

مدل‌سازی تشخیصی شناختی (CDM) ریاضیات پایه اول دبیرستان

افشین افزالی^۱، علی دلاور^۲، محمدرضا فلسفی نژاد^۳، احمد برجعلی^۴

تاریخ دریافت: ۹۵/۰۲/۰۶

تاریخ پذیرش: ۹۵/۱۰/۱۵

چکیده

مدل‌های تشخیصی شناختی (CDM) مدل‌های متغیر مکنون چندبعدی تأییدی با ساختاری پیچیده هستند. در این پژوهش از این مدل‌ها برای بررسی وضعیت دانش‌آموزان پایه اول دبیرستان در درس ریاضیات استفاده شد. سنجش تشخیصی شناختی بر اساس ۸ صفت اصلی، شامل ۳۲ سؤال بر روی نمونه‌ای به حجم ۵۰۹ دانش‌آموز که بر اساس روش نمونه‌گیری خوشه‌ای چندمرحله‌ای از بین دانش‌آموزان شهر تهران انتخاب شده بودند، اجرا گردید. از مدل‌های IRT برای تعیین ویژگی‌های روان‌سنجی سؤالات استفاده شد. تجزیه و تحلیل داده‌ها با به‌کارگیری مدل DINA در مدل‌سازی تشخیصی شناختی ریاضیات نشان داد که هشت صفت زیربنایی تبیین‌کننده عملکرد ریاضی دانش‌آموزان پایه اول دبیرستان است. بر اساس نتایج پژوهش مشخص گردید که آزمودنی‌ها تنها در یک صفت از هشت صفت به حد تسلط (۰/۷) دست یافته‌اند و کمترین میزان تسلط مربوط به مهارت‌های درک مفهوم تعاریف (۰/۴۹۴)، عملیات پیشرفته ریاضی (۰/۴۹۸) و کاربرد آموخته‌ها در مسائل واقعی (۰/۵۰۵) است.

واژگان کلیدی: مدل‌های تشخیصی شناختی، مدل DINA، ریاضیات، دبیرستان

۱. استادیار گروه علوم تربیتی، دانشکده علوم انسانی، دانشگاه بوعلی سینا، (نویسنده مسئول)

afzali.afshin@yahoo.com

۲. استاد گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی

۳. دانشیار گروه سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی

۴. دانشیار گروه روانشناسی، دانشگاه علامه طباطبائی

مقدمه

ریاضیات یکی از دروس اصلی و مهم برنامه‌های آموزشی در کشورهای پیشرفته است تا آنجا که در اغلب کشورها موفقیت در ریاضیات پیش شرطی برای ورود به مقاطع تحصیلی بالاتر است. هم‌چنین بسیاری از صاحب‌نظران آموزش معتقدند که لازمه زندگی در جهان پیشرفته و پیچیده امروز، برخورداری از تفکر خلاق و اندیشه پویا و مولد است و فراگیری ریاضی می‌تواند به شکل‌گیری و رشد این تفکر کمک کند (رضویه و همکاران، ۱۳۸۴).

باین حال بررسی‌های انجام‌شده ملی و بین‌المللی در دوره‌های مختلف آموزشی حاکی از مشکلات زیاد دانش‌آموزان ایرانی در یادگیری این درس است. سومین مطالعه جهانی تیمز در سال ۱۳۷۴ و مطالعه مجدد آن در سال ۱۳۷۸ عملکرد ضعیف دانش‌آموزان ایرانی در حوزه ریاضیات را نشان داد. به‌نحوی که ایران از بین ۴۱ کشور شرکت‌کننده رتبه‌های ۳۷ و ۳۸ را کسب نمود (کیامنش و نوری، ۱۳۷۶) همچنین در بررسی اخیر در سال‌های ۲۰۰۷ و ۲۰۰۱ نیز دانش‌آموزان ایرانی رتبه‌ای بهتر از ۵۹ و ۴۲ کسب نمودند (کریمی و همکاران، ۱۳۹۱).

دشواری‌های یادگیری ریاضی در دانش‌آموزان ایرانی علل متعدد و متنوعی دارد، سلسله‌ای و همکاران (۱۳۸۴)، مشکلات برنامه درسی را به‌عنوان اصلی‌ترین عامل افت تحصیلی در این درس تعیین کردند، اصلی‌ترین اشکال در برنامه‌های درسی نیز اشکال در روش‌های یادگیری - یاددهی ذکر شده است، همچنین بر اساس نتایج پژوهش مذکور نارسایی‌های ارزشیابی از آموخته‌های دانش‌آموزان، عدم توجه به ارزشیابی تشخیصی و نداشتن الگوی مناسب ارزشیابی نیز از عوامل مؤثر در عدم موفقیت دانش‌آموزان بوده‌اند.

سنجش میزان فهم دانش‌آموز، از ریاضیات نیازمند ابزارهای اندازه‌گیری جایگزینی است که ادراک مفهومی را علاوه بر تسلط ورزی^۱ مورد پرسش قرار دهد (وب ورمبرگ، ۱۹۹۲) فرآیندهای اندازه‌گیری که برای هر سنجش مورد استفاده قرار می‌گیرد، باید ماهیت

1. Procedural fluency

چگونگی مهم اشیاء، چگونگی نشان دادن و چگونگی تفسیر مشاهدات پاسخ‌های آزمون را در نظر بگیرید. (میسلوی، الموندو لوکاس، ۲۰۰۳)

لذا تدوین و ارزیابی یک مدل تشخیصی شناختی^۱ (CDM) برای ریاضیات پایه اول دبیرستان که بتواند ساختارهای دانش و فرآیندهای مهارت‌های خاص در دانش‌آموزان با هدف فراهم نمودن اطلاعات در رابطه با ضعف و قوت‌های آن‌ها مدل‌سازی نماید و همچنین اطلاعات معتبر آموزشی را فراهم نماید که با استفاده از آن‌ها معلم بتواند به‌طور مؤثر آموزش‌های بعدی و مکمل را برای پاسخ‌گویی به نیازهای یادگیری هر دانش‌آموز خاص طراحی نماید، هدف این پژوهش است.

