

سنجش تشخیصی صلاحیت‌های سواد ریاضی

- مریم محسن پور*
- زهرا گویا**
- محسن شکوهی یکتا***
- علیرضا کیامنش****
- عباس بازرگان*****

چکیده: [۱]

در دهه اخیر، رویکرد سنجش شناختی تشخیصی، به منظور تلفیق نظریه‌های شناختی با نظریه‌های یادگیری و ارائه نیمرخ از صلاحیت‌های شناختی دانش‌آموزان، برخلاف روش‌های مرسوم اندازه‌گیری آموزشی، مانند نظریه کلاسیک و نظریه سؤال پاسخ، جایگاه ویژه‌ای پیدا کرده است. با عنایت به چنین دیدگاهی، پژوهشی با عنوان «آزمون تشخیصی صلاحیت‌های شناختی سواد ریاضی» حاوی بیست سؤال، بر مبنای مدل شناختی مطالعه پیزا (۲۰۱۲) و متشکل از صلاحیت‌های شناختی ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی؛ ریاضی‌وار کردن و بازنمایی؛ استدلال و بحث و طراحی راهبرد برای حل مسائل، طراحی گردید. مقاله حاضر نتایج اجرای این آزمون را تبیین می‌کند. جامعه آماری این پژوهش، دانش‌آموزان پایه دهم (دوم دبیرستان) شهر تهران در سال تحصیلی ۹۳-۱۳۹۲ در سه رشته ریاضی فیزیک، علوم تجربی و ادبیات و علوم انسانی بود. با استفاده از رویکرد نمونه‌گیری احتمال متناسب با حجم، نمونه‌ای ۶۸۸ نفری انتخاب شد و آزمون برای این نمونه اجرا گردید. پاسخ‌های سؤال‌ها با استفاده از مدل دینا و به‌کارگیری نرم‌افزار آر تحلیل شده و ظرفیت تشخیصی آزمون در سطح هر سؤال و نحوه ایجاد نیمرخ صلاحیت شناختی دانش‌آموزان ارائه شده است. نتایج این مطالعه نشان داد که با طراحی یک آزمون تشخیصی بر مبنای صلاحیت‌های شناختی می‌توان نیمرخ صلاحیتی دانش‌آموزان را با استفاده از مدل دینا شناسایی کرد.

کلید واژه‌ها: سنجش شناختی تشخیصی، مدل دینا، مطالعه پیزا، سواد ریاضی، مدل‌سازی.

کلید واژه‌ها:

□ تاریخ دریافت مقاله: ۹۲/۷/۳۰

□ تاریخ شروع بررسی: ۹۲/۱۰/۲۱

□ تاریخ پذیرش مقاله: ۹۳/۲/۱۶

□ دانشجوی دکتری سنجش آموزش، دانشگاه تهران mohsenpor@ut.ac.ir

□ استاد دانشگاه شهیدبهبشتی

□□□ دانشیار دانشگاه تهران

□□□□ استاد دانشگاه آزاد اسلامی واحد علوم تحقیقات تهران

□□□□□□ استاد دانشگاه تهران

مقدمه و بیان مسئله

به استناد آخرین ویرایش سند برنامه درسی ملی، توانمندسازی دانش‌آموزان در به‌کارگیری ریاضی در حل مسائل روزمره و امور انتزاعی، یکی از اهداف اساسی آموزش ریاضی در نظام آموزشی ایران است. به این معنا، دانش‌آموزان باید با فرایندهای ریاضی نظیر حل مسئله و به‌کارگیری راهبردهای حل مسئله و مدل‌سازی (مسائل واقعی و پدیده‌ها) آشنا شده و در آن‌ها مهارت یابند (دبیرخانه طرح تولید برنامه درسی ملی، ۱۳۹۰، ص. ۹۲). در سطح بین‌المللی نیز، یکی از اهداف مهم آموزش ریاضی در سطوح مختلف آموزشی، رشد و توسعه صلاحیت مدل‌سازی ریاضی^۱ دانش‌آموزان است (جنسن^۲، ۲۰۰۷؛ نیس، بلوم و گالبرایت^۳، ۲۰۰۷). منظور از صلاحیت مدل‌سازی ریاضی، «توانایی شناسایی سؤال‌ها، متغیرها، روابط و مفروضه‌های^۴ مرتبط در یک موقعیت واقعی، ترجمه آن‌ها در دنیای ریاضی و تفسیر و تأیید راه‌حل به‌دست‌آمده از نتایج مسئله ریاضی در ارتباط با موقعیت دنیای واقعی است» (نیس، بلوم و گالبرایت، ۲۰۰۷، ص. ۱۲).

هنگامی که دانش‌آموزان درگیر مدل‌سازی ریاضی می‌شوند، عمدتاً نیاز به استفاده از مجموعه‌ای صلاحیت‌های شناختی^۵ دارند که نیس (۲۰۰۳) در پروژه‌ای با عنوان صلاحیت‌ها و یادگیری ریاضی، آن‌ها را معرفی کرده است (نیس، ۲۰۱۰). به عبارت دیگر، برای این که فردی به‌طور موفقیت‌آمیز در طیفی از موقعیت‌های حل مسئله (فرا ریاضی یا درون ریاضی)، درگیر فرایند مدل‌سازی شود، ممکن است نیازمند به‌کارگیری همه یا برخی از صلاحیت‌های شناختی باشد (ترنر، دوسی، بلام و نیس^۶، ۲۰۱۲). با این تعبیرها، ایجاد صلاحیت‌های شناختی مدل‌سازی در دانش‌آموزان، به‌منظور حل مسائل زندگی واقعی، ضروری است و زمانی می‌توان از کسب آن‌ها اطمینان حاصل کرد که آزمونی به‌منظور سنجش صلاحیت‌های شناختی ریاضی دانش‌آموزان در اختیار باشد.

امروزه، استفاده از آزمون‌های شناختی تشخیصی^۷ جهت سنجش توانایی‌های آزمون‌شوندگان بیشتر مورد توجه است، زیرا اطلاعات به‌دست‌آمده از این نوع آزمون‌ها، به‌طور معناداری می‌تواند به وسیله نفع‌بران^۸ آموزشی، تفسیر شود (راپ، تمپلین و هنسون^۹، ۲۰۱۰). نکته دیگری که لازم است به آن اشاره شود، این است که تهیه و تدوین ارزشیابی‌های آموزشی بر مبنای سنجش شناختی تشخیصی، مزایای زیادی دارد که شاید بتوان گفت مهم‌ترین آن‌ها، پیوند دادن نظریه‌های شناختی و یادگیری با آموزش باشد، زیرا نتایج این نوع سنجش، گزارشی از نیم‌رخ شناختی دانش‌آموزان در زمینه تکلیف‌های درسی فراهم می‌آورد که بتوان با طراحی‌های مناسب، نقاط ضعف شناسایی شده دانش‌آموزان را برطرف نمود (گیرل، وانگ و ژو^{۱۰}، ۲۰۰۸).

یکی از مطالعات بین‌المللی که بر رویکرد مدل‌سازی و صلاحیت‌های شناختی تأکید داشته، برنامه سنجش بین‌المللی دانش‌آموزان (پیزا)^{۱۱} است. این مطالعه، با هدف ارزیابی بازده‌های نظام‌های آموزشی برحسب پیشرفت دانش‌آموزان در یک چارچوب بین‌المللی متداول در سه حوزه سواد ریاضی^{۱۲}،

خواندن و علوم ایجاد شده است (سازمان همکاری و توسعه اقتصادی، ۲۰۰۰). در مطالعات پیزا و در بخش سواد ریاضی، فرایند مدل‌سازی از مطالعه ۲۰۰۳ تاکنون، یک فرایند بنیادی در چارچوب مطالعه سواد ریاضی محسوب شده است. در چارچوب مطالعه پیزا (۲۰۱۲)، در واقع تعریف سواد ریاضی با مفهوم مدل‌سازی ریاضی به‌طور ویژه تلفیق شده است و دربرگیرنده «توانایی فرد در صورت‌بندی^{۱۳}، به‌کارگیری^{۱۴} و تفسیر^{۱۵} ریاضی در طیفی از زمینه‌های مختلف» است. با این تعبیر، فرایند زیربنایی سنجش سواد ریاضی، مدل‌سازی است که با یک مسئله در موقعیت دنیای واقعی، شروع می‌شود. سپس مسئله از دنیای واقعی به زبان ریاضی صورت‌بندی^{۱۶} می‌گردد. در گام بعدی، مسئله ریاضی حل می‌شود و جواب آن، به زبان دنیای واقعی ترجمه و تفسیر می‌شود (استیسی^{۱۷}، ۲۰۱۲). یک دهه تجربه در تدوین سؤال‌های پیزا و تحلیل روش‌های پاسخ‌گویی دانش‌آموزان به سؤال‌ها، نشان داده است که مجموعه صلاحیت‌های شناختی ریاضی برگرفته از پژوهش‌های نیس و همکاران دانمارکی او (نیس، ۲۰۰۳؛ نیس و جنسن، ۲۰۰۲؛ نیس و هوگارد^{۱۸}، ۲۰۱۱)، حامی ایجاد سواد ریاضی در دانش‌آموزان است (سازمان همکاری و توسعه اقتصادی، ۲۰۱۳).

ایران تاکنون در مطالعه بین‌المللی پیزا شرکت نکرده است تا چگونگی عملکرد دانش‌آموزان ایرانی در حل مسائل دنیای واقعی یعنی سواد ریاضی آن‌ها، تبیین شود. همچنین، بررسی پیشینه پژوهش‌ها در ایران نشان می‌دهد که تاکنون، پژوهشی برای تدوین یک آزمون شناختی تشخیصی، به‌منظور سنجش صلاحیت‌های شناختی سواد ریاضی دانش‌آموزان ایرانی انجام نگرفته است. بنابراین، طراحی یک آزمون به‌منظور سنجش صلاحیت‌های شناختی سواد ریاضی دانش‌آموزان ایرانی در پایان دوره آموزش عمومی بر مبنای مطالعات پیزا، می‌تواند گامی مؤثر در شناسایی ضعف‌ها و قوت‌های صلاحیت‌های شناختی ضروری برای ایجاد سواد ریاضی در دانش‌آموزان ایرانی و حرکت در جهت ارتقای کیفیت سواد ریاضی باشد. بر این اساس، این مقاله که بخشی از یک مطالعه وسیع‌تر است، ظرفیت تشخیصی آزمون سنجش صلاحیت‌های شناختی سواد ریاضی دانش‌آموزان ایرانی را در ابتدای پایه دهم بررسی می‌کند.

■ ضرورت انجام پژوهش

به‌طور کلی، رویکرد سنجش شناختی تشخیصی، با هدف ارتقای سنجش برای یادگیری و فرایند یادگیری^{۱۹}، در مقابله با سنجش بازده‌های یادگیری^{۲۰} از طریق فراهم کردن اطلاعات مورد نیاز برای اصلاح آموزش و یادگیری در کلاس درس، به‌وسیله معلم است (جانگ^{۲۱}، ۲۰۰۸). به‌طور کلی، سنجشی که به معلم اجازه می‌دهد که 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 نه تنها به محتوای ریاضی، بلکه هم‌زمان به بررسی توانایی‌های شناختی دانش‌آموز توجه داشته باشد (اشمن و کانوی^{۲۲}، ۱۳۸۴). در نتیجه، با اندازه‌گیری صلاحیت‌های شناختی بر مبنای سنجش شناختی تشخیصی، ضعف‌ها و قوت‌های آزمون‌شوندگان شناسایی می‌شود و سرانجام، می‌توان درباره صلاحیت‌های حل مسئله آن‌ها، استنباط‌های تشخیصی نمود.

