و لغزش برازش مناسب بین داده‌ها، مدل DINA و ساختار سنجش شناختی را نشان می‌دهد (محسن‌پور، ۱۳۹۳). پارامتر IDI یا شاخص تمیز، توانایی سؤال در تفکیک افراد مسلط و غیرمسلط در مهارت و RMSEA برازش سؤال‌ها با مدل را نشان می‌دهد. نتایج پژوهش نشان داد که همه سؤال‌ها دارای شاخص تمیز مثبت هستند و ۵ سؤال شاخص تمیز بالاتری نسبت به بقیه سؤال‌ها دارند و ۱۷ سؤال دارای شاخص تمیزی بین ۰/۳ تا ۰/۵ هستند و بقیه مقادیری پایین‌تر از ۰/۳ دارند. مقادیر بالای شاخص IDI (مقادیر نزدیک به ۱) بیانگر آن است که سؤال توانسته است بین افراد مسلط و غیر مسلط در سؤال به‌خوبی تفکیک قائل شود. سؤال‌های شماره ۵۲، ۵۳، ۴۵، ۵۶ و ۱۲ دارای بالاترین ضریب تشخیص در این مجموعه سؤال بودند. سؤال‌های ۱ و ۲ به دلیل سادگی بیش‌ازحد دارای کمترین ضریب تشخیص هستند. یافته‌های پژوهش نشان داد که ۳۷ سؤال دارای مقادیر RMSEA کوچک‌تر از ۰/۰۵ هستند که نشانگر برازش بالای مدل با این سؤال‌ها است، همچنین مقادیر RMSEA ۲۱

1. Ivie
2. Roussos

سؤال دیگر در دامنه $0/05$ تا $0/1$ می‌باشد که نشان‌دهنده برآزش متوسط است. سؤالات ۴۷ و ۵۱ (با مقادیر بیشتر از $0/1$) کمترین برآزش را با مدل DINA نشان دادند. نتایج احتمال حاشیه‌ای (کناری) نشان داد که آزمودنی‌ها بیشترین احتمال تسلط را به ترتیب در مهارت‌های ثبات و پیشرفت دارند. چنین نتیجه‌ای قابل‌انتظار و منطقی است؛ زیرا همان‌طور که در ماتریس Q مشاهده می‌شود سؤال‌های ابتدایی که آسان‌تر بوده و تقریباً توسط اکثر آزمودنی‌ها پاسخ داده می‌شود، با مهارت ثبات رابطه دارند. از طرف دیگر احتمال تسلط آزمودنی‌ها در دو متغیر استدلال کمی و تسلط اندیشه کمتر از سایر مهارت‌ها به دست آمد. مهارت تسلط اندیشه جزء مهارت‌های موردنیاز برای پاسخگویی صحیح به سؤال‌های پایانی است که این سؤال‌ها نیز در زمره سؤال‌های دشوار قرار دارند. در مورد مهارت استدلال کمی باید به این نکته اشاره کرد که این مهارت تنها با ۴ سؤال ارتباط دارد و از طرف دیگر براساس تجربه و پیشینه موجود ضعف دانش‌آموزان در استدلال‌های کمی و ریاضی امری آشکار است؛ بنابراین ضعف دانش‌آموزان در این مهارت نسبت به سایر مهارت‌ها دور از انتظار نیست.

به‌منظور تعیین وضعیت تسلط آزمودنی‌ها ابتدا با پسینی مورد انتظار احتمال تسلط هر آزمودنی در هر مهارت محاسبه گردید. با در نظر گرفتن پیشینه‌های پژوهشی موجود نقطه برش $0/7$ برای حد تسلط تعیین شد؛ افرادی که احتمال تسلطشان بر یک مهارت بیشتر از $0/7$ بود، به‌عنوان افراد مسلط در آن مهارت و در صورتی که این احتمال کمتر از $0/7$ بود، به‌عنوان افراد غیرمسلط تقسیم‌بندی شدند. سپس فراوانی و درصد فراوانی افراد در هر یک از دو گروه مسلط و غیرمسلط محاسبه شد. نتایج پژوهش نشان داد که بیشترین احتمال تسلط مربوط به مهارت‌های ثبات و پیشرفت است که حاکی از نقطه قوت دانش‌آموزان در این دو مهارت است؛ در مجموع وضعیت آزمودنی‌ها در همه مهارت‌ها به‌غیراز دو مهارت استدلال کمی و تسلط اندیشه مطلوب است. در واقع دانش‌آموزان در تولید ایده سریع و سیالی اندیشه و کشف روابط ریاضی ضعف دارند؛ بنابراین پیشنهاد می‌شود که در مدارس به پرورش خلاقیت دانش‌آموزان توجه بیشتری شود و تکالیفی که نیاز به ایده پردازی دارد به عهده آن‌ها سپرده شود. لازم به ذکر است که در این پژوهش فقط به نتیجه نهایی آزمون توجه شد و فرآیند رسیدن به پاسخ بررسی نشد. در جهت رفع این نقصان و پژوهش دقیق‌تر پیشنهاد می‌شود که طی یک پژوهش چگونگی رسیدن به پاسخ صحیح توسط افراد بررسی و مهارت‌های بکار گرفته شده توسط آن‌ها مشخص گردد.