در طول سالیان اخیر علاقه متخصصین روان‌سنجی در زمینه مدل‌های تشخیصی شناختی (CDMs)، افزایش یافته است. این مدل‌ها شامل مدل‌هایی آماری هستند که به پژوهشگران اجازه می‌دهند که به‌صورت تجربی فرضیات نسبتاً دقیقی را درباره طبیعت فرآیندهای پاسخ‌دهی که آزمودنی‌ها در زمان پاسخ‌گویی به یک سنجش یا سؤالات یک پرسشنامه استفاده می‌کنند، آزمون نمایند. اگر طرح جمع‌آوری داده‌ها و تئوری زیرساخت، درجه مناسبی از کفایت را دارا باشند، اطلاعات تجربی دقیقی درباره مؤلفه‌های ذهنی مداخله‌کننده در فرآیندهای پاسخ‌دهی و چگونگی تعامل میان این مؤلفه‌ها توسط این مدل‌ها قابل‌دستیابی است. همچنین، مدل‌های CDM یک طبقه‌بندی چند متغیره از آزمونی‌ها ایجاد می‌نماید که آن‌ها را بر اساس مؤلفه‌های ذهنی که زمینه‌ساز فرآیندهای پاسخ‌دهی آن‌هاست، نیمرخ‌بندی می‌نماید که بنابراین می‌تواند به‌منظور شناسایی مسیرهای اصلاحی برای شایستگی در همه مؤلفه‌ها، کمک‌کننده باشد. (راپ^۲، ۲۰۰۷)

استفاده از روش‌های CDM در سنجش مزایایی دارد که در سایر روش‌ها وجود ندارد، در درجه نخست این که اکثر مدل‌های دیگر، از جمله IRT، تک‌بعدی بودن آماری مجموعه داده‌ها را فرض نموده و آن را به‌عنوان شرط قبلی برای مدرج نمودن سؤالات و برآورد پارامترها نیاز دارند. در اکثر مدل‌ها، تک‌بعدی بودن به‌عنوان لازمه اصلی برای مشخص

1. Cognitive Diagnostic Models
2. Rupp

نمودن جایگاه آزمودنی‌ها در طی یک پیوستار فرضی، ضروری است. یکی از ویژگی‌های مهم CDM عدم نیاز به تک‌بعدی بودن است. به نظر می‌رسد تک‌بعدی بودن تا حدودی در بخش‌های تحصیلی مشکل‌ساز باشد چراکه تحقیقات نشان داده است که ابزارهای سنجش تحصیلی نوعاً مجموعه‌ای از صفات یا خرده مهارت‌ها را مورد توجه قرار می‌دهند که هر کدام از آن‌ها می‌تواند یک بعد آماری جداگانه را ایجاد نماید (آریا دوست، ۲۰۱۱).

هم‌چنین به‌عنوان یک از مدل‌های سؤال - پاسخ چندبعدی، CDM یکی از روش‌های طراحی شده در میان ارزشیابی‌های صفت‌مکون است که هم‌زمان هم از روش‌های ریاضی برای برآورد پارامترها و هم از اصول روان‌شناسی شناختی برای متمایز نمودن افراد مسلط و غیر مسلط استفاده می‌نماید. (گیرل، سیوی و ژو، ۲۰۰۹) مدل‌های تشخیصی شناختی از این جهت ساخته شده‌اند تا از طریق اطلاعات هدفمندتری که به شکل پروفایل‌های نمره ارائه می‌دهند، نقص و محدودیت مدل‌های IRT را برطرف نمایند (دی لا توره، ۲۰۰۹).

مدل‌های تشخیصی شناختی ((CDM مدل‌های متغیر مکون چندبعدی تأییدی احتمالاتی با یک ساختار بارگذاری پیچیده هستند. این مدل‌ها برای مدل‌سازی متغیرهایی با پاسخ طبقه‌ای مناسب هستند و شامل متغیرهای پیش‌بینی‌کننده مکون طبقه‌ای هستند که طبقات نهفته را تولید می‌نمایند. این مدل‌ها قادر به تولید تفسیرها و بازخوردهای چند ملاکی برای اهداف تشخیصی هستند که به یک تئوری زمینه‌ای - شناختی برای فرآیندهای پاسخ در اندازه ریز اشاره دارد (راپ، ۲۰۰۷).

شکل‌های متنوعی از مدل‌های تشخیصی شناختی در تاریخچه سنجش مطرح شده است. به‌طور کلی این مدل‌ها انواع موقعیت‌هایی (مثل انواع سازه، پاسخ و ابعاد) را پوشش می‌دهند که موردعلاقه پژوهشگران در روان‌سنجی و علوم شناختی و یادگیری است. در حال حاضر پژوهشگران به دلیل گستردگی این مدل‌ها در تلاش‌اند تا روابطی را بین این مدل‌ها بیابند و شباهت‌هایی را جهت یکسان‌سازی آن‌ها به دست بیاورند. این تلاش‌ها منجر به محبوبیت

فزاینده این مدل‌ها شد و نیز موجب تحولی در بین پژوهشگران شد تا مدل‌های CDM را برای محققان کاربردی ساده‌تر و امکان‌پذیرتر نمایند (یانگ سان^۱ و همکاران، ۲۰۱۱). به‌طور کلی دو روش برای استفاده از سنجش تشخیصی شناختی برای سنجش مهارت‌ها وجود دارد:

الف) تحلیل داده‌های حاصل از یک سنجش موجود با استفاده از مدل‌های پیچیده‌تر مهارت محور به امید استخراج اطلاعات بیش‌تر نسبت به مقیاس سازی‌های تک‌بعدی یا سایر روش‌های موجود.

ب) طراحی یک آزمون از ابتدا برای یک هدف تشخیص مهارت محور. در طراحی یک سنجش تشخیصی مؤثر برآورده کردن اهداف سنجشی، انگیزه اصلی است. در اغلب کارهایی که در زمینه‌ی مدل‌های تشخیصی انجام شده و در ادبیات پژوهش یافت می‌شود از سنجش تشخیصی به‌عنوان یک روش پس از وقوع (راه اول) استفاده شده است (دیلبو و دیگران، ۲۰۰۷).

در پژوهش حاضر از روش دوم استفاده شده و یک مدل تشخیصی شناختی با استفاده از نظرات کارشناسان ریاضی، شامل ۸ صفت سلسله‌مراتبی خطی، طراحی شد سپس بر اساس این مدل، یک سنجش تشخیصی شناختی (CDA) ساخته و بر روی گروه نمونه اجرا گردید. به‌منظور ارزیابی مدل و سنجش شناختی تشخیصی از مدل DINA^۲ استفاده شده است. مدل DINA یک از پرکاربردترین مدل‌های CDM در ادبیات این روش‌هاست. این مدل بدون توجه به تعداد صفت‌های بررسی شده در هر ارزیابی به‌تنهایی برای هر آیت‌م نیازمند به تخمین دو پارامتر است. بعلاوه زمانی که صفت‌های موردنیاز برای یک آیت‌م دارای اهمیت برابر و مساوی هستند، مدل DINA مناسب به نظر می‌رسد. مطالعات مختلفی مدل DINA را بررسی نموده‌اند. یک بحث کلی از مدل DINA که شامل کاربردها و مدل‌های طبقه‌بندی مرتبط و پنهانی آن است در کارهای دلاتوره و داگلاس^۳ (۲۰۰۸)، داگنون و فالماگن

1. Young-sun
2. deterministic input noisy and gate model
3. de la Torre & Douglas

(۱۹۹۹)، هر تل (۱۹۸۹)، جانکر و سیجسما^۱ (۲۰۰۱)، مک رییدی و دایتون (۱۹۷۷) و تاتسوکا (۲۰۰۲) قابل دستیابی است.