■ برنامه بین‌المللی سنجش دانش‌آموزان (پیزا)

برنامه بین‌المللی سنجش دانش‌آموزان (پیزا)، بر کاربرد ریاضی در زندگی واقعی تأکید دارد. پیزا به‌عنوان یکی از پژوهش‌های پیمایشی «سازمان همکاری و توسعه اقتصادی»، به‌طور ویژه برای دانش‌آموزان ۱۵ ساله در سه حوزه سواد ریاضی، خواندن و علوم، از سال ۲۰۰۰ آغاز شده و هر سه سال یک‌بار، برگزار می‌شود^{۲۳}. به‌طور کلی، آزمون‌های پیزا طوری طراحی می‌شوند تا معیارهایی ارائه دهند که بتوان معلوم کرد دانش‌آموزان تا چه حد می‌توانند با به‌کارگیری آنچه در مدرسه آموخته‌اند، به حل مسائل واقعی و چالش‌هایی که احتمالاً در زندگی روزمره تجربه می‌کنند، بپردازند (سازمان همکاری و توسعه اقتصادی، ۲۰۰۹).

در مطالعه پیزا (۲۰۱۲)، سواد ریاضی دارای جنبه‌های از درون مرتبط زیر است:

- فرایندهای ریاضی^{۲۴} آن چیزی را توصیف می‌کند که فرد برای مرتبط کردن زمینه مسئله با ریاضی انجام می‌دهد تا مسئله را حل کند.
- صلاحیت‌های شناختی^{۲۵} نیز زیربنای فرایندهای ریاضی را توصیف می‌کند.
- دانش محتوای^{۲۶} ریاضی برای پاسخ به سؤال‌های آزمون استفاده می‌شود.
- زمینه^{۲۷} که سؤال‌های آزمون در آن قرار گرفته است.

۱. فرایندهای ریاضی

فرایندهای ریاضی در پیزا، شامل سه بُعد صورت‌بندی موقعیت‌ها به‌صورت ریاضی؛ به‌کارگیری مفاهیم، حقایق، رویه‌ها و استدلال‌های ریاضی؛ و تفسیر کردن، کاربرد و ارزیابی نتایج ریاضی است. منظور از صورت‌بندی این است که دانش‌آموزان، بتوانند محتوا و ساختار ریاضی مورد نیاز برای حل مسائل زمینه‌مدار را تشخیص دهند؛ سپس، قادر باشند مفاهیم، حقایق و رویه‌های ریاضی را برای حل مسائلی که به شکل ریاضی صورت‌بندی شده‌اند، به‌کار گیرند. بالاخره، بُعد تفسیر کردن، بر توانایی دانش‌آموزان برای بررسی راه‌حل‌ها و نتایج ریاضی و تفسیر کردن آن‌ها در زمینه مسائل زندگی واقعی، تمرکز دارد. این کار، شامل ترجمه کردن راه‌حل‌ها یا استدلال‌های ریاضی به دنیای واقعی و میزان تطابق آن‌ها در دنیای واقعی است (سازمان همکاری و توسعه اقتصادی،

۲۰۱۳ الف).

۲. صلاحیت‌های شناختی

در چارچوب مطالعه پیزا (۲۰۰۳)، نیس، توانمندی‌ها را به عنوان صلاحیت‌ها^{۳۸} معرفی کرده است و این صلاحیت‌ها عبارت‌اند از ارتباطات^{۳۹}، مدل‌سازی^{۴۰}، بازنمایی^{۴۱}، تفکر و استدلال^{۴۲}، بحث^{۴۳}، طرح و حل مسئله^{۴۴}، استفاده از زبان و عملیات نمادین، رسمی و فنی^{۴۵} و استفاده از ابزار و وسایل کمکی^{۴۶} (سازمان همکاری و توسعه اقتصادی، ۲۰۰۳، صص. ۴۰-۴۱). در چارچوب مطالعه پیزا (۲۰۱۲)، صلاحیت‌های استدلال و بحث در هم ادغام شده‌اند. در این چارچوب صلاحیت‌ها با نام‌های ارتباطات، ریاضی‌وار^{۴۷} کردن، بازنمایی، استدلال و بحث^{۴۸}، طراحی راهبرد برای حل مسائل^{۴۹}، استفاده از زبان و عملیات نمادین، رسمی و فنی و استفاده از ابزار و وسایل ریاضی^{۵۰} معرفی شده‌اند (استیسی، ۲۰۱۲). منظور از ارتباطات، رمزگشایی و تفسیر اجزای سؤال، به‌منظور معنا بخشیدن به اطلاعات فراهم شده است. علاوه بر این، ریاضی‌وار کردن، شامل تبدیل یک مسئله تعریف‌شده در دنیای واقعی به زبان ریاضی، یا تفسیر یک راه‌حل ریاضی یا یک مدل ریاضی در ارتباط با مسئله اصلی است. همچنین، صلاحیت بازنمایی دربرگیرنده طراحی و ترجمه، به‌کاربردن انواعی از نمایش‌های مختلف برای درک یک موقعیت است. به‌علاوه، استدلال صلاحیتی است که کمک می‌کند تا اجزای گوناگون مسئله کشف و به هم، مرتبط شوند تا بتوان براساس آن‌ها، یک حدسیه را واریسی نمود و به استنباط درست برای یک مسئله رسید. همچنین، طراحی راهبرد برای حل مسائل، توانایی طراحی راهبردهای متنوعی برای حل مسئله‌های ریاضی است و متشکل از راهبردهایی است که شخص را قادر می‌سازد تا ماهیت مسائل را تشخیص داده، صورت‌بندی نموده و حل کند. افزون بر این‌ها، استفاده از زبان و عملیات نمادین، رسمی و فنی، درک کردن، دست‌ورزی کردن و استفاده از عبارات نمادین در یک زمینه ریاضی مانند عبارت‌های حسابی و جبری و به‌کارگیری قواعد ریاضی است. صلاحیت استفاده از ابزار و وسایل ریاضی شامل شناخت و توانایی به‌کارگیری ابزارهای ریاضیات مانند ابزار اندازه‌گیری، ماشین حساب هستند (سازمان همکاری و توسعه اقتصادی، ۲۰۱۳). همچنین بنا بر یافته‌های ترنر^{۴۱} (۲۰۱۲)، ترنر و آدامز^{۴۲} (۲۰۱۲) و ترنر و همکاران (۲۰۱۳) شش صلاحیت شناختی ارتباطات، ریاضی‌وار کردن، بازنمایی، استدلال و بحث، طراحی راهبرد برای حل مسائل، استفاده از زبان و عملیات نمادین، رسمی و فنی در هر یک از سه مرحله فرایندهای ریاضی به‌کار گرفته می‌شوند و با افزایش سطح سواد ریاضی شخص، میزان به‌کارگیری صلاحیت‌های شناختی توسط وی افزایش می‌یابد. یعنی افزایش سطح صلاحیت‌های شناختی، با سطح دشواری سؤال‌ها مرتبط است.

۳. دانش محتوایی ریاضی

تغییر و روابط^{۴۳}، فضا و شکل^{۴۴}، کمیت^{۴۵} و عدم قطعیت^{۴۶}، چهار حیطه محتوایی برای تدوین

سنجش تشخیصی صلاحیت‌های سواد ریاضی

سؤال‌های ریاضی پیزا هستند. منظور از تغییر و روابط، درک تغییرات بنیادین و تشخیص زمان رخ دادن آن‌ها است تا شخص بتواند به‌طور مناسب، از مدل‌های ریاضی برای توصیف و پیش‌بینی متغیرها، استفاده کند. همچنین، فضا و شکل به طیفی از فعالیت‌ها مانند درک منظره‌ها، خواندن نقشه‌ها، شکل‌های در حال تغییر، تفسیر نماهای سه بعدی از زاویه‌های مختلف و ساختن بازنمایی اشکال، اطلاق می‌شود. به‌علاوه، کمیت مانند کمی سازی ویژگی‌های اشیاء، روابط و موقعیت‌ها در دنیای واقعی است. علاوه بر این‌ها، درک عدد، بازنمایی‌های چندگانه اعداد، محاسبات ذهنی و تخمین و ارزشیابی هم در این حیطه قرار می‌گیرند. بالاخره عدم قطعیت، به تشخیص جایگاه تغییر در فرایندها اطلاق می‌شود (سازمان همکاری و توسعه اقتصادی، ۲۰۱۳ الف).

۴. زمینه

زمینه، جنبه‌ای از دنیای واقعی فرد است که مسئله‌ها، در آن واقع می‌شوند. انتخاب راهبردها و بازنمایی‌های مناسب، اغلب بستگی به زمینه‌ای دارند که مسئله از آن نشأت گرفته است. در مطالعه پیزا (۲۰۱۲)، چهار زمینه شخصی، شغلی، اجتماعی و علمی، مورد نظر است. مسائل طبقه‌بندی شده در زمینه شخصی، بر فعالیت‌های شخصی، خانوادگی یا گروه همسالان متمرکز است. همچنین، مسئله‌های زمینه شغلی، عمدتاً مرتبط با بازار کار و اشتغال هستند. زمینه اجتماعی مسائلی را در بر می‌گیرد که بر وجوه مختلف جامعه‌ای که فرد در آن زندگی می‌کند - بومی، ملی یا جهانی - متمرکزند. به‌علاوه، مسائل دارای زمینه علمی با کاربردهای ریاضی در دنیای واقعی مسائل و موضوعات مرتبط با علوم و فناوری مرتبطاند (سازمان همکاری و توسعه اقتصادی، ۲۰۱۳ الف).

■ سنجش شناختی تشخیصی^{۴۷} علوم انسانی و مطالعات فرهنگی

در چند دهه اخیر؛ با توجه به نتایج مطالعات انجام‌شده توسط محققان آموزشی و روان‌شناسان به‌ویژه در حوزه علوم شناختی، یکی از رویکردهای پیشنهادی و قابل دفاع در فعالیت‌های آموزشی، رویکرد شناختی بوده است (غلام آزاد، ۱۳۹۱ به نقل از کلمنتس و سامارا، ۲۰۰۹). موضوع روان‌شناسی شناختی^{۴۸}، فعالیت‌های ذهنی یا فکری انسان است. به عبارت دیگر، موضوع این علم، فرایندهایی است که از طریق آن‌ها، انواع اطلاعات ورودی به مغز، دچار دگرگونی، کاستی، افزایش یا بسط و ذخیره و بازیابی می‌شوند و مورد استفاده قرار می‌گیرند. این فرایندها شامل ادراک، بازشناسی الگو، حافظه، یادگیری و زبان و فرایندهای دیگری هستند که همیشه مورد توجه پژوهشگران بوده‌اند. کنش‌های عالی‌تر شناختی مانند حل مسئله و تصمیم‌گیری که خود متکی بر کنش‌های ساده‌ترند، در این علم مورد تحلیل قرار می‌گیرند (خرازی و دولتی، ۱۳۸۸).