منابع

- اشمن، آدریان. اف و کانوی، روبرت. ان. اف. (۱۳۸۴). *مقدمه‌ای بر آموزش و پرورش شناختی، نظریه و کاربرد*. ترجمه سید کمال خرازی. تهران: سنا.
- استرنبرگ، رابرت. (۱۳۸۷). *روانشناسی شناختی*. ترجمه سید کمال خرازی، الهه حجازی. تهران: سمت.
- امیر تیموری، محمدحسن. (۱۳۸۲). «نظریه هوش‌های چندگانه و برنامه درسی». *تعلیم و تربیت*، ۷۶، ۳۳-۶۶.
- مارنات، گری گراث. (۱۳۹۱). *راهنمای سنجش روانی (جلد اول)*. ترجمه حسن پاشا شریفی و محمدرضا نیکخو (چاپ ششم). تهران: رشد.
- محسن پور، مریم. (۱۳۹۳). *طراحی و ساخت آزمون شناختی- تشخیصی سواد ریاضی کاربردی دانش‌آموزان سال اول دبیرستان و سنجش اثربخشی بسته آموزشی جبرانی مبتنی بر آن*. پایان‌نامه دکتری: دانشگاه تهران.
- مینایی، اصغر. (۱۳۹۱). *مدل‌پردازی تشخیصی شناختی (CDM) سؤال‌های ریاضیات تیمز ۲۰۰۷ در دانش‌آموزان پایه هشتم ایران با استفاده از مدل یکپارچه با پارامتر پردازی مجدد (RUM) و مقایسه مهارت‌های ریاضی دانش‌آموزان دختر و پسر (رساله دکتری سنجش و اندازه‌گیری)*. دانشکده روانشناسی و علوم تربیتی. دانشگاه علامه طباطبائی.
- Armstrong, M. (2011). *Armstrong's handbook of strategic human resource management*. Kogan Page Publishers.
- Chen, C. M. C. (2013). *Examining Uncertainty and Misspecification of Attributes in Cognitive Diagnostic Models* (Doctoral dissertation, Columbia University).
- DeCarlo, L. T. (2011). "On the analysis of fraction subtraction data: The DINA model, classification, latent class sizes, and the Q-matrix". *Applied Psychological Measurement*, 35(1), 8-26.
- Gierl, M. J., Wang, C., Zhou, J. (2008). "Using the Attribute Hierarchy Method to Make Diagnostic Inferences about Examinees' Cognitive Skills in Algebra on the SAT". *Journal of Technology, Learning, and Assessment*, 6(6), n6.
- Henson, R., Douglas, J. (2005). "Test construction for cognitive diagnosis". *Applied Psychological Measurement*, 29(4), 262-277.
- Huo, Y., de la Torre, J. (2014). "Estimating a cognitive diagnostic model for multiple strategies via the EM algorithm". *Applied Psychological Measurement*, 38(6), 464-485.
- Jang, E. E. (2008). "A framework for cognitive diagnostic assessment". *Towards adaptive CALL: Natural language processing for diagnostic language assessment*, 117-131.
- Jang, E. E., Dunlop, M., Park, G., van der Boom, E. H. (2015). "How do young students with different profiles of reading skill mastery, perceived ability, and goal orientation respond to holistic diagnostic feedback?". *Language Testing*, 32(3), 359-383.
- Kunina-Habenicht, O., Rupp, A. A., Wilhelm, O. (2009). "A practical illustration of multidimensional diagnostic skills profiling: Comparing results from

- confirmatory factor analysis and diagnostic classification models". *Studies in Educational Evaluation*, 35(2), 64-70.
- McGlohen, M. K. (2004). *The application of cognitive diagnosis and computerized adaptive testing to a large-scale assessment* (Doctoral dissertation).
- Roussos, L. A., Templin, J. L., Henson, R. A. (2007). "Skills diagnosis using IRT-based latent class models". *Journal of Educational Measurement*, 44(4), 293-311.
- Rupp, A., A., Templin, J., Henson, R., A. (2010). *Diagnostics measurement*. USA: Guilford Press.
- Templin, J. L., Ivie, J. L. (2006, April). *Analysis of the Raven's Progressive Matrices (RPM) scale using skills assessment*. In annual meeting of the National Council on Measurement in Education (NCME), San Francisco, CA.