مدل DINA شامل دو بخش است: (۱) فرایند قطعی یا دور از خطا^۲ و (۲) فرایند تصادفی^۳. فرایند قطعی با بردار پاسخ پنهان (η_i) نشان داده می‌شود که با استفاده از بردار برآورد شده α_i محاسبه می‌شود. ارزش عددی ۱ به پاسخ پنهان نشان‌دهنده این است که آزمودنی دارای همه صفتهای موردنیاز جهت حل آیتم ز است درحالی که ارزش عددی صفر به معنای آن است که آزمودنی یکی از صفتهای موردنیاز برای آیتم را دارا نیست، نه اینکه فاقد همه صفتهای k باشد (دی لا تور ۲۰۰۹). همچنین دو پارامتر جنبه‌های مزاحم را مشخص می‌کنند: پارامترهای حدس^۴ و پارامترهای لغزش^۵. مدل DINA پارامتر لغزش را با عنوان $s_j = P(X_{ij} = 0 | \eta_{ij} = 1)$ و پارامتر حدس را با عنوان $g_j = P(X_{ij} = 1 | \eta_{ij} = 0)$ برای آیتم j مشخص می‌کند. دانش‌آموزانی که دارای همه صفتهای نیستند و گاهی اوقات حدس می‌زنند (پارامتر حدس) و پاسخ صحیح به یک آیتم می‌دهند و دانش‌آموزانی که دارای همه صفتهای هستند و گاهی لغزش می‌کنند (پارامتر لغزش) و غلط به یک آیتم پاسخ می‌دهند. مدل DINA با ترکیب این دو پارامتر این احتمال را محاسبه می‌کند که آزمودنی i چگونه آیتم j را حل می‌نماید.

از این رو ارائه پاسخ صحیح به یک سؤال نیازمند آن است که یک آزمودنی همه صفتهای لازم برای حل آیتم j را داشته باشد آن گونه که ماتریس Q با اجتناب از خطا آن را تعریف می‌کند و برای یک آزمودنی که فاقد صفت خاصی است حدس صحیحی بزند. همچنین اگر هیچ کدام از پارامترهای حدس و لغزش وجود نداشتند، پاسخ آزمودنی کاملاً قطعی می‌شود و نتایج بر تعامل بین مسیر α و بردار Q برای آیتم متکی است. (دی لاتور، ۲۰۰۹). تخمین تأثیرات هم پارامترهای حدس و هم پارامترهای لغزش می‌تواند منجر به

-
1. Junker & sijtsma
 2. deterministic
 3. stochastic
 4. guessing
 5. slipping

شناسایی یکسری روابط غیردقیق بین تسلط دانش آموز بر صفت موردنیاز و پاسخ‌دهی به یک آیتم گردد. در نتیجه مربیان و معلمان می‌توانند بفهمند که آیا یک صفت یا مهارت ویژه جهت دستیابی به یک سطح رتبه‌بندی معین بر یک دانش‌آموز یا گروهی از دانش‌آموزان تأثیر دارد.

گیرل و همکاران (۲۰۰۸) از روش سلسله‌مراتبی صفات برای بررسی سؤالات جبر در آزمون استعداد تحصیلی^۱ استفاده کردند. در این پژوهش ابتدا متخصصان مهارت‌های لازم برای حل سؤالات آزمون مذکور را استخراج و آن‌ها را بر اساس یک ارتباط سلسله‌مراتبی منظم کردند.

سپس نتایج تجربی با استفاده از شاخص ثبات سلسله‌مراتب مورد ارزیابی قرار گرفت. بر اساس نتایج متوسط این شاخص ۰/۸٪ بود که نشان‌دهنده برازش مدل سلسله‌مراتبی با داده‌های حاصل از ۵۰۰۰ دانش‌آموز شرکت‌کننده در SAT بود.

همچنین سو (۲۰۱۳) از مدل DINA با ساختار سلسله‌مراتبی برای بررسی نتایج آزمون تیمز (۲۰۰۳) استفاده کرد بر اساس نتایج این پژوهش مدل DINA با ساختار سلسله‌مراتبی نسبت به سایر مدل‌ها از برازش مناسب‌تری با داده‌های آزمون تیمز برخوردار است. سو (۲۰۱۳) استفاده از مدل DINA سلسله‌مراتبی به‌عنوان مکملی برای روش سلسله‌مراتبی صفات به دلیل محدودیت‌های این روش در برآورد پارامترهای سؤال و تغییرپذیری احتمال صفات بر اساس حجم نمونه را پیشنهاد کرده است.

لی و همکاران (۲۰۱۱) از مدل DINA برای مقایسه عملکرد دانش‌آموزان آمریکا در سؤال‌های ریاضی تیمز (۲۰۰۷) استفاده کردند. بر اساس مدل، آن‌ها ۱۵ صفت را برای این سؤالات شناسایی و تعیین کردند بر اساس نتایج پژوهش مذکور دانش‌آموزان آمریکایی در ۶ مهارت از ۱۵ مهارت به تسلط رسیده‌اند.

بایرن بام و همکاران (۲۰۰۴) نیز از مدل RUM برای مقایسه عملکرد دانش‌آموزان آمریکا، ژاپن و استرالیا استفاده کردند. بر اساس نتایج مطالعه فوق دانش‌آموزان ژاپنی در ۲۲

1. SAT

مهارت از ۲۳ مهارت به تسلط رسیده‌اند در حالی که دانش آموزان آمریکایی و اسرائیلی در ۱۷ صفت به حد تسلط رسیده‌اند.

ایم و پارک (۲۰۱۰) نیز از یک مدل تشخیصی شناختی، شامل ۱۰ مهارت برای بررسی نتایج تیمز (۲۰۰۳) با هدف مقایسه دانش آموزان کره‌ای جنوبی و آمریکا استفاده کرده‌اند. بر اساس نتایج این پژوهش ترجمه‌ی کلمات به معادلات، بهره‌گیری از اشکال و کردارها و احتمال برای هر دو گروه دشوار و صفات اعداد، به کارگیری مفاهیم و انجام محاسبات برای هر دو گروه ساده بوده است. در مجموع عملکرد دانش آموزان کره‌ای در هر ۱۰ مهارت بهتر از دانش آموزان آمریکایی بوده است.

در ایران پژوهش‌های چندانی در زمینه مدل‌های تشخیصی شناختی به‌ویژه در زمینه درس ریاضیات، انجام نشده است. تنها پژوهش انتشار یافته‌ای که در زمینه‌ای مشابه انجام شده است توسط مینایی (۱۳۹۱) با موضوع مدل پردازی تشخیصی شناختی سؤالات تیمز ۲۰۰۷ در دانش آموزان پایه هشتم ایران انجام شده است که بر اساس نتایج آن متوسط دانش آموزان ایرانی در هیچ‌یک از صفاتی که در مدل شناختی مشخص شده بود، به‌ویژه صفاتی که معطوف به کاربرد آموخته‌های ریاضیاتی بود، به حد تسلط نرسیده‌اند. در این پژوهش به‌منظور تحلیل نتایج از بسته CDM در نرم‌افزار R استفاده گردید.