از جمله ابعاد مهم فعالیت‌های آموزشی، اندازه‌گیری آموزشی است. نقش مهمی که نظریه

شناختی می‌تواند در اندازه‌گیری‌های آموزشی داشته باشد، برای بسیاری از متخصصان آموزشی آشکار گردیده است (لیتون^{۴۹}، گیرل و هونکا^{۵۱}، ۲۰۰۴). به‌عنوان نمونه، اسنو و لومن^{۵۱} (۱۹۸۹) در راستای کمک به بهبود سنجش و اندازه‌گیری آموزشی با استفاده از روان‌شناسی شناختی، گامی بزرگ برداشته و در فصل هفتم کتاب اندازه‌گیری آموزشی لین (۱۹۸۹)، به دلالت‌های روان‌شناسی شناختی برای اندازه‌گیری آموزشی پرداخته‌اند. از نظر این دو نویسنده، پژوهش‌های روان‌سنجی، مدل‌های اندازه‌گیری روان‌سنجی آموزشی^{۵۲} قدرتمند متنوعی مانند نظریه کلاسیک آزمون، مدل راش و نظریه سؤال پاسخ برای اندازه‌گیری‌های آموزشی ایجاد کرده است که با وجود تفاوت مفروضه‌های آنها، همگی به دنبال برآورد جایگاه فرد در یک متغیر مکنون^{۵۳} زیربنایی هستند، با این حال، از دیدگاه روان‌شناسی شناختی، اسنو و لومن (۱۹۸۹) سه محدودیت مهمی که مدل‌های اندازه‌گیری روان‌سنجی آموزشی در اندازه‌گیری آموزشی دارند، را مطرح می‌کنند که عبارت‌اند از لحاظ نکردن توجهات روان‌شناسی زیربنایی در سطح عملکرد یک سؤال، در نظر گرفتن مفروضات ساده در زمینه روان‌شناسی سؤال‌ها و ضمنی شدن یا حذف شدن فرایندهای روان‌شناسی آزمون. به‌عنوان مثال در زمینه محدودیت اول، توجه اولیه‌ای در استفاده از مدل سه پارامتری لجستیک برای تحلیل داده‌ها وجود ندارد. به عبارت دیگر، این مدل باید براساس نتایجی که از تحلیل داده‌ها به‌دست می‌دهد، توجه شود نه براساس معقول بودن روان‌شناختی‌اش. در خصوص محدودیت دوم، هر دو نظریه کلاسیک آزمون و تعمیم‌پذیری، مفروضه ناهمبسته بودن خطاها را می‌پذیرند، با این حال در مورد تکلیف‌های شناختی، آزمون‌شوندگان اغلب در حین آزمون نسبت به موفقیت یا شکست در مورد سؤال قبلی، عکس‌العمل نشان می‌دهند و احتمالاً به‌گونه‌ای عمل می‌کنند تا مفروضه‌های استقلال را زیر پا بگذارند. همچنین، محدودیت سوم به مسائل اعتبار سازه مرتبط است. مدل‌های اندازه‌گیری روان‌سنجی آموزشی سنتی، به‌طور صریح تبیین‌های اساسی را از عملکردهایی که آنها مدل‌سازی می‌کنند، لحاظ نمی‌کنند، اگرچه ضرورت تبیین معتبر، از نمره‌های آزمون به عنوان یک استاندارد برای اندازه‌گیری‌های آموزشی و روان‌شناختی پذیرفته شده، عمل اعتباریابی آزمون، عمدتاً تأیید^{۵۴} بیرونی از تفسیر آزمون است و به ندرت به درون آزمون نفوذ پیدا می‌کند.

از نظر لیتون و گیرل (۲۰۰۷)، ایده‌های بیان شده به‌وسیله اسنو و لومن (۱۹۸۹)، بسیاری از پژوهشگران آموزشی را به بررسی بالقوه یک شاخه نسبتاً جدید از روان‌شناسی یعنی «روان‌شناسی شناختی» در غنی ساختن آزمون‌های شناختی ترغیب کرد. لذا، از سال ۱۹۸۹، مقالات و کتاب‌های متنوعی در زمینه سنجش شناختی تشخیصی نوشته شد که از آن بین، می‌توان به فردریکسون، گلیرز، لسگلد و شافتو^{۵۵}، (۱۹۹۰)؛ نیکولز^{۵۶}، (۱۹۹۴)؛ نیکولز، چیپ‌من و برنان^{۵۷}، (۱۹۹۵) اشاره نمود. به‌طور کلی، هدف از طراحی سنجش شناختی تشخیصی، اندازه‌گیری ساختار دانش و مهارت‌های پردازشی آزمون‌شوندگان به‌منظور فراهم آوردن اطلاعاتی در زمینه نقاط قوت و ضعف شناختی آنهاست.

ساختار دانش شامل اطلاعات رویه‌ای و حقایقی هستند، در حالی که صلاحیت‌های پردازشی دربرگیرنده تحولات و راهبردهای مورد نیاز برای دست‌ورزی با این اطلاعات‌اند (لهمن، ۲۰۰۰، به نقل از لیتون و گیرل، ۲۰۰۷).

۱. مدل‌های آماری سنجش شناختی تشخیصی

برای تحلیل داده‌های آزمون‌شوندگان در سنجش شناختی تشخیصی، از مدل‌های آماری روان‌سنجی استفاده می‌شود که به طبقه‌بندی پاسخ‌های آزمون‌شوندگان براساس متغیرهای مکنون چندگانه می‌پردازد که معمولاً، مدل‌های طبقه‌بندی تشخیصی^{۵۸} نامیده می‌شوند. در هر یک از این مدل‌ها، عملکرد هر آزمون‌شونده، برحسب احتمال تسلط وی بر هر یک از صلاحیت‌ها یا احتمال تعلق به یک کلاس مکنون تعیین می‌شود (راپ، تمپلین و هسون، ۲۰۱۰).

در مدل‌های آماری سنجش شناختی تشخیصی، صلاحیت‌های شناختی در قالب نیمرخ‌های صلاحیتی سؤال‌ها، نمایش داده می‌شوند. منظور از نیمرخ صلاحیتی در هر سؤال، الگویی ترکیبی از مقادیر ویژه برای صلاحیت‌هاست. بنابراین نیمرخ صلاحیتی در هر سؤال، نمایانگر نیاز به تسلط بر هر صلاحیت است. گفتنی است، نیمرخ صلاحیت برای مجموعه سؤال‌ها در یک ماتریس، به نام ماتریس کیو^{۶۰}، نمایش داده می‌شود. بنابراین ماتریس کیو، نشان‌دهنده ساختار بارگیری در مدل‌های شناختی است. هر سطر این ماتریس، یک فرضیه درباره صلاحیت‌های مورد نیاز برای کسب پاسخ درست به یک سؤال است و دارای مرتبه $n \times k$ است که در آن n ، تعداد سؤال‌ها و k ، تعداد صلاحیت‌های مورد نیاز برای سؤال‌هاست. تمام درایه‌های این ماتریس، اعداد صفر و یک هستند. برای یک درایه خاص ماتریس Q در سطر n ام و ستون k ام، عدد ۱ نشانگر آن است که سؤال n ام خصیصه یا صلاحیت k ام را اندازه می‌گیرد و عدد صفر نشان می‌دهد که سؤال، صلاحیت مورد نظر را اندازه نمی‌گیرد (همان).

۱-۱. ویژگی‌های مدل‌های شناختی تشخیصی

از ویژگی‌های مهم مدل‌های سنجش شناختی تشخیصی، چگونگی تعامل متغیرهای مکنون با یکدیگر، برای تولید پاسخ قابل مشاهده است. نوع تعامل متغیرهای مکنون با هم، مدل‌ها را به دو دسته‌بندی مدل‌های جبرانی^{۶۱} و مدل‌های غیرجبرانی تقسیم کرده است. در مدل‌های جبرانی، نقص در یک صلاحیت از طریق قوت در صلاحیت دیگر جبران می‌شود، در حالی که در مدل‌های غیرجبرانی لازم است تا همه صلاحیت‌ها وجود داشته باشند تا بتوان به سؤال پاسخ درست داد (راپ و همکاران، ۲۰۱۰). به‌عنوان مثال، برای به‌دست آوردن پاسخ درست مسئله، $\frac{3}{4} - \frac{3}{8}$ ، شخص باید بداند که چگونه مخرج مشترک بگیرد و سپس صورت کسرها را از هم کم کند. در دیدگاه غیرجبرانی، اگر شخص فاقد تسلط به هر یک از این صلاحیت‌ها باشد، صرف‌نظر از هر صلاحیت دیگری که داشته باشد، با احتمال زیاد نمی‌تواند پاسخ درست سؤال را به‌دست آورد (تمپلین، ۲۰۰۶).

همچنین، یک خصوصیت مهم مدل‌های شناختی تشخیصی، ماهیت چند بعدی بودن آن‌ها است. به دلیل اینکه این مدل‌ها مبتنی بر متغیرهای مکنون چندگانه به منظور طبقه‌بندی پاسخ‌دهندگان هستند، بر حسب تعریف، مدل‌های چندبعدی نیز نامیده می‌شوند. ویژگی دیگر مدل‌های شناختی تشخیصی، تأییدی بودن آن‌ها از دو دیدگاه است؛ یکی دیدگاه اساسی از کاربرد و دیگری دیدگاه آماری از ساختار مدل. از دیدگاه کاربردی، چون مدل‌ها برای تأیید یا رد کردن یک فرضیه ویژه درباره روابط بین اینکه چگونه پاسخ‌دهندگان نسبت به سؤال عکس‌العمل نشان می‌دهند و ویژگی‌های ذهنی زیربنایی. از دیدگاه آماری، مدل‌ها تأییدی هستند زیرا در آن‌ها ساختار بارگیری از قبل در قالب ماتریس Q تعیین می‌شود، بنابراین تعیین ساختار ماتریس کیو، این نوع مدل‌ها را به مدل‌های تحلیل عاملی تأییدی شبیه ساخته است (راپ و همکاران، ۲۰۱۰).

۲. سنجش شناختی تشخیصی در ریاضی

دو رویکرد در طراحی مدل‌های سنجش شناختی تشخیصی وجود دارد که عبارت‌اند از استفاده از یک رویکرد استقرایی^{۶۲} به منظور ساخت یک مجموعه سؤال تشخیصی برای استنباط فرایندهای صلاحیتی و ساختار دانش مورد نظر و استفاده از یک رویکرد مهندسی معکوس^{۶۳} به منظور استخراج فرایندهای شناختی و صلاحیت‌ها از آزمون‌های موجود، به قصد کسب اطلاعات غنی‌تر، در مقایسه با آنچه که مقیاس‌پردازی تک بعدی به دست می‌دهد^{۶۴} (جانگ، ۲۰۰۸). تاکنون در بیشتر پژوهش‌های انجام‌شده در حوزه سنجش شناختی تشخیصی و در زمینه ریاضی، از رویکرد مهندسی معکوس برای آزمون‌هایی که اساساً با هدف شناختی تشخیصی طراحی نشده‌اند، استفاده شده است (گیرل، ونگ و ژو، ۲۰۰۸؛ لی، پارک و تایلان^{۶۵}، ۲۰۱۱؛ باسوکو، اگرتمن و قلعج‌اوغلو^{۶۶}، ۲۰۱۳؛ مینایی، ۱۳۹۱) و پژوهش‌های کمتری در زمینه طراحی آزمون‌های شناختی تشخیصی با زمینه ریاضی صورت گرفته است (گیرل، آلویز و ماجیو، ۲۰۱۰؛ روبرتز و گیرل^{۶۷}، ۲۰۱۱ و بردشاو، ایساک، تمپلین و جاکویسون^{۶۸}، ۲۰۱۳). در یک مطالعه شناختی تشخیصی، گیرل، ونگ و ژو (۲۰۰۸) با استفاده از رویکرد مهندسی معکوس و به‌کارگیری روش سلسله‌مراتبی خصیصه^{۶۹} به‌عنوان یکی از مدل‌های آماری سنجش شناختی تشخیصی، به بررسی سؤال‌های جبر در آزمون استعداد تحصیلی^{۷۰} که در سال ۲۰۰۵ اجرا شده بود، پرداختند. این مطالعه در دو مرحله انجام شد که در مرحله اول، متخصصان حوزه محتوایی ریاضی به روش تحلیل تکلیف، سؤال‌های منتشرشده جبر را بررسی و صلاحیت‌های مهم مبتنی بر آن سؤال‌ها را شناسایی کردند. سپس صلاحیت‌ها را در یک سلسله مراتب منظم کرده و چندین مدل شناختی از آن، استخراج کردند. به عنوان نمونه، «نمایش دادن صلاحیت‌های عملیات پایه‌ای حساب»^{۷۱}، «دانش درباره ویژگی‌های عامل‌ها»^{۷۲} و صلاحیت «به‌کارگیری قواعد فاکتورگیری»^{۷۳} از جمله صلاحیت‌های به‌دست‌آمده برای جبر بود. در مرحله دوم، تحلیل‌های روان‌سنجی روی

سنجش تشخیصی صلاحیت‌های سواد ریاضی

مدل‌های شناختی انجام شد. ارزیابی برازش این مدل‌ها با داده‌ها به وسیله مقایسه الگوی پاسخ مورد انتظار تولیدشده توسط مدل شناختی و الگوی پاسخ مشاهده‌شده تولیدشده به وسیله نمونه تصادفی ۵۰۰۰ نفری با استفاده از شاخص سازگاری سلسله مراتبی^۴، انجام شد. متوسط این شاخص حدود ۰/۸ بود که نشان از برازش خوب مدل با داده‌ها داشت.

در مطالعه شناختی تشخیصی دیگری، گیرل، آلویز و ماجیو (۲۰۱۰)، با هدف ارتقای استنباط‌های شناختی درباره صلاحیت‌های حل مسئله دانش‌آموزان پایه‌های سوم و ششم، از رویکرد استقرایی استفاده کردند. آنان ابتدا، با روش تحلیل تکلیف، صلاحیت‌های شناختی را در حیطه‌های اعداد، الگو و روابط، شکل و فضا و آمار و احتمال، تدوین کردند که در پایه سوم، ۲۶ مدل شناختی شامل ۱۷۸ صلاحیت و در پایه ششم ۲۶ مدل شناختی با ۱۵۰ صلاحیت به دست آمد. سپس یک بانک سؤال برای اندازه‌گیری این صلاحیت‌ها طراحی کردند. در مرحله آخر، تحلیل‌های روان‌سنجی تأییدی، با استفاده از برازش مدل با داده‌ها در نمونه‌های ۳۳۸ و ۱۸۴ نفری، به ترتیب در پایه‌های سوم و ششم صورت گرفت. شاخص سازگاری سلسله مراتبی برای دو پایه سوم و ششم به ترتیب ۰/۶۶ و ۰/۶۷ بود که نشان‌دهنده برازش متوسط تا خوب مدل با داده‌ها بود.

■ روش پژوهش

پژوهش حاضر ظرفیت تشخیصی آزمون سنجش صلاحیت‌های شناختی سواد ریاضی دانش‌آموزان ایرانی را در ابتدای پایه دهم بررسی می‌کند. در ادامه، به معرفی پژوهش و جزئیات لازم آن، پرداخته می‌شود.

۱. ساخت آزمون صلاحیت‌های شناختی سواد ریاضی

در این پژوهش، مبنای ساخت آزمون صلاحیت‌های شناختی سواد ریاضی، استفاده از چارچوب مطالعه پیزا (۲۰۱۲) بود. همچنین، رویکرد مورد استفاده در چارچوب سنجش شناختی تشخیصی، رویکرد استقرایی بود. بنابراین، به منظور بررسی شش صلاحیت شناختی سواد ریاضی^۵، یک آزمون شناختی تشخیصی برای سنجش سواد ریاضی، در راستای آزمون پیزا، طراحی شد. ابتدا در یک مطالعه مقدماتی، مجموعه‌ای شامل ۳۳ سؤال تهیه شد و با استفاده از نظر چند دبیر ریاضی، سؤال‌ها جرح و تعدیل شدند و پس از آن براساس راهنمای بررسی^۶ سطوح صلاحیت‌های شناختی (پیوست الف)، با چهار سطح تعریف‌شده ناچیز (۰)، کم (۱)، متوسط (۲) و زیاد (۳)، درجه‌بندی شدند. پس از اجرای مقدماتی و بازنگری مجدد در ساختار سؤال‌های آزمون، مجموعاً ۲۰ سؤال انتخاب شد.^۷ ملاک انتخاب سؤال‌ها، براساس پوشش الف، سطوح صلاحیت‌های شناختی، ب. ابعاد مدل‌سازی در سازه سواد ریاضی، ج. زمینه و د. محتوای سؤال‌ها بود.^۸

جدول (۱)، ویژگی‌های سؤال‌های آزمون صلاحیت‌های شناختی را نشان می‌دهد.

جدول ۱. ویژگی ابعاد آزمون صلاحیت‌های شناختی سواد ریاضی

واحد	سؤال	فرایند	محتوا	زمینه	نوع سؤال	درصد پاسخ درست
فروش گوشی تلفن همراه	۱	تفسیر	عدم قطعیت	شغلی	چند گزینه‌ای	۹۵
	۲	تفسیر	عدم قطعیت	شغلی	چند گزینه‌ای	۷۸/۲
	۳	به‌کارگیری	عدم قطعیت	شغلی	چند گزینه‌ای	۶۱
خرید خانه	۴	صورت‌بندی	فضا و شکل	شخصی	پاسخ ساختنی	۳۹/۷
	۵	صورت‌بندی	فضا و شکل	شخصی	پاسخ ساختنی	۱۳/۲
کوهنوردی	۶	صورت‌بندی	کمیت	اجتماعی	چند گزینه‌ای	۷۴/۶
	۷	صورت‌بندی	تغییر و رابطه	اجتماعی	پاسخ ساختنی	۳۶/۴
	۸	به‌کارگیری	کمیت	اجتماعی	پاسخ ساختنی	۳۱/۲
سوله	۹	تفسیر	فضا و شکل	شغلی	چند گزینه‌ای	۸۸
	۱۰	به‌کارگیری	فضا و شکل	شغلی	پاسخ ساختنی	۱۴/۹
کدام ماشین	۱۱	تفسیر	عدم قطعیت	شخصی	چند گزینه‌ای	۸۶/۵
	۱۲	به‌کارگیری	کمیت	شخصی	چند گزینه‌ای	۶۵/۸
	۱۳	به‌کارگیری	کمیت	شخصی	پاسخ ساختنی	۴۹/۲
نجار	۱۴	به‌کارگیری	فضا و شکل	شغلی	چند گزینه‌ای	۱۰/۷
استقبال	۱۵	صورت‌بندی	کمیت	اجتماعی	چند گزینه‌ای	۱۸/۳
کشتی دریانوردی	۱۶	به‌کارگیری	کمیت	علمی	چند گزینه‌ای	۴۸/۶
	۱۷	به‌کارگیری	فضا و شکل	علمی	چند گزینه‌ای	۵۱
	۱۸	صورت‌بندی	تغییر و روابط	علمی	پاسخ ساختنی	۱۸/۹
زباله	۱۹	تفسیر	عدم قطعیت	علمی	پاسخ ساختنی	۳۰/۷
تانک آب	۲۰	صورت‌بندی	تغییر و روابط	علمی	چند گزینه‌ای	۳۲/۹

۲. ساخت ماتریس کیو

اگرچه در بخش طراحی آزمون سواد ریاضی، مجموعه سؤال‌های آزمون توسط دبیران داوطلب و براساس راهنمای بررسی صلاحیت‌های شناختی، طبق سطوح این صلاحیت‌ها، درجه‌بندی شده بود. اما پس از نهایی شدن سؤال‌های آزمون نیز، مجدداً این سؤال‌ها، براساس راهنمای بررسی سطوح

سنجش تحصیلی صلاحیت‌های سواد ریاضی

صلاحیت‌های شناختی، توسط یک استاد دانشگاه با گرایش آموزش ریاضی و دو دانشجوی دکتری آموزش ریاضی که از دبیران ریاضی باتجربه در دوره متوسطه هستند، و یک دبیر باتجربه ریاضی و یک دانشجوی دکتری سنجش آموزش، درجه‌بندی شدند. سپس میانگین سطوح درجه‌بندی این افراد محاسبه شد و بر مبنای آن، ماتریس کیو با شش صلاحیت شکل گرفت.

با توجه به مداخله اغلب صلاحیت‌ها در بیشتر سؤال‌ها، ماتریس کیو دارای پیچیدگی زیادی بود و احتمال عدم تعیین^۹ مدل وجود داشت. بنابراین، تصمیم به ترکیب برخی صلاحیت‌های شناختی قابل تلفیق گرفته شد. بررسی ماتریس کیوی شش صلاحیتی نشان داد که دو صلاحیت ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی در بیشتر سؤال‌ها با هم ظاهر می‌شوند، همچنین، بررسی همبستگی صلاحیت‌ها نشان داد که در بین ضرایب همبستگی تراکوریک بین صلاحیت‌ها، دو صلاحیت مذکور، بیشترین ضریب همبستگی را با هم دارند (۰/۷۶). از طرفی، ادبیات مربوط به مدل شناختی به‌کار برده شده در این پژوهش، حاکی از این بود که مجموعه شش صلاحیت شناختی در دو گروه دسته‌بندی می‌شوند که دو صلاحیت ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی، در یک گروه مشترک قرار داشتند. از این گذشته، پیشینه صلاحیت‌های شناختی نشان می‌داد که اصولاً برای به‌کارگیری صلاحیت ریاضی‌وار کردن، نیاز به استفاده از صلاحیت بازنمایی وجود دارد (نیس و هوچگارد، ۲۰۱۱). بنابراین تصمیم به تلفیق دو صلاحیت ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی با هم، و نیز ادغام دو صلاحیت ریاضی‌وار کردن و بازنمایی در یک صلاحیت، گرفته شد. سپس یک جلسه هم‌اندیشی (گروه کانونی) به‌منظور تدوین نهایی ماتریس کیو تشکیل گردید. در ادامه، بخشی از ماتریس کیو در جدول (۲) ارائه شده است.

جدول ۲. بخشی از ماتریس کیو در این پژوهش

صلاحیت سؤال	ارتباطات و استفاده از زبان و عملیات	ریاضی‌وار کردن و بازنمایی	استدلال	طراحی راهبرد
۱	۱	۰	۰	۰
۲	۱	۱	۰	۰
۳	۱	۱	۰	۱
۴	۱	۱	۱	۰
۵	۱	۱	۱	۱

در جدول (۲)، پنج سؤال اول آزمون ارائه شده است. سه سؤال اول، مربوط به واحد فروش گوشی تلفن همراه هستند که دو سؤال اول و دوم، به ترتیب نیاز به یک و دو صلاحیت شناختی و سؤال سوم نیاز به سه صلاحیت شناختی دارد. سؤال چهارم و پنجم نیز مربوط به واحد خرید خانه هستند که به ترتیب سه و چهار صلاحیت شناختی را طلب می‌کنند.

۳. جامعه آماری و نمونه

جامعه آماری این پژوهش، کلیه دانش‌آموزان پایه دهم شهر تهران در سال تحصیلی ۱۳۹۳-۱۳۹۲ در سه رشته ریاضی فیزیک، علوم تجربی و ادبیات و علوم انسانی بودند. انتخاب پایه دهم به دو دلیل صورت گرفت؛ نخست آن‌که در این پایه، دانش‌آموزان به تازگی آموزش عمومی را به پایان می‌رسانند و دلیل دوم این‌که طیف سنی دانش‌آموزان در زمان آزمون، ۱۵/۲ سال تا ۱۶/۲ سال است که با جمعیت هدف مطالعات پی‌زده همخوانی دارد. علاوه بر این، در مورد آزمون‌های شناختی تشخیصی، که بزرگی حجم نمونه با توجه به توصیه شو، هنسون و ویلز^{۸۱} (۲۰۱۳) لازم است تا بتوان برآورد قابل‌اعتمادی از پارامترهای مدل شناختی به دست آورد. از این رو، از فرمول‌های متداول نمونه‌گیری که حداکثر حجم نمونه را با جمعیتی حدود ۶۰,۰۰۰ نفر بین ۳۰۰ تا ۴۰۰ نفر تعیین می‌کنند استفاده نشد. در عوض، برای تعیین حجم مناسبی از جامعه آماری، با نویسندگان دوم و سوم کتاب سنجش شناختی^{۸۱} مکاتبه شد و بر مبنای راهنمایی آن‌ها، یک حجم نمونه ۶۰۰ نفری در نظر گرفته شد. رویکرد نمونه‌گیری در این پژوهش، رویکرد احتمال متناسب با حجم^{۸۲} بود. در این روش، حجم مناطق و مدارس، عاملی در افزایش احتمال انتخاب منطقه و مدرسه است.