روش

تدوین و ارزیابی مدل تشخیصی شناختی یادگیری ریاضیات پایه اول دبیرستان طی پنج مرحله به شرح زیر انجام شد:

الف: طراحی مدل شناختی و تدوین صفات زیربنایی. در طراحی مدل شناختی از روش گیرل و آلوز (۲۰۱۰) استفاده شد. بر اساس این الگو مدل‌ها توسط ۴ متخصص ریاضی با مدرک کارشناسی ارشد ریاضی و آموزش ریاضی و با سوابق ۱۴، ۱۲، ۸ و ۱۶ سال تدریس و تألیف طراحی شد. پژوهشگر با این چهار نفر به‌منظور تدوین مدل شناختی و همچنین سؤالات آزمون تشخیصی همکاری نمود.

بر اساس مدل پیشنهادی ۸ صفت به عنوان صفات زیر بنایی و اصلی برای ریاضیات پایه اول دبیرستان معرفی شدند. در جدول شماره یک، لیست هشت صفت و خرده مهارت‌های مرتبط با آن‌ها آمده است.

جدول ۱. مهارت‌ها و خرده مهارت‌های ریاضیات پایه اول دبیرستان

ردیف	مهارت	خرده مهارت
۱	عملیات پایه ریاضی (مهارتی)	<ul style="list-style-type: none"> - اولویت اعمال ریاضی و محاسبات ریاضی (جمع، تفریق، ضرب و تقسیم) - محاسبات ریاضی روی اعداد گویا (جمع، تفریق، ضرب و تقسیم) - ریشه گیری و توان - ویژگی طرفین وسطین تناسب‌ها
۲	درک مفهوم تعاریف (محتوایی)	<ul style="list-style-type: none"> - شیب - تعاریف مثلثاتی (tan , cos , sin) - معادله و جواب آن - اتحاد
۳	به خاطر سپاری تعاریف (محتوایی)	<ul style="list-style-type: none"> - فرمول شیب - تعاریف مثلثاتی (tan , cos , sin) - معادله - اتحاد
۴	استفاده از تعاریف (فرآیندی)	<ul style="list-style-type: none"> - معادله خط و کاربرد فرمول شیب - تعاریف مثلثاتی در مثلث قائم‌الزاویه - تشخیص معادله و حل آن - اتحاد مزدوج و دو جمله مشترک
۵	عملیات ریاضی پایه‌ای (مهارتی)	<ul style="list-style-type: none"> - عملیات جبری ساده ۱ (جمع و تفریق عدد ثابت به طرفین معادله) - آشنایی با نسبت‌های مثلثاتی در مثلث قائم‌الزاویه و درک ارتباط بین آن‌ها - به کمک قضیه فیثاغورث - عملیات جبری ساده ۲ (ضرب و تقسیم عدد ثابت به طرفین معادله) - تجزیه چندجمله‌ای‌ها با استفاده از اتحادهای یک جمله مشترک و مزدوج - و حل معادله درجه دوم
۶	عملیات ریاضی پیشرفته (مهارتی)	<ul style="list-style-type: none"> - درک ارتباط شیب دو خط موازی - استفاده از فرمول‌های اساسی و درک روابط مستقل از θ - جمع و تفریق یک عبارت جبری به طرفین معادله

		- عبارات گویا و ساده کردن آنها
		- درک مفهوم شرط تعلق یک نقطه به خط
۷	محاسبات عددی (فرآیندی)	- استفاده از نسبت های مثلثاتی معروف
		- درک مفهوم جواب معادله
		- کاربرد اتحاد مزدوج در محاسبات ریاضی
		- کسب مهارت برای تعیین شیب و استفاده از این مفهوم برای دنیای واقعی
		- حل مسائل با استفاده از مثلثات و درک کاربرد مثلثات در دنیای واقعی
۸	کاربرد (مهارتی)	- کسب مهارت تشکیل و سپس حل معادله درجه اول در دنیای واقعی
		- کسب مهارت حل معادله درجه دوم و استفاده از اتحاد یک جمله ای
		- مشترک برای حل مسائل واقعی

ب: طراحی سؤالات (سنجش شناختی تشخیصی). در گام دوم و در ادامه جلسات گروه به منظور اندازه گیری هر یک از صفات گام اول ۴ سؤال با سطوح دشواری متفاوت طراحی گردید و در مجموع یک آزمون شامل ۳۲ آیتم تدوین گردید.

ج: تدوین ماتریس Q. در این مرحله ماتریس Q به صورت یک ماتریس (k, p) که در آن k تعداد صفات و p تعداد تمام سؤالات شناختی ممکن است، ساخته شد. در ماتریس Q بر اساس مدل شناختی که از قبل طراحی شده چنانچه سؤالی به یک صفت مرتبط باشد به آن عدد ۱ در صورت عدم ارتباط عدد صفر اختصاص می یابد. در فایل ماتریس Q سؤالات در سطرها و صفات در ستون ها قرار می گیرند. ماتریس Q پژوهش حاضر به شکل زیر تدوین شد.

1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0
0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	0	0	1

ماتریس Q

د: اجرای CDA بر روی گروه نمونه. پس از نهایی سازی، سنجش شناختی تشخیصی بر روی گروه نمونه که شامل ۲۳۱ دانش‌آموز دختر و ۲۷۸ دانش‌آموز پسر پایه اول دبیرستان که در هشت مدرسه در مناطق مختلف شهر تهران مشغول به تحصیل بودند اجرا گردید.

ه: برآورد پارامترها. مدل DINA دو پارامتر سؤال دارد، لغزش (S_j) و حدس (g_j). «لغزش» به این احتمال اشاره دارد که یک آزمودنی تمام صفات موردنیاز برای یک سؤال را داشته باشد اما در پاسخ‌گویی صحیح به آن سؤال شکست بخورد. «حدس» به احتمال پاسخ‌دهی درست به سؤال در غیاب یک یا بیش از یک صفت موردنیاز به آن سؤال اشاره دارد؛ اما دو پارامتر سؤال، سایر پارامترهای مزاحم را نیز در برمی‌گیرند. این عوامل مزاحم، دلایل این را

که چرا آزمودنی‌هایی که در برخی صفات مورد نیاز تسلط ندارند، می‌توانند به سؤال پاسخ درست دهند و همچنین دلایل این را که آزمودنی‌هایی که در همه صفات مورد نیاز به تسلط رسیده‌اند، پاسخ درست سؤال را از دست می‌دهند، مخدوش می‌نمایند. به‌عنوان دو مثال از عوامل مزاحم معمول شامل بد تشخیصی در ماتریس Q و استفاده از راهبردهای جایگزین را می‌توان نام برد (سو، ۲۰۱۳).

ز: برآورد احتمال موفقیت در هر یک از صفات. در پایان احتمال موفقیت دانش‌آموزان در هشت صفت بر اساس برآورد بیشینه احتمال کناری برآورد گردید.