براساس رویکرد مذکور، ابتدا آمار دانش‌آموزان نواحی ۱۹ گانه آموزش و پرورش تهران، به تفکیک هر ناحیه، تعیین شد. سپس براساس جمعیت آماری و با استفاده از جدول اعداد تصادفی، کد پنج‌انگاره آموزش تعیین و براساس آن، منطقه‌ای که دانش‌آموز مورد نظر در آن قرار داشت مشخص شد. براین اساس مناطق ۳، ۴، ۵، ۶ و ۱۸ انتخاب شدند. در مرحله بعد، با توجه به حجم نواحی منتخب و تعداد کلاس‌های موجود در این نواحی، میانگین کلاسی به دست آمد. همچنین، مطابق با حجم نمونه ۶۰۰ نفری، تعداد کلاس‌ها تعیین و میانگین کلاس‌ها در هر مدرسه مشخص شد. در نهایت، براساس حجم نواحی منتخب و تعداد مدارس، فاصله طبقاتی محاسبه شد (پیوست ب). براساس جدول اعداد تصادفی، اولین انتخاب نفر ۶۳۴ بود که مدرسه این فرد شناسایی شد و جزو نمونه قرار گرفت. با اضافه کردن عدد فاصله طبقاتی به اولین انتخاب، نفر بعدی و به تبع آن مدرسه این فرد تعیین شد. این فرایند تا انتخاب ۱۹ مدرسه ادامه یافت. با مشخص شدن مدارس، در گام بعدی و تعیین کلاس‌های مدارس منتخب، در صورت وجود یک یا دو کلاس در مدرسه، یک کلاس

و در صورت وجود بیش از دو کلاس، دو کلاس انتخاب شدند. به این ترتیب، حجم احتمالی نمونه ۷۰۰ نفر به دست آمد. موقع اجرای آزمون، یک نمونه ۶۸۸ نفری به سؤال‌های آزمون پاسخ دادند که پس از حذف ۳۱ برگه مخدوش (شامل سفید، نیمه کاره و...)، تعداد نهایی به ۶۵۷ نفر رسید.

۴. مدل روان‌سنجی: دینا

ترنر، دوسی، بلام و نیس، (۲۰۱۳) معتقدند که صلاحیت‌های شناختی، کاملاً منفصل از هم نیستند بلکه به میزان معینی هم‌پوشانی دارند و تقریباً به‌طور مشترک، در فرایند حل مسئله ریاضی فعال می‌شوند و بدین سبب، غیرجبرانی هستند. پس با عنایت به حجم نمونه، مناسب‌ترین مدل برای تحلیل داده‌ها، مدلی غیرجبرانی و غیر حساس به حجم نمونه شناخته شد. براین اساس، مدل دینا^{۸۳} که غیرجبرانی است، برای تحلیل داده‌ها انتخاب شد.

مدل غیرجبرانی دینا اولین بار، توسط جانکر و ستیجما^{۸۴} (۲۰۰۱) معرفی شد. این مدل در مقایسه با سایر مدل‌های شناختی تشخیصی، نه تنها مقرون به صرفه^{۸۵} است، بلکه به سادگی قابل تفسیر است. واژه Deterministic، به جنبه تعینی^{۸۶} مدل، واژه And-Gate، به فرایند ربطی^{۸۷} مدل و واژه Noisy به عنصر تطوری^{۸۸} (احتمالی) مدل اشاره دارد (به نقل از دلاتوره^{۸۹}، ۲۰۰۹).

جنبه تعینی (قطعی) مدل به تولید پاسخ مکنون η_{ij} مرتبط است که به وسیله دو مقدار α_i (الگوی صلاحیت برای فرد i ام) و q_j (ردیفی از ماتریس کیو که به سؤال j ام مربوط است) از طریق معادله زیر تعیین می‌شود.

$$\eta_{ij} = \prod_{k=1}^K \alpha_k^{q_{jk}}$$

پاسخ مکنون تعینی η_{ij} ، دلالت بر این دارد که آیا فرد i ام همه صلاحیت‌های لازم برای سؤال j ام را دارد یا خیر. اگر فردی به همه صلاحیت‌های لازم برای یک سؤال تسلط داشته باشد، η_{ij} برابر یک و اگر حداقل به یک صلاحیت تسلط نداشته باشد، η_{ij} برابر صفر است.

ربطی بودن این مدل بدین معناست که تسلط بر همه صلاحیت‌های مشخص شده برای یک سؤال در ماتریس کیو، برای پاسخگویی درست به سؤال لازم است و اگر فردی به برخی از این صلاحیت‌ها تسلط داشته باشد، احتمال موفقیتش برابر با فردی است که به هیچ یک از صلاحیت‌ها تسلط ندارد. بر این اساس در این مدل، آزمودنی‌ها در هر سؤال، به دو کلاس تقسیم می‌شوند. یک کلاس آزمودنی‌هایی که به همه صلاحیت‌ها تسلط دارند و کلاس دیگر آزمودنی‌هایی که حداقل در

یکی از صلاحیت‌ها به حد تسلط نرسیده‌اند.

همچنین، عنصر تطوری (احتمالی) مدل نیز بدین معناست که تسلط داشتن به همه صلاحیت‌های لازم برای یک سؤال، پاسخ درست دادن را تضمین نمی‌کند و تسلط نداشتن به همه صلاحیت‌های یک سؤال، پاسخ نادرست دادن را تضمین می‌کند. با توجه به جنبه تطوری مدل، دو پارامتر s_j و g_j در یک سؤال ایجاد می‌شود. پارامتر s_j به احتمال لغزش^{۹۱} و پاسخ نادرست دادن به سؤال هنگامی که η_{ij} برابر یک است، اشاره دارد. پارامتر g_j به احتمال درست حدس^{۹۱} زدن پاسخ سؤال هنگامی که η_{ij} برابر صفر است، اشاره دارد. در زیر تعریف این دو پارامتر ارائه شده است:

$$s_j = P(X_j = 0 | \eta_j = 1)$$

$$g_j = P(X_j = 1 | \eta_j = 0)$$

در مدل دینا، احتمال پاسخ درست آزمون شونده α_i با الگوی صلاحیت α_i در سؤال j ام، به وسیله معادله زیر تعریف می‌شود:

$$p_j(\alpha_i) = P(X_j = 1 | \alpha_i) = (1 - s_j)^{\eta_{ij}} g_j^{1 - \eta_{ij}}$$

در این معادله، P احتمال، $p_j(\alpha_i)$ احتمال پاسخ درست برای سؤال j ام با الگوی صلاحیت α_i و X_{ij} پاسخ مشاهده شده فرد α_i برای سؤال j ام است (دی‌لتوره و داگلاس^{۹۲}، ۲۰۰۴). با توجه به سنجش صلاحیت‌های چندگانه در مدل‌های شناختی تشخیصی، شاخص تمیز هر سؤال در این مدل‌ها، نشان می‌دهد که «چقدر یک سؤال به خوبی بین افرادی که به صلاحیت‌های بیشتری تسلط دارند و افرادی که به صلاحیت‌های کمتری تسلط دارند، تمایز ایجاد می‌کند» (راب، و همکاران، ۲۰۱۰، ص. ۲۸۱). در مدل دینا، شاخص تمیز کلی برای هر سؤال، از طریق محاسبه اختلاف احتمال پاسخ درست به یک سؤال برای آزمودنی‌هایی که به همه صلاحیت‌ها تسلط یافته‌اند و احتمال پاسخ درست به یک سؤال برای آزمودنی‌هایی که به همه صلاحیت‌ها تسلط نیافته‌اند، محاسبه می‌شود. بنابراین:

$$d_{i,DINA} = (1 - s_i) - g_i$$

نتایج

با استفاده از رویکرد مبتنی بر غیرجبرانی بودن ساختار ماتریس کیو، داده‌های آزمون ۲۰ سؤالی براساس مدل دینا و نرم‌افزار آر^{۹۴} تحلیل و پارامترها، برآورد شدند^{۹۵}. جدول (۳)، دو پارامتر حدس و لغزش و خطای استاندارد را برای ۲۰ سؤال آزمون نشان می‌دهد.

جدول ۳. پارامترهای حدس و لغزش برای مدل دینا

سؤال	حدس	خطای حدس	لغزش	خطای لغزش
۱	۰/۹۲	۰/۰۲	۰/۰۲	۰/۰۲
۲	۰/۵۸	۰/۰۳	۰/۰۷	۰/۰۲
۳	۰/۵۱	۰/۰۳	۰/۲۴	۰/۰۲
۴	۰/۳	۰/۰۲	۰/۳۶	۰/۰۸
۵	۰/۰۱	۰/۰۱	۰/۶۳	۰/۰۳
۶	۰/۶۳	۰/۰۳	۰/۰۹	۰/۰۳
۷	۰/۲	۰/۰۴	۰/۲۸	۰/۰۲
۸	۰/۱۱	۰/۰۱	۰/۲۸	۰/۰۶
۹	۰/۸	۰/۰۳	۰/۰۳	۰/۰۱
۱۰	۰/۰۱	۰/۰۱	۰/۶	۰/۰۳
۱۱	۰/۸۱	۰/۰۳	۰/۰۷	۰/۰۱
۱۲	۰/۳۵	۰/۰۲	۰/۱۴	۰/۰۴
۱۳	۰/۱۷	۰/۰۱	۰/۲۵	۰/۰۶
۱۴	۰/۰۲	۰/۰۱	۰/۷۶	۰/۰۳
۱۵	۰/۱	۰/۰۲	۰/۶۷	۰/۰۳
۱۶	۰/۲۷	۰/۰۲	۰/۱۵	۰/۰۳
۱۷	۰/۴	۰/۰۴	۰/۱۴	۰/۰۲
۱۸	۰/۱۳	۰/۰۳	۰/۴۴	۰/۰۴
۱۹	۰/۲	۰/۰۲	۰/۴۶	۰/۰۹
۲۰	۰/۲۲	۰/۰۲	۰/۴۳	۰/۰۳
میانگین	۰/۳۴		۰/۳	

بر اساس جدول (۳)، متوسط مقادیر پارامترهای حدس و لغزش به ترتیب برابر ۰/۳۴ و ۰/۳ است. این بدین معناست که دانش‌آموزانی که به همه صلاحیت‌های لازم برای یک سؤال تسلط نداشتند، به طور متوسط ۳۴ درصد شانس پاسخگویی صحیح به آن سؤال برایشان وجود داشت. همچنین، دانش‌آموزانی که به همه صلاحیت‌های لازم برای یک سؤال تسلط داشتند، به طور متوسط ۳۰ درصد شانس پاسخگویی غلط به سؤال، برای آن سؤال، برای ایشان وجود داشت. از نظر راپ و همکاران (۲۰۱۰)، هر اندازه مقادیر حدس و لغزش در سؤالی «کوچک» باشد، آن سؤال حاوی اطلاعات «مفید»تری است. البته قاعده دقیقی برای این که مقادیر حدس و لغزش «کوچک» تلقی شوند، وجود ندارد. به عبارت دیگر، پارامترهای حدس و لغزش «کوچک»، نشان‌دهنده برآزش مناسب بین ساختار سنجش شناختی، داده‌ها و مدل دینا است. سؤال‌های ۱، ۲، ۳، ۶، ۹ و ۱۱ (که همگی چندگزینه‌ای هستند)، پارامترهای حدس بالاتری نسبت به بقیه سؤال‌ها دارند (۰/۵ >). همچنین مطابق جدول (۱) در

سؤال‌های مذکور، درصد پاسخگویی درست تقریباً بین ۷۵ درصد تا ۹۵ درصد است که به نظر می‌رسد این سؤال‌ها، برای دانش‌آموزان، سؤال‌هایی سخت محسوب نمی‌شوند. از طرفی، سؤال‌های ۵، ۱۰، ۱۴ و ۱۵ (که همگی پاسخ ساختنی هستند)، پارامترهای لغزش بالاتری نسبت به بقیه سؤال‌ها دارند (۰/۵ >). بررسی درصد پاسخگویی درست این مجموعه سؤال‌ها در جدول (۱) نشان می‌دهد که تقریباً بین ۱۱ درصد تا ۱۹ درصد دانش‌آموزان به این سؤال‌ها، پاسخ درست داده‌اند که به نظر سؤال‌های آسانی محسوب نمی‌شوند.

از آنجا که با داشتن K ، صلاحیت شناختی دو ارزشی، تعداد کلاس‌های مکنون (الگوهای صلاحیت) برابر 2^K است، در مطالعه حاضر، با داشتن چهار صلاحیت شناختی، تعداد کلاس‌های مکنون برابر ۱۶ است. به عنوان نمونه در کلاس ششم الگوی صلاحیت (۱۱۰۰) نشان‌دهنده تسلط بر دو صلاحیت شناختی ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی و ریاضی وار کردن و بازنمایی و عدم تسلط بر دو صلاحیت شناختی استدلال و طراحی راهبرد است. مقادیر جدول (۴) احتمال پسین^{۹۶} را برای چهار دانش‌آموز در هر یک از ۱۶ کلاس مکنون نشان می‌دهد.

جدول ۴. احتمال پسین برای چهار دانش‌آموز در کلاس‌های مکنون

کلاس مکنون	الگوی صلاحیت	دانش آموز ۱	دانش آموز ۲	دانش آموز ۳	دانش آموز ۴
۱	۰۰۰۰	۰/۰۲۶	۰/۰۳۹	۰/۱۲۴	۰/۰۰۰
۲	۱۰۰۰	۰/۰۰۳	۰/۰۰۱	۰/۰۰۴	۰/۰۰۰
۳	۰۱۰۰	۰/۰۲۶	۰/۰۳۹	۰/۰۱۲۴	۰/۰۰۰
۴	۰۰۱۰	۰/۰۲۶	۰/۰۳۹	۰/۰۱۲۴	۰/۰۰۰
۵	۰۰۰۱	۰/۰۲۶	۰/۰۲۵	۰/۰۱۲۴	۰/۰۰۰
۶	۱۱۰۰	۰/۰۵۸	۰/۰۰۰	۰/۰۱۳	۰/۰۰۰
۷	۱۰۱۰	۰/۰۰۳	۰/۰۰۲	۰/۰۰۰	۰/۰۰۰
۸	۱۰۰۱	۰/۰۰۵	۰/۰۲۹	۰/۰۰۷	۰/۰۰۰
۹	۰۱۱۰	۰/۰۲	۰/۰۳۹	۰/۰۹۳	۰/۰۰۰
۱۰	۰۱۰۱	۰/۰۲۶	۰/۰۳۹	۰/۱۲۴	۰/۰۰۰
۱۱	۰۰۱۱	۰/۰۲۶	۰/۰۰۵	۰/۱۲۴	۰/۰۰۰
۱۲	۱۱۱۰	۰/۶۴۴	۰/۶۶۸	۰/۰۰۵	۰/۳۶۲
۱۳	۱۱۰۱	۰/۰۴۹	۰/۰۰۱	۰/۰۳۵	۰/۰۰۷
۱۴	۱۰۱۱	۰/۰۲۲	۰/۰۲۹	۰/۰۰۲	۰/۰۰۹
۱۵	۰۱۱۱	۰/۰۲	۰/۰۰۴	۰/۰۹۳	۰/۰۰۰
۱۶	۱۱۱۱	۰/۰۱۷	۰/۵۷۴	۰/۰۰۱	۰/۶۱۷

سنجش تشخیصی صلاحیت‌های سواد ریاضی

به عنوان نمونه، برای دانش‌آموز شماره یک، احتمال تعلق فرد به کلاس مکنون دوازدهم بیشتر از همه است، پس می‌توان نتیجه گرفت که این فرد، به الگوی صلاحیتی (۱۱۱۰) تعلق دارد. یعنی دانش‌آموز شماره یک، ۶۴ درصد شانس تعلق به کلاس مکنون دوازدهم با الگوی صلاحیت (۱۱۱۰) را دارد. به تعبیر دیگر این فرد ۶۴ درصد احتمال دارد که به سه صلاحیت ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی، ریاضی وار کردن بازنمایی و استدلال دست یافته باشد.

جدول ۵. احتمال پسین برای چهار دانش‌آموز در کلاس‌های مکنون

الگوی صلاحیت	فراوانی مورد انتظار پاسخ‌دهندگان	درصد مورد انتظار پاسخ‌دهندگان
۰۰۰۰	۳۵/۴	۰/۰۵۳۸
۱۰۰۰	۱/۳	۰/۰۰۱۹
۰۱۰۰	۳۵/۴	۰/۰۵۳۸
۰۰۱۰	۳۵/۴	۰/۰۵۳۸
۰۰۰۱	۳۵/۴	۰/۰۵۳۸
۱۱۰۰	۱۲/۶	۰/۰۱۹۲
۱۰۱۰	۰/۵	۰/۰۰۰۸
۱۰۰۱	۲/۱	۰/۰۰۳۱
۰۱۱۰	۲۱	۰/۰۳۲۱
۰۱۰۱	۳۵/۴	۰/۰۵۳۸
۰۰۱۱	۳۵/۴	۰/۰۵۳۸
۱۱۱۰	۵۶/۶	۰/۰۸۶۲
۱۱۰۱	۵۵/۸	۰/۰۸۴۹
۱۰۱۱	۱۳/۷	۰/۰۲۱۱
۰۱۱۱	۲۱	۰/۰۳۲۱
۱۱۱۱	۲۶۰	۰/۳۹۶

بنابر یافته‌های جدول (۵) انتظار می‌رود که تقریباً ۵/۴ درصد دانش‌آموزان در این مطالعه، به اولین کلاس مکنون یا الگوی صلاحیتی (۰۰۰۰) تعلق یابند. یعنی ۵/۴ درصد از آزمودنی‌ها (یا ۳۵ نفر) در این مطالعه، به هیچ یک از صلاحیت‌ها دست نیافته‌اند. همچنین، ۳۹/۶ درصد دانش‌آموزان (یا ۲۶۰ نفر)، به

۱۶ امین کلاس مکنون با الگوی صلاحیتی (۱۱۱۱) تعلق دارند یا به عبارتی ۳۹/۶ درصد دانش‌آموزان به همه صلاحیت‌ها دست یافته‌اند.

جدول ۶. شاخص‌های تمیز سؤال‌های آزمون

سؤال	حدس	لغزش	تمیز
۱	۰/۹۲	۰/۰۲	۰/۰۶
۲	۰/۵۶	۰/۰۷	۰/۳۷
۳	۰/۵۰	۰/۲۴	۰/۲۶
۴	۰/۳	۰/۳۶	۰/۳۴
۵	۰/۰۱	۰/۶۳	۰/۳۶
۶	۰/۶۴	۰/۰۹	۰/۲۷
۷	۰/۱۹	۰/۲۸	۰/۵۲
۸	۰/۱	۰/۲۸	۰/۶۲
۹	۰/۸	۰/۰۳	۰/۱۷
۱۰	۰/۰۱	۰/۶	۰/۳۹
۱۱	۰/۸	۰/۰۷	۰/۱۳
۱۲	۰/۳۵	۰/۱۴	۰/۵۱
۱۳	۰/۱۷	۰/۲۵	۰/۵۸
۱۴	۰/۰۲	۰/۷۶	۰/۲۲
۱۵	۰/۱	۰/۶۷	۰/۲۳
۱۶	۰/۳۱	۰/۱۵	۰/۵۴
۱۷	۰/۳۹	۰/۱۴	۰/۴۷
۱۸	۰/۱۲	۰/۴۴	۰/۴۴
۱۹	۰/۲۱	۰/۴۶	۰/۳۳
۲۰	۰/۲۲	۰/۴۳	۰/۳۵

باتوجه به جدول (۶)، هر چه مقادیر شاخص‌های حدس و لغزش کمتر باشد، مقادیر شاخص‌های تمیز بیشتر می‌شود، اما قاعده صریحی برای تعیین یک نقطه برش برای شاخص تمیز وجود ندارد (دی‌لتوره، ۲۰۱۲). مطابق این جدول، به دلیل بالا بودن مقادیر پارامتر حدس ($>0/8$) در سه سؤال ۱، ۹،

و ۱۱، شاخص تمیز این سؤال‌ها به ترتیب برابر $۰/۰۶$ ، $۰/۱۷$ و $۰/۱۳$ و دارای کمترین مقادیر شاخص تمیز هستند، در حالی که مقادیر شاخص تمیز برای سؤال‌های ۷، ۸، ۱۲، ۱۳ و ۱۶ بیش از $۰/۵$ به دست آمده‌اند. ۱۲ سؤال دیگر دارای شاخص تمیز بین $۰/۲$ تا $۰/۵$ بودند.

یکی از محاسن مدل‌های شناختی این است که برای هر دانش‌آموز احتمال تسلط بر هر یک از صلاحیت‌ها تعیین می‌شود. هوبنر (۲۰۱۰) استدلال می‌کند که چون کلاس‌های مکنون دوه‌دو ناسازگار و جامع^{۹۷} هستند، می‌توان احتمال کلاس‌های مکنون مرتبط با هر صلاحیت را جمع کرد. به عنوان مثال، در جدول (۵) برای دانش‌آموز شماره یک، احتمال تسلط این فرد به صلاحیت اول (یعنی ارتباطات) برابر مجموع احتمال‌های پسین این فرد در کلاس‌های مرتبط با صلاحیت اول است. براساس جدول (۵)، صلاحیت اول با کلاس‌های مکنون دوم، ششم، هفتم، هشتم، دوازدهم، سیزدهم، چهاردهم و شانزدهم مرتبط است. اکنون مجموع احتمال پسین دانش‌آموز شماره یک برای صلاحیت اول برابر $۰/۸$ است. پس ۸۰ درصد احتمال دارد که دانش‌آموز شماره یک، به صلاحیت ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی تسلط یافته باشد. افزون بر این، صلاحیت دوم با کلاس‌های مکنون سوم، ششم، نهم، دهم، دوازدهم، سیزدهم، پانزدهم و شانزدهم مرتبط است. بنابراین، مجموع احتمال پسین دانش‌آموز شماره یک برای صلاحیت دوم برابر $۰/۸۶$ است. یعنی ۸۶ درصد احتمال دارد که دانش‌آموز شماره یک، به صلاحیت ریاضی وار کردن و بازنمایی تسلط یافته باشد.

■ بحث و نتیجه‌گیری ■

در دهه اخیر، رویکرد سنجش شناختی تشخیصی به‌منظور تلفیق نظریه‌های شناختی با یادگیری و ارائه نمرخی از صلاحیت‌های شناختی دانش‌آموزان برخلاف روش‌های مرسوم اندازه‌گیری آموزشی مانند نظریه کلاسیک و نظریه سؤال پاسخ، جایگاه ویژه‌ای پیدا کرده است. علت این امر این است که نتایج تحلیل‌های آزمون شناختی تشخیصی، می‌تواند حاوی بینش‌های جدیدی برای درک معلمان از صلاحیت‌های شناختی دانش‌آموزان در حل مسائل کاربردی باشد. براین اساس با عنایت به چنین دیدگاهی، آزمونی به‌منظور بررسی صلاحیت‌های شناختی سواد ریاضی دانش‌آموزان ایرانی پایه دهم اجرا گردید.

از نتایج مهم این پژوهش، بالا بودن شاخص تمیز پنج سؤال مربوط به سه واحد کوهنوردی، کدام ماشین و کشتی دریانوردی بود که همگی دارای مقادیری بالاتر از $۰/۵$ بودند. یعنی این سؤال‌ها، بیشترین اطلاعات تشخیصی را درباره آزمودنی‌هایی که به اکثر صلاحیت‌ها تسلط یافته بودند و آزمودنی‌هایی که به صلاحیت‌های کمتری تسلط یافته بودند، فراهم نمود. همچنین، ۱۲ سؤال دارای شاخص تمیزی بین $۰/۲$ تا

۰/۵ بودند. تنها سه سؤال از سه واحد فروش گوشی تلفن همراه، سوله و کدام ماشین، به دلیل بالا بودن پارامتر حدس، دارای شاخص تمیز پائینی بودند. از این گذشته، نتایج تحلیل سؤال‌ها نشان داد که از بین ۲۰ سؤال این آزمون، در شش سؤال که همگی چندگزینه‌ای بودند، احتمال پاسخگویی درست به این سؤال‌ها برای دانش‌آموزانی که به همه صلاحیت‌های لازم برای یک سؤال تسلط نداشتند، بیش از ۰/۵ بود. بررسی ماتریس کیو این سؤال‌ها نشان داد که در همه آن‌ها، بین یک تا دو صلاحیت شناختی برای حل سؤال، مورد نیاز بود. علاوه بر این، درجه بندی سطوح صلاحیت‌های شناختی این سؤال‌ها براساس نظر متخصصان اغلب کم گزارش شده بود. با توجه به این که سؤال‌های مذکور سطوح بالایی از صلاحیت‌های شناختی را طلب نمی‌کردند، به این معنا که براساس یافته‌های ترنر (۲۰۱۲)، ترنر و آدامز (۲۰۱۲) و ترنر، دوسی، بلام و نیس (۲۰۱۳)، سطح دشواری سؤال‌ها پائین بود، احتمال پاسخگویی درست به چنین سؤال‌هایی افزایش یافته بود.