به‌منظور بررسی پایایی طبقه‌بندی از شاخص ثبات سلسله‌مراتب (HCI) با فرمول

$$HCI_i = 1 - \frac{\sum_{J \in S_{correct_i}} \sum_{g \in S_J} X_{ij}(1 - X_{ig})}{N_{C_i}}$$

استفاده گردید که متوسط مقدار این شاخص برای همه آزمودنی‌ها برابر با ۰/۴۸+ به دست آمد که در محدوده مثبت و مناسب است.

یافته‌ها

در ادامه نتایج مدل‌سازی تشخیصی شناختی بر اساس مدل DINA با استفاده از داده‌های تجربی ارائه شده است. پاسخ‌های ۵۰۹ دانش‌آموز پسر و دختر در سنجش تشخیصی شناختی ریاضیات پایه اول دبیرستان بر اساس ماتریس Q شامل ۸ صفت سلسله‌مراتبی که در بخش روش معرفی شده است، با استفاده از بسته CDM نرم‌افزار R تحلیل شده است.

جدول ۲. پارامترهای حدس و لغزش سؤالات

سؤال	حدس	خطای حدس	لغزش	خطای لغزش	سؤال	حدس	خطای حدس	لغزش	خطای لغزش
۱	۰/۸۳	۰/۰۰۷	۰/۰۴۹	۰/۰۰۹	۱۷	۰/۳۷۴	۰/۰۰۶	۰/۱۱۵	۰/۰۱۲
۲	۰/۷۹۵	۰/۰۰۷	۰/۰۳۶	۰/۰۰۸	۱۸	۰/۳۴	۰/۰۰۷	۰/۱۲۳	۰/۰۰۵
۳	۰/۳۷۸	۰/۰۰۸	۰/۱۶۶	۰/۰۰۷	۱۹	۰/۲۳۵	۰/۰۰۶	۰/۲۰۱	۰/۰۰۶
۴	۰/۴۷۴	۰/۰۰۹	۰/۰۹۱	۰/۰۰۷	۲۰	۰/۱۳۱	۰/۰۰۹	۰/۳۳۷	۰/۰۱
۵	۰/۴۲۲	۰/۰۱	۰/۱۵۲	۰/۰۰۷	۲۱	۰/۱۲۳	۰/۰۱	۰/۲۲	۰/۰۱۳
۶	۰/۳۵۲	۰/۰۰۶	۰/۲۷	۰/۰۰۹	۲۲	۰/۱۰۲	۰/۰۱	۰/۳۷۵	۰/۰۱۹
۷	۰/۵۵۱	۰/۰۰۹	۰/۱۷۸	۰/۰۰۸	۲۳	۰/۰۵۵	۰/۰۱	۰/۳۰۵	۰/۰۲۴
۸	۰/۲۶۶	۰/۰۰۹	۰/۳۹۴	۰/۰۰۹	۲۴	۰/۰۵۹	۰/۰۱	۰/۲۰۴	۰/۰۳۱
۹	۰/۸۴۷	۰/۰۰۹	۰/۰۴۸	۰/۰۰۸	۲۵	۰/۰۱۷	۰/۰۰۹	۰/۲۶۹	۰/۰۲۵
۱۰	۰/۵۴۱	۰/۰۰۸	۰/۱۱۵	۰/۰۰۸	۲۶	۰/۰۱۸	۰/۰۰۹	۰/۴۷۳	۰/۰۰۹
۱۱	۰/۴۴۳	۰/۰۰۹	۰/۱۱۲	۰/۰۱	۲۷	۰/۰۳۴	۰/۰۰۸	۰/۱۱۷	۰/۰۳۳
۱۲	۰/۷۶۷	۰/۰۰۸	۰/۰۳۱	۰/۰۰۹	۲۸	۰/۰۳۶	۰/۰۰۹	۰/۳۴۵	۰/۰۲۳
۱۳	۰/۷۱	۰/۰۱	۰/۰۳۹	۰/۰۱۳	۲۹	۰/۰۲۲	۰/۰۰۸	۰/۶۳۸	۰/۰۱۵
۱۴	۰/۶۴۱	۰/۰۰۸	۰/۰۷۲	۰/۰۱	۳۰	۰/۰۲۳	۰/۰۰۹	۰/۷۷	۰/۰۲۴
۱۵	۰/۴۹۴	۰/۰۰۹	۰/۰۹۶	۰/۰۱۲	۳۱	۰/۰۳۵	۰/۰۰۶	۰/۵۲۷	۰/۰۱۲
۱۶	۰/۴۱۶	۰/۰۱	۰/۱۳۶	۰/۰۰۸	۳۲	۰/۰۵۳	۰/۰۰۷	۰/۳۱۸	۰/۰۰۵

جدول فوق پارامترهای حدس و لغزش بر اساس مدل DINA را نشان می‌دهد، بر اساس اطلاعات مندرج در این جدول پایین‌ترین ضریب حدس مربوط به سؤال‌های ۲۵ و ۲۶ با اندازه‌های ۰/۰۱۷ و ۰/۰۱۸ و بالاترین ضریب حدس مربوط به سؤال‌های ۹ و ۱ با مقادیر ۰/۸۴۷ و ۰/۸۳ است. این ضرایب احتمال پاسخ‌گویی درست به سؤال برای دانش‌آموزانی است که به مهارت‌های موردنیاز برای پاسخ‌گویی به سؤال تسلط ندارند را نشان می‌دهد. همچنین پایین‌ترین مقدار لغزش مربوط به سؤالات ۱۲ و ۲ با مقادیر ۰/۰۳۱ و ۰/۰۳۶ است و بالاترین ضریب لغزش مربوط به سؤال‌های ۳۰ و ۲۹ با مقادیر ۰/۷۷ و ۰/۶۳۸ است. این ضریب نشان‌دهنده احتمال پاسخ‌گویی غلط به سؤال برای دانش‌آموزانی است که به مهارت‌های موردنیاز برای پاسخ‌گویی به سؤال تسلط دارند.

هرچه پارامترهای حدس و لغزشی در سؤال کوچک تر باشد، نشان دهنده برازش بهتر میان طرح سنجش تشخیصی، داده های تجربی و مدل شناختی است (راوند و همکاران، ۲۰۱۳).