در بین سؤال‌های پاسخ‌ساختنی نیز، برای چهار سؤال، دانش‌آموزانی که به همه صلاحیت‌های لازم برای یک سؤال تسلط داشتند، بیش از ۰/۵ درصد احتمال پاسخگویی نادرست به سؤال برای آن‌ها وجود داشت. مجدداً در بررسی ماتریس کیو مشخص شد که همه این سؤال‌ها، تمام صلاحیت‌های شناختی را طلب می‌کنند. از طرف دیگر، سطوح صلاحیت‌های شناختی این سؤال‌ها براساس نظر متخصصان اغلب متوسط یا بالا گزارش شده بود. لذا، براساس یافته‌های ترنر (۲۰۱۲)، ترنر و آدامز (۲۰۱۲) و ترنر، دوسی، بلام و نیس (۲۰۱۳) می‌توان نتیجه گرفت که چنین سؤال‌هایی، دارای سطح دشواری بالاتری نسبت به بقیه سؤال‌ها بودند. همچنین، بنابر یافته‌های پژوهش رفیع‌پور و گویا (۱۳۸۹)، عادت داشتن دانش‌آموزان ایرانی به حل مسائل با بهره‌گیری از فرمول‌ها و کلیشه‌های خاص و مشکل در حل مسائلی که نیازمند تجزیه و تحلیل هستند، احتمال پاسخگویی نادرست به چنین سؤال‌هایی افزایش یافته بود.

بنا به نظر راب و تمپلین (۲۰۰۸ الف)، یک دلیل محتمل برای بالا بودن پارامترهای حدس و لغزش، شاید بد توصیف شدن ماتریس کیو^{۹۹} باشد. از نظر آن‌ها، حذف نادرست یک صلاحیت در ساختار ماتریس کیو برای یک سؤال، باعث بیش‌برآورد در پارامتر لغزش آن سؤال می‌شود، زیرا سؤال به ظاهر ساده‌تر جلوه می‌کند و در نتیجه، وقوع لغزش، محتمل‌تر می‌شود. در مقابل، گنجاندن نادرست یک صلاحیت در ساختار ماتریس کیو برای یک سؤال، باعث بیش‌برآورد^{۱۰۰} در پارامتر حدس آن سؤال می‌شود، زیرا سؤال به ظاهر سخت‌تر جلوه می‌کند و در نتیجه، وقوع حدس محتمل‌تر خواهد

بود. با توجه به این که در مطالعه حاضر ساختار ماتریس کیو در ابتدای این مطالعه چندسطحی بوده است و به منظور تحلیل داده‌ها و استفاده از مدل دینا به ماتریسی دو ارزشی تبدیل شده است، بنابراین، توصیف دقیق ساختار صلاحیتی‌های شناختی کمرنگ شده، زیرا در این ماتریس همه سطوح شناختی اعم از بالا، متوسط و کم در هم ادغام شده است. به عنوان نمونه براساس جدول (۲)، در مقایسه دو سؤال اول و پنجم آزمون، هر دو سؤال دارای صلاحیت شناختی ارتباطات و استفاده از زبان و عملیات نمادین، رسمی و فنی هستند، اما سطوح این صلاحیت در این دو سؤال از نظر متخصصان، متفاوت است، بنابراین در سؤال اول، بیش برآوردی از پارامتر حدس و در سؤال پنجم بیش برآوردی در پارامتر لغزش محتمل تر خواهد بود.

به طور کلی نتایج این پژوهش نشان داد که پایین بودن سطح صلاحیت‌های شناختی در یک سؤال، احتمال پاسخگویی درست به آن سؤال و نیز بالا بودن سطح صلاحیت‌های شناختی، احتمال پاسخگویی غلط به سؤال را افزایش می‌دهد. با توجه به تأکید نداشتن محتوای کتاب‌های درسی ایران به سؤال‌های زمینه‌مدار و کمرنگ بودن فعالیت‌های مدل‌سازی در این کتاب‌ها و گرایش دانش‌آموزان نوعی ایرانی به حل مسائلی که قبلاً با زمینه آن آشنا شده‌اند (رفیع‌پور و گویا، ۱۳۸۹)، رسیدن به چنین نتایجی دور از ذهن نبود. با توجه به نقش مهم صلاحیت‌های شناختی در به‌کارگیری دانش ریاضی برای حل مسائل زمینه‌مدار مدل‌سازی، پیشنهاد می‌شود آموزش چنین صلاحیت‌هایی در برنامه درسی ریاضی از طریق گنجاندن تکلیف‌ها و مسائل پاسخ‌ساختنی که ابعاد مختلف مدل‌سازی را در بر می‌گیرد و همچنین آموزش این صلاحیت‌ها به معلمان، لحاظ شود. به نظر می‌رسد پرداختن به چنین مسائلی منجر به تغییر نگرش دانش‌آموزان از تمایل به حل مسائل کلیشه‌ای به سوی گرایش به حل مسائل کاربردی با زمینه‌های مختلف و در نتیجه بهبود عملکرد آن‌ها در حل مسائل مدل‌سازی خواهد شد.

محدودیت‌های پژوهش

از محدودیت‌های مهم این پژوهش، تغییر سطوح صلاحیت شناختی از چندارزشی به دوارزشی و در نتیجه تغییر ساختار ماتریس کیو و کمرنگ شدن تفاوت بین سطوح صلاحیت‌های شناختی است. همچنین محدودیت دیگر این پژوهش، تغییر نمره‌گذاری چهار سؤال از چهار واحد نجار، سوله، کوهنوردی و خرید خانه از چندارزشی (۰، ۱، ۲) به دوارزشی (۰ و ۱) به منظور استفاده از مدل دینا بود. پیشنهاد می‌شود در مطالعات آتی، تحلیل ساختار دقیق سطوح صلاحیت‌های شناختی صورت گیرد.

پیوست (الف) راهنمای بررسی صلاحیت‌های شناختی

ارتباطات				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
ایجاد یک توصیف یا توضیح شفاف، منسجم، کامل و باصرفه، برای یک راه‌حل، فرایند یا استدلال	تفسیر اطلاعات استخراج شده، برقراری ارتباطات از طریق توضیح و توصیف منظم	شناسایی و استخراج اطلاعات حاصل شده از ارتباطات	آشکار بودن ارتباط بین اطلاعات موجود در مسئله	رمزگشایی و تفسیر اجزای سؤال، شامل تصورکردن موقعیت ارائه شده به منظور معنابخشیدن به اطلاعات فراهم شده

ریاضی‌وار کردن				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
تولید یک مدل و تبیین متغیرها، روابط و محدودیت‌های آن، ارزیابی یا مقایسه مدل‌ها	جرح و تعدیل یک مدل مفروض یا ایجاد مدلی با متغیرها، روابط و محدودیت‌های صریح	استنباط مستقیم از یک مدل مفروض	عدم نیاز به استفاده از مدل	استفاده کردن از مدل‌های مفروض یا ساخته شده

طراحی راهبرد برای حل مسئله				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
ساختن راهبرد تفصیلی برای یک راه حل جامع	ساختن راهبرد مورد نیاز	انتخاب راهبرد مناسب	انجام دادن عملیاتی بدون استفاده از راهبرد	انتخاب طراحی و همچنین اجرای یک راهبرد ریاضیاتی برای حل مسئله

استفاده از زبان و عملیات نمادین، رسمی و فنی				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
کاربرد چند مرحله‌ای رویه‌های ریاضیاتی رسمی و روابط جبری	استفاده از نمادها و فرمول‌ها	استفاده مستقیم از نمادهای ریاضی برای انجام محاسبات	انجام عملیات حسابی	درک کردن و استفاده از نمادها، قواعد و تعریف‌های رسمی

سنجش تشخیصی صلاحیت‌های سواد ریاضی

بازنمایی				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
فهمیدن و استفاده از یک بازنمایی غیراستاندارد که نیازمند رمزگشایی‌ها و تفسیرهای اساسی است	استفاده از دو یا چند بازنمایی و حرکت بین آن‌ها/جرح و تعدیل یک بازنمایی / طراحی یک بازنمایی ساده از یک موقعیت	انتخاب و تفسیر یک بازنمایی استاندارد	استفاده مستقیم از یک بازنمایی موجود	طراحی بازنمایی‌ها و حرکت بین بازنمایی‌های مختلف (بازنمایی‌ها تصویری از روابط یا ریاضیاتی هستند که شامل نمودارها، جداول، دیاگرام‌ها، تصاویر، معادلات، فرمول‌ها، توصیفات متنی و مواد عینی هستند)

استدلال				
سطح ۳	سطح ۲	سطح ۱	سطح ۰	تعریف متغیر
ارزشیابی و تولید زنجیره‌های استدلال برای حالت‌های کلی	تجزیه و تحلیل اطلاعات برای استدلال‌های چند مرحله‌ای	ترکیب اطلاعات برای استدلال‌های یک مرحله‌ای	اجرای مستقیم دستورالعمل‌های داده شده	فرایندهای فکری که عناصر مسئله را کشف کرده و به هم پیوند می‌زند

پیوست (ب)

۲۰۷۷۸	حجم پنج ناحیه
۱۹۰	تعداد کلاس‌ها در پنج ناحیه
۲۳/۳۴	میانگین دانش‌آموزان هر کلاس در پنج ناحیه
۲۵/۷۱	تعداد کلاس‌های تعیین شده در نمونه
۱۹	تعداد مدارس تعیین شده در نمونه

منابع

- اشمن، آدریان. اف. وکانوی، روبرت. ان. اف. (۱۳۸۴). مقدمه‌ای بر آموزش و پرورش شناختی، نظریه و کاربرد (ترجمه سیدکمال خرازی). تهران: انتشارات سنا.
- خرازی، کمال و دولتی، رمضان. (۱۳۸۸). راهنمای روان‌شناسی شناختی و علوم شناختی. تهران: انتشارات جهاد دانشگاهی - پژوهشکده علوم شناختی.
- رفیع پورگتایی، ابوالفضل و گویا، زهرا. (۱۳۸۹). ضرورت و جهت تغییرات در برنامه درسی ریاضی مدرسه‌ای در ایران از دیدگاه معلمان. فصلنامه نوآوری‌های آموزشی، ۹(۳۳)، ۹۱-۱۲۰.
- دبیرخانه طرح تولید برنامه درسی ملی. (۱۳۹۰). برنامه درسی ملی جمهوری اسلامی ایران (طرح جامع تحول بنیادین برنامه‌های درسی و تربیتی). نگاشت پنجم. سازمان پژوهش و برنامه‌ریزی آموزشی. وزارت آموزش و پرورش.
- غلام‌آزاد، سهیلا. (۱۳۹۱). رویکرد شناختی به آموزش ریاضی در دوره ابتدایی. فصلنامه مطالعات برنامه درسی ایران: ویژه مقالات برنامه درسی ریاضی، ۶(۲۴)، ۷-۳۲.
- 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 منتشر نشده دکتری در رشته سنجش و اندازه‌گیری). دانشکده روان‌شناسی و علوم تربیتی. دانشگاه علامه طباطبائی.
- Basoku, T., Ogretmen, T., Kelecioğlu, H. (2013). Model data fit comparison between DINA and G-DINA in cognitive diagnostic models. *Educational Journal*, 2(6), 256-262.
- Bradshaw, L., Izsák, A., Templin, J. and Jacobson E. (2013). Diagnosing Teachers' Understandings of Rational Numbers: Building a Multidimensional Test within the Diagnostic Classification Framework. *Educational Measurement: Issues and Practice*, 33 (1), 2-14.
- De la Torre, J., and Douglas (2004). Higher-order Latent Trait Models for Cognitive Diagnosis. *Psychometrika*, 69, 333-353
- De la Torre, J. (2009). DINA Model and Parameter Estimation: A Didactic. *Journal of Educational and Behavioral Statistics*, 34 (1), 115-130
- De la Torre, J. (2012). *Recent developments in the G-DINA model framework*. Paper presented at the V European Congress of Methodology, Santiago de Compostela, Spain.
- Gierl, M. J., Alves, C. & Majeau, R. T. (2010). *Using Principled Test Design to Develop and Evaluate a Diagnostic Mathematics Assessment in Grades 3 and 6*. Paper Presented at the Annual Meeting of the American Educational Research Association, Denver, CO, USA.
- 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. *The Journal of Technology, Learning, and Assessment*, 6 (6), 3-48. Retrieved from <http://www.jtla.org>
- Huebner, A. (2010). An overview of recent developments in cognitive diagnostic computer adaptive assessments. *Practical Assessment, Research & Evaluation*, 15, 1.
- Jang E. E. (2008). A Framework for Cognitive Diagnostic Assessment. In C. A. Chapelle, Y.R. Chung, & J. Xu (Eds.), *Towards adaptive CALL: Natural language processing for diagnostic language assessment* (pp. 117-131). Ames, IA: Iowa State University.
- Jensen, T. H. (2007). *ASSESSING MATHEMATICAL MODELING COMPETENCY*. Mathematical Modelling (ICTMA12) Education, Engineering and Economics.
- Lee, Y., Park, Y., and Taylan, D. (2010). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. national sample using the IMSS 2007. *International Journal of Testing*, 11, 144-177.
- Leighton, J. P. & Gierl, M. J. (2007). Why Cognitive Diagnostic Assessment. In J. P. Leighton & M. J. Gierl, (Eds) *Cognitive Diagnostic Assessment for Education: Theory and Applications* (pp. 3-18). New York: Cambridge University Press.
- Leighton, J. P., Gierl, M. J. & Hunka, S. M. (2004). The Attribute Hierarchy Method for Cognitive Assessment: A Variation on Tatsuoka's Rule-Space Approach. *Journal of Educational Measurement*, 41 (3), 205-237.
- Niss, M. Blum, W. & Galbraith, P. (2007). Introduction. In W. Blum, P. Galbraith, H. W. Henn and M. Niss (Eds.), *Modelling and applications in mathematics education, the 14th ICMI study* (pp. 3-32). New York: Springer.
- Niss, M. (2010). Modeling a Crucial Aspect of Students' Mathematical Modeling. In R. Lesh, P. L. Galbraith, C. R. Haines and A. Hurford (Eds), *Modeling Students' Mathematical Modeling Competencies*, (ICTMA 13) (pp. 43-60). New York: Springer.

سنجش تشخیصی صلاحیت‌های سواد ریاضی

- Niss, M & Højgaard, T. (2011). *Competencies and Mathematical Learning, Ideas and inspiration for the development of mathematics teaching and learning in Denmark* (Tekster fra IMFUFA, no 485). Roskilde: Roskilde University, IMFUFA.
- OECD (2000). *Measuring Student Knowledge and Skills: The PISA 2000 Assessment of Reading, Mathematical and Scientific Literacy*. Retrieved from <http://www.oecd.org/>
- OECD (2009). *Learning Mathematics for Life: A Perspective from PISA*. Paris: OECD Publications.
- OECD (2013). *PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy*. Paris: OECD Publications.
- Roberts, M. Gierl, M. (2011). *Developing and evaluating score reports for a diagnostic mathematics assessment*. Paper Presented in Symposium titled “Communicating Assessment Results to Particular Audiences” Annual meeting of the American Educational Research Association New Orleans, LA.
- Rupp, A. A., & Templin, J. (2008a). The Effects of Q-matrix Misspecification on Parameter Estimates and Classification Accuracy in the DINA Model. *Educational and Psychological Measurement*, 68, 78–98.
- Rupp, A.A. & Templin, J. (2008b). Unique characteristics of diagnostic classification models: A comprehensive review of the current state-of-the-art. *Measurement: Interdisciplinary Research and Perspectives*, 6 (4), 219-262.
- Rupp, A.A., Templin, J. & Henson, R.A. (2010). *Diagnostic Measurement, Theory, Methods, and Applications*. New York: The Guilford Press.
- Shu, Z., Henson, R. and Willse, J. (2013). Using Neural Network Analysis to Define Methods of DINA Model Estimation for Small Sample Sizes. *Journal of Classification*, 30, 173-194.
- Snow, R.E., & Lohman, D.F. (1989). Implication of cognitive psychology for education measurement. In R.L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 263–331). New York: Macmillan.
- Stacey, K. (2012). *The International Assessment of Mathematical literacy: PISA 2012 Framework and Items*. 12th International Congress on Mathematical Education. 8 July – 15 July, 2012, COEX, Seoul, Korea.
- Templin, J. & Henson, R.A. (2006). Measurement of Psychological Disorders Using Cognitive Diagnosis Models. *Psychological Methods*, 11(3), 287–305.
- Turner, R. (2010). *Identifying cognitive processes important to mathematics learning but often overlooked*. ACER Research Conference.
- Turner, R. and Adams, R.J. (2012). *Some drivers of test item difficulty in mathematics: an analysis of the competency rubric*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA), 13-17 April 2012, Vancouver, <http://research.acer.edu.au/pisa/7/>.
- Turner, R., Dossey, J., Blum, W. and Niss, M. (2013). Using Mathematical Competencies to Predict Item Difficulty in PISA: A MEG Study. In M. Prenzel, M. Kobarg, K. Schöps and S. Rönnebeck (ED), *Research on PISA, Research Outcomes of the PISA Research Conference 2009* (pp. 23–37). New York: Springer Dordrecht Heidelberg.

پی‌نوشت‌ها

14. Employ
15. Interpret
16. Formulation
17. Stacey
18. Højgaard
19. Assessment Used for Learning and as Learning Process
20. Assessment of Learning Outcomes
21. Jang
22. Ashman & Conway
23. علت انتخاب سن ۱۵ سالگی این است که در اغلب کشورهای عضو سازمان همکاری و توسعه اقتصادی، سن ۱۵ سالگی پایان آموزش اجباری است.
24. Mathematical Process
25. Cognitive Competencies

[۱] این مقاله، مستخرج از رساله دکتری نویسنده اول است.

1. Mathematical Modelling Competency
2. Jensen
3. Niss, Blum and Galbraith
4. Assumptions
5. Cognitive Competencies
6. Turner, Dossey, Blum and Niss
7. Cognitive Diagnostic Measurement
8. Stakeholders
9. Rupp, Templin & Henson
10. Gierl, Wang & Zhou
11. Programme for International Students Assessment (PISA)
12. Mathematics Literacy
13. Formulate

67. Roberts & Gierl
 68. Bradshaw, Iz's'ak, Templin and Jacobson
 69. Attribute Hierarchy Method
 70. Scholastic Aptitude Test (SAT)
 71. Basic Arithmetic Operation Skills
 72. Knowledge about the Property of Factors
 73. Applying the Rules of Factoring
 74. Hierarchy Consistency Index
 ۷۵. صلاحیت استفاده از ابزار و وسایل ریاضی، مورد نظر پژوهشگران این پژوهش نبود.
 76. Rubric
 ۷۷. برای نهایی کردن سؤال‌ها، توصیه راب و تمپلین (۲۰۰۸ ب) مبنی بر تعداد سؤال‌ها در یک آزمون شناختی تشخیصی متعادل که بین ۲۰ تا ۴۰ سؤال است، در نظر گرفته شد و مجموعاً ۲۰ سؤال انتخاب شد.
 ۷۸. مراحل طراحی آزمون و چگونگی انتخاب سؤال‌های نهایی، در مقاله دیگری آمده است.
 79. Nonidentifiability
 80. Shu, Henson and Willse
 81. Diagnostic Measurement, Theory, Methods, and Applications. (2010)..Rupp, A.A., Templin, J. & Henson, R.A.
 82. Probability Proportional to Size (PPS)
 83. Deterministic Inputs, Noisy And-Gate (DINA)
 84. Junker and Sijtsma
 85. Parsimonious
 86. Deterministic
 87. Conjuuctive
 88. Stochastic
 89. De La Torre
 90. Slipping
 92. Guessing
 93. Douglas
 94. R Software
 ۹۵. آر، یک زبان برنامه‌نویسی و محیط نرم‌افزاری برای محاسبات آماری و تحلیل داده از جمله مدل‌سازی خطی و غیرخطی، آزمون‌های کلاسیک آماری، تحلیل سری‌های زمانی، خوشه‌بندی و غیره است، که بر اساس زبان‌های اس و اسکیم پیاده‌سازی شده است. این نرم‌افزار با متن، به طور رایگان از طریق سایت اصلی آن یعنی www.r-project.org قابل دانلود است. امکان توسعه قابلیت‌های آر، با افزودن بسته‌های ایجادشده توسط کاربران آن، یکی از ویژگی‌های مهم این نرم‌افزار است. در این پژوهش، از بسته‌ای تحت عنوان مدل‌سازی شناختی (Cognitive Diagnosis Modeling) که توسط روبیتز و همکارانش تهیه شده و در سایت <https://sites.google.com/site/alexanderrobitzsch/software> قابل دسترسی است، برای تحلیل در محیط آر، استفاده شده است.
 96. Posterior Probability
 97. Huebner
 98. Mutually Exclusive and Exhaustive
 99. Misspecification of the Q-Matrix
 100. Overestimation
 26. Content
 27. Context
 28. Competencies
 29. Communication
 30. Modelling
 31. Representation
 32. Thinking and Reasoning
 33. Argumentation
 34. Problem posing and solving
 35. Using symbolic, formal and technical language and operations
 36. Use of aids and tools
 37. Mathematising
 38. Reasoning and Argument
 39. Devising strategies for solving problems
 40. Using mathematical tools
 41. Turner
 42. Adams
 43. Change and Relationships
 44. Space and Shape
 45. Quantity
 46. Uncertainty
 ۴۷. در ادبیات سنجش شناختی تشخیصی از واژه خصیصه (Attribute) برای توصیف ویژگی‌های مکنون آزمودنی‌ها استفاده می‌شود. با توجه به این‌که در این پژوهش، صلاحیت‌های شناختی نقش ویژگی‌های مکنون دانش‌آموزان را ایفا می‌کنند، به منظور یکسان‌سازی واژه‌ها در این نوشتار به جای واژه خصیصه از صلاحیت استفاده می‌شود.
 48. Cognitive Psychology
 49. Leighton
 50. Hunka
 51. Snow & Lohman
 52. Educational Psychometric Measurement Models
 ۵۳. یک متغیر مکنون (Latent Variable)، متغیری آماری است که نمایانگر ویژگی مکنون غیر قابل مشاهده پاسخ‌دهندگان (صلاحیت) است.
 54. Corroboration
 55. Frederiksen, Glaser, Lesgold & Shafto
 56. Nichols
 57. Chipman & Brennan
 58. Diagnostic Classification Models
 ۵۹. در صورت داشتن یک فضای مکنون k بعدی (مجموعه k صلاحیتی)، که همگی دوازده‌گانه باشند، ۲^k کلاس مکنون (یا الگوی صلاحیت) وجود دارد. هر کلاس مکنون، برداری ترکیبی از مقادیر مختلف K صلاحیت است.
 60. Q-Matrix
 61. Compensatory
 62. Inductive Approach
 63. Reverse-engineered Approach
 ۶۴. در ادبیات سنجش شناختی تشخیصی، از واژه Retrofit نیز به جای مهندسی معکوس استفاده می‌شود.
 65. Lee, Park, and Taylan
 66. Basokcu, Ogretmen, & Kelecioğlu