جدول ۳. ضرایب تشخیص و RMSEA سؤالات

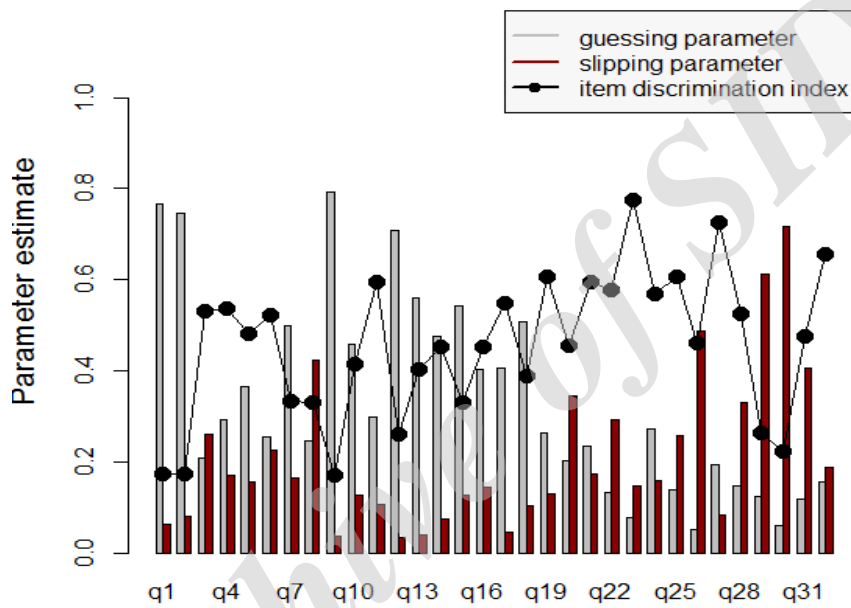
سؤال	IDI	RMSEA	سؤال	IDI	RMSEA
۱	۰/۱۲۱	۰/۰۲۸	۱۷	۰/۵۱۱	۰/۰۴۱
۲	۰/۱۷	۰/۰۷۳	۱۸	۰/۵۳۷	۰/۰۷۶
۳	۰/۴۵۶	۰/۰۵۷	۱۹	۰/۵۶۴	۰/۰۴۲
۴	۰/۴۳۶	۰/۰۳۹	۲۰	۰/۵۳۲	۰/۰۵۶
۵	۰/۴۲۶	۰/۱۰۵	۲۱	۰/۶۵۷	۰/۰۸۸
۶	۰/۳۷۸	۰/۰۴۷	۲۲	۰/۵۲۳	۰/۰۳۲
۷	۰/۲۷۱	۰/۰۴۱	۲۳	۰/۶۴	۰/۰۳۱
۸	۰/۳۳۹	۰/۰۲۵	۲۴	۰/۷۳۷	۰/۰۴۲
۹	۰/۱۰۵	۰/۰۲۸	۲۵	۰/۷۱۴	۰/۱۰۲
۱۰	۰/۳۴۴	۰/۰۳۴	۲۶	۰/۵۰۹	۰/۰۸
۱۱	۰/۴۴۵	۰/۰۳۷	۲۷	۰/۸۴۹	۰/۰۳۷
۱۲	۰/۲۰۲	۰/۰۳۱	۲۸	۰/۶۱۸	۰/۰۳۴
۱۳	۰/۲۵	۰/۰۴۴	۲۹	۰/۳۴۱	۰/۰۶۸
۱۴	۰/۲۸۷	۰/۰۲۵	۳۰	۰/۲۰۷	۰/۰۶۶
۱۵	۰/۴۱	۰/۰۲۹	۳۱	۰/۴۳۸	۰/۰۹
۱۶	۰/۴۴۹	۰/۰۵۷	۳۲	۰/۶۲۸	۰/۰۶۹

جدول شماره ۵ شاخص های تشخیص سؤال^۱ (IDI) و شاخص برازندگی سؤال (RMSEA) را ارائه می نماید. شاخص IDI رابطه معکوسی با شاخص های حدس و لغزش دارد و هرچه این شاخص کوچک تر باشند، شاخص بهتری تشخیص سؤال بالاتر است. شاخص RMSEA نیز میزان برازندگی سؤال با مدل شناختی را نشان می دهد، این شاخص بر اساس توزیع حتی دولت هرچه کوچک تر باشد، نشان دهنده برازش بالاتر سؤال با مدل است به صورت معمول سؤالات با شاخص RMSEA کوچک تر از ۰/۰۵ را سؤال

1. Item Discrimination Index

دارای بهترین برازندگی با مدل و سؤالات با RMSEA بزرگ‌تر از ۰/۱ را فاقد برآزش با مدل معرفی می‌نمایند.

بر اساس نتایج مندرج در جدول سؤالات ۲۷ و ۲۴ با شاخص‌های IDI برابر با ۰/۸۴۹ و ۰/۷۳۷ بهترین شاخص تشخیص و سؤالات ۱، ۱۴ و ۹ با شاخص $RMSEA$ برابر با ۰/۰۲۵ و ۰/۰۲۸ بیش‌ترین میزان برآزش با مدل را دارا هستند.



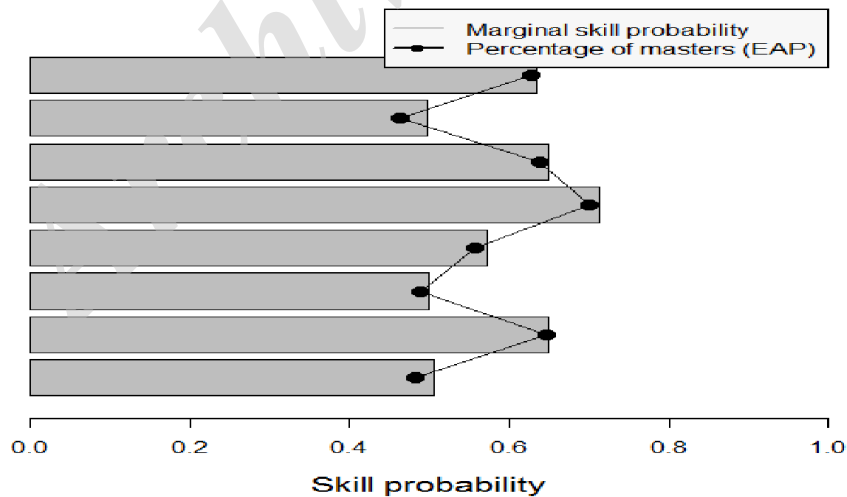
شکل ۲. پارامترهای حدس و لغزش و ضرایب تشخیص سؤالات

شکل ۲ پارامترهای حدس و لغزش و ضرایب تشخیص هر یک از ۳۲ سؤال را به صورت نمودار نمایش می‌دهد. همان‌گونه که در نمودار مشخص است هر چه حدس و لغزش کمتر باشند، ضریب تشخیص سؤال بالاتر است.

جدول ۴. احتمال تسلط به صفات هشت گانه در نمونه مورد مطالعه

صفت	عنوان	احتمال
صفت ۱	عملیات پایه ریاضی	۰/۶۳۴
صفت ۲	درک مفهوم تعاریف	۰/۴۹۴
صفت ۳	به خاطر سپاری تعاریف	۰/۶۴۹
صفت ۴	استفاده از تعاریف	۰/۷۱۴
صفت ۵	عملیات میانه ریاضی	۰/۵۷۲
صفت ۶	عملیات پیشرفته ریاضی	۰/۴۹۸
صفت ۷	محاسبات عددی	۰/۶۵۰
صفت ۸	کاربرد در مسائل واقعی	۰/۵۰۵

جدول شماره ۶ احتمال کناری تسلط به هر یک از مهارت های هشت گانه را نمایش می دهد با توجه به نتایج مندرج در جدول، آزمودنی ها بیشترین احتمال تسلط را به ترتیب در مهارت های استفاده از تعاریف (۰/۷۱۴) محاسبات عددی (۰/۶۵۰) و به خاطر سپاری تعاریف (۰/۶۴۹) و کمترین احتمال تسلط را به ترتیب در مهارت های درک مفهوم تعاریف (۰/۴۹۴) عملیات ریاضی پیشرفته (۰/۴۹۸) و کاربرد آموخته ها در مسائل دنیای واقعی (۰/۵۰۵) داشته اند.



شکل ۳. احتمال تسلط به مهارت های هشت گانه

شکل شماره ۳ احتمال هر یک از صفات را به صورت نمودار نمایش می‌دهد

جدول ۵. همبستگی تراکوریک صفات هشت‌گانه

صفت ۱	صفت ۲	صفت ۳	صفت ۴	صفت ۵	صفت ۶	صفت ۷	صفت ۸
صفت ۱	۰/۶۷۳۹	۰/۷۵۱۸	۰/۸۱۵۴	۰/۷۸۲۴	۰/۶۸۰۷	۰/۸۰۰۳	۰/۷۴۷۹
صفت ۲	۰/۶۷۳۹	۱	۰/۶۳۴۳	۰/۸۴۹۱	۰/۶۷۸۴	۰/۷۱۵۱	۰/۷۵۹۱
صفت ۳	۰/۷۵۱۸	۰/۶۳۴۳	۱	۰/۸۶۰۶	۰/۸۲۰۶	۰/۹۱۰۲	۰/۷۴۹۲
صفت ۴	۰/۸۱۵۴	۰/۸۴۹۱	۰/۸۶۰۶	۱	۰/۹۱۱۲	۰/۸۸۳۴	۰/۷۸۰۷
صفت ۵	۰/۷۸۲۴	۰/۶۷۸۴	۰/۸۲۰۶	۰/۹۱۱۲	۱	۰/۷۹۵۴	۰/۷۹۶۶
صفت ۶	۰/۶۸۰۷	۰/۷۶۶۳	۰/۸۰۹۴	۰/۷۴۴۷	۰/۷۵۳۷	۱	۰/۷۵۳۷
صفت ۷	۰/۸۰۰۳	۰/۷۱۵۱	۰/۹۱۰۲	۰/۸۸۳۴	۰/۷۹۵۴	۰/۸۱۵۸	۱
صفت ۸	۰/۷۴۷۹	۰/۷۵۹۱	۰/۷۴۹۲	۰/۷۸۰۷	۰/۷۹۶۶	۰/۷۵۳۷	۰/۸۳۶۷

جدول شماره ۵ همبستگی تراکوریک میان صفت‌ها را نشان می‌دهد همان‌گونه که در جدول مشخص است صفت‌های ۴ و ۵ دارای بالاترین همبستگی با مقدار ۰/۹۱۱ و صفت‌های ۱ و ۲ دارای کم‌ترین همبستگی با مقدار ۰/۶۷۳ هستند.

بحث و نتیجه‌گیری

در پژوهش حاضر از مدل‌سازی تشخیصی شناختی به‌عنوان روشی به‌منظور ساخت و طراحی یک ابزار و مدل تشخیصی شناختی برای درس ریاضیات پایه اول دبیرستان و همچنین بررسی احتمال دستیابی آزمودنی‌ها به تسلط در هر یک از صفات زیربنایی و اصلی این درس استفاده شد.

بر اساس روش سلسله‌مراتبی صفات (AHM)، هشت مهارت و صفت که در جدول شماره یک ذکر شده‌اند به‌عنوان صفات زیربنایی و اصلی با ارتباط سلسله‌مراتبی برای ریاضیات پایه اول دبیرستان تعیین گردید، سپس برای اندازه‌گیری هر صفت ۴ سؤال طراحی شد و یک سنجش شناختی تشخیصی شامل ۳۲ سؤال بر روی گروه نمونه اجرا گردید. به‌منظور ارزیابی سؤالات از روش‌های مبتنی بر تئوری سؤال پاسخ و مدل تشخیصی شناختی DINA استفاده شد. بر اساس نتایج IRT، سؤالات سنجش، دامنه تشخیص و دشواری

مناسبی را نشان دادند (ضریب تشخیص از ۰/۴۸ تا ۱/۵۸ و ضریب دشواری از ۳/۳۵- تا ۳/۴۶۷+). در بررسی بر اساس مدل DINA نیز پارامترهای حدس و دشواری سؤالات با میزان دشواری بر اساس IRT رابطه‌ی منطقی را نشان دادند، به عبارت دیگر سؤالات دشوار پارامتر حدس کوچک‌تر و لغزش بزرگ‌تر و سؤالات ساده پارامتر حدس بزرگ‌تر و لغزش کوچک‌تری دارند.

همچنین توانایی تشخیص سؤالات بر اساس شاخص IDI نیز بررسی شد که بر اساس نتایج آن سؤالات در مجموع توانایی تشخیص مناسبی میان آزمودنی‌هایی که در صفات به تسلط رسیده‌اند و آزمودنی‌هایی که به تسلط نرسیده‌اند را نشان دادند.

به منظور بررسی برازش سؤالات با مدل تشخیصی شناختی از شاخص RMSEA استفاده گردید. بر اساس این شاخص اغلب سؤالات غیر از سؤالات ۵ و ۲۵ (شاخص ۰/۱۰۵ و ۰/۱۰۲)، مقدار RMSEA کوچک‌تر از یک داشتند که نشان‌دهنده برازش مناسب آن‌ها با مدل است.

در پایان احتمال تسلط آزمودنی‌ها در صفات هشت‌گانه مورد بررسی قرار گرفت. بر اساس نتایج پژوهش، مندرج در جدول شماره پنج، تنها در صفت چهارم بر اساس نقطه برش ۰/۷ (مینایی، ۹۱) آزمودنی‌ها به حد قابل قبولی از تسلط رسیده‌اند. کمترین میزان تسلط در صفات درک مفهوم تعاریف (۰/۴۹۴)، عملیات پیشرفته ریاضی (۰/۴۹۸) و کاربرد آموخته‌ها در مسائل واقعی (۰/۵۰۵) بود.

با توجه به همبستگی‌های بالای میان صفات هشت‌گانه، مندرج در جدول شماره هفت، مشخص می‌شود که بخش قابل توجهی از عدم توانایی دانش‌آموزان در صفات مرتبه بالاتر ناشی از ضعف یا عدم یادگیری در مهارت‌های پایه و زیر بنایی است که به نوعی پیش‌نیاز دستیابی به موفقیت در ریاضیات پایه اول دبیرستان است. همچنین دانش‌آموزان پایه اول دبیرستان توفیق چندانی در کاربرد آموخته‌های خود برای حل مسائل دنیای واقعی نداشته‌اند. در ایران پژوهش‌های چندانی در زمینه مدل‌های تشخیصی شناختی به ویژه در زمینه درس ریاضیات، انجام نشده است که به توان نتایج این پژوهش را با آن‌ها مقایسه نمود، تنها پژوهش انتشار یافته‌ای که در زمینه‌ای مشابه انجام شده است توسط مینایی (۱۳۹۱) با موضوع مدل

پردازی تشخیصی شناختی سؤالات تیمز ۲۰۰۷ در دانش‌آموزان پایه هشتم ایران انجام شده است که بر اساس نتایج آن متوسط دانش‌آموزان ایرانی در هیچ‌یک از صفاتی که در مدل شناختی مشخص شده بود، به‌ویژه صفاتی که معطوف به کاربرد آموخته‌های ریاضیاتی بود، به حد تسلط نرسیده‌اند که نتایج این پژوهش با نتایج مینایی (۱۳۹۱) همسو است.

بر اساس نتایج این پژوهش و با توجه به مدل شناختی درس ریاضیات بررسی شده در آن، می‌توان به این نتیجه رسید که اغلب دانش‌آموزان ایرانی در مهارت‌های پایه‌ای و مبنایی درس ریاضیات دارای ضعف‌های عمده‌ای هستند و بخش قابل توجهی از عدم توفیق دانش‌آموزان در یادگیری مباحث سطح بالاتر و همچنین افت تحصیلی در این درس و در نتیجه کسب نتایج ضعیف در آزمون‌های ملی و بین‌المللی، ناشی از این ضعف است. لذا توجه به ساختار سلسله‌مراتبی این درس در زمان تدوین برنامه‌های آموزشی و پیروی از مدل شناختی متناسب با این ساختار در فرآیند یاددهی - یادگیری و همچنین برنامه‌ریزی جهت اطمینان حصول دانش‌پیش‌نیاز برای مفاهیم درسی و برنامه‌ریزی در راستای رفع مشکلات دانش‌آموزان در مهارت‌های مرتبه پایین‌تر قبل از آموزش مهارت‌های سطح بالا می‌تواند تأثیر قابل توجهی در رشد کیفی آموزش و یادگیری ریاضیات در دانش‌آموزان ایرانی داشته باشد.

منابع

- رضویه، اصغر. سیف، دیبا. طاهری، عبدالمحمد (۱۳۸۴). بررسی تأثیر مؤلفه‌های اضطراب و نگرش ریاضی بر پیشرفت تحصیلی دانش‌آموزان دبیرستانی در درس ریاضی. فصلنامه تعلیم و تربیت. شماره مسلسل ۸۲
- سلسبیلی، نادر. قاسمی، نرجس (۱۳۸۴). عوامل مؤثر بر شکست تحصیلی: نگاهی دوباره به عوامل درونی و بیرونی. تهران. فصلنامه تعلیم و تربیت. شماره مسلسل ۸۳
- کریمی، عبدالعظیم. بخشعلی زاده، شهرناز؛ و کبیری مسعود (۱۳۹۱). گزارش اجمالی از مهم‌ترین نتایج تیمز و پرلز ۲۰۱۱. تهران. پژوهشگاه مطالعات آموزش و پرورش
- کیامنش، علیرضا. نوری رحمان (۱۳۷۶). یافته‌های سومین مطالعه بین‌المللی تیمز. تک‌نگاشت ۲۲. پژوهشکده تعلیم و تربیت.
- مینایی، اصغر (۱۳۹۱). مدل پردازش تشخیصی شناختی سؤالات تیمز ۲۰۰۷ در دانش‌آموزان پایه هشتم ایران. پایان‌نامه دکترا. دانشگاه علامه طباطبایی
- ورما سورندرا (۱۳۸۰). برنامه درسی و چارچوب استانداردها. ترجمه نادرقلی قورچیان و حسن ملکی. تهران: انتشارات فراشناختی اندیشه.
- Aryadoust, V. (2011). Cogniittiive diagnostic assessmentt as an alltternattiive measurementt model. *JALT Testing & Evaluation SIG Newsletter*. March 2011. Vol. 15 #1. pp. 2-6.
- Bartolucci, F. (2007). A class of multidimensional IRT models for testing unidimensionality and clustering items. *Psychometrika*, 72, 141-157.
- Chen, J., & de la Torre, J. (2013). A general cognitive diagnosis model for expert-defined polytomous attributes. *Applied Psychological Measurement*, 37, 419-437.
- De la Torre, J. (2009). DINA model and parameter estimation. *Journal of Educational and Behavioral Statistics*, 34, 115-130.
- De la Torre, J. (2013). The generalized DINA model framework. *Psychometrika*.
- De la Torre, J., & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69(3), 333-353.
- De la Torre, J., & Douglas, J. (2004). Higher-order latent trait models for cognitive diagnosis. *Psychometrika*, 69, 333-353.

- DiBello, L.V. Roussos, L, A. Stout, W. (2007). Review of Cognitively Diagnostic Assessment and a Summary of Psychometric Models. *Handbook of statistics*, 26
- Gierl, M, J. Alves, C. (2010) Using principled test design to develop and evaluate a diagnostic mathematics assessment in grade 3 and 6. Paper presented at the annual meeting of the American educational research association. Denever, Co, USA
- Henson, R., Templin, J., & Willse, J. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, 74(2), 191-210.
- Junker, B. W., & Sijtsma, K. (2001). Cognitive assessment models with few assumptions, and connections with nonparametric item response theory. *Applied Psychological Measurement*, 25, 258-272.
- Mislevy, R. J., Steinberg, L. & Almond, R. (2003). On the structure of educational assessments. *Measurement: Interdisciplinary Research and Perspective*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Nichols, P. D. (1994).
- Nesher, P., & Kilpatrick, J. (Eds.). (1990). *Mathematics and cognition: A research synthesis by the International Group for the Psychology of Mathematics Education*. ICMI Study Series. Cambridge: Cambridge University Press.
- Ravand, H. Barati, H & Widhiarso, W. (2013) Exploring Diagnostic Capacity of a High Stakes Reading Comprehension Test: A Pedagogical Demonstration. *Iranian Journal of Language Testing* Vol. 3, No. 1
- Robitzsch, A. Kiefer, T. George, A. Uenlue, A. (2013). Package 'CDM'. <https://sites.google.com/site/alexanderrobitzsch/software>
- Rupp, A, A. Templin J, L. (2007). Review Article: Unique Characteristics of Cognitive Diagnosis Models. *Educational and Psychological Measurement*
- Su, Yu-lan. (2013). Cognitive diagnostic analysis using hierarchically structured skills. Dissertation, University of Iowa, <http://ir.uiowa.edu/etd/2640>.
- Von Davier, M. (2005). A general diagnostic model applied to language testing data (RR-05-16). Princeton, NJ: Educational Testing Service.
- Von Davier, M. (2008). A general diagnostic model applied to language testing data. *British Journal of Mathematical and Statistical Psychology*, 61, 287-307.
- Young-sun, L. & Yoon Soo, P. (2011). A cognitive diagnostic modeling of attribute mastery in Massachusetts, Minnesota, and the U.S. National sample using the TIMSS 2007. *International journal of testing*, 11:2, 144-177