

شناسایی صفات زیربنایی سؤال‌های آزمون هوش و کسلر چهار کودکان بر اساس توانایی‌های باریک نظریه کتل-هورن-کارول با استفاده از مدل تشخیصی شناختی جی دینا

نوشیروان محمدی^۱، علی دلاور^{۲*}، نورعلی فرخی^۳، اصغر مینایی^۴

تاریخ دریافت: ۹۵/۰۶/۰۳

تاریخ پذیرش: ۹۶/۰۵/۱۵

چکیده

هدف از پژوهش حاضر شناسایی صفات زیربنایی آزمون هوش و کسلر کودکان نسخه چهار با توجه به توانایی‌های باریک نظریه کتل-هورن-کارول و با استفاده از مدل تشخیصی شناختی جی دینا بود. جامعه آماری در این پژوهش کلیه دانش‌آموزان ابتدایی کشور بودند که به صورت تصادفی خوشه‌ای ۱۲۲۲ نفر از آن‌ها انتخاب شدند. ابزار اندازه‌گیری نیز آزمون هوش و کسلر کودکان نسخه ۴ بود. نتایج تحلیل نشان داد که از بین نزدیک به ۷۰ توانایی باریک که در مدل کتل-هورن-کارول به آن‌ها اشاره شده است، ۹ توانایی باریک شامل تصویرسازی، استقراء، حافظه فعال، اطلاعات عمومی کلامی، دانش واژگانی، انعطاف‌پذیری در مشابهت، فراختای حافظه، رشد زبانی، استدلال عمومی زنجیره‌ای به‌عنوان صفت‌های زیربنایی سؤال‌های خرده‌مقیاس آزمون هوش و کسلر شناسایی شدند. همچنین از بین مدل‌های محدود در مدل عمومی، مدل LLM برای تحلیل داده‌ها با توجه به آماره والد انتخاب شد. بررسی میزان تسلط دانش‌آموزان در صفت‌های شناسایی شده نشان داد که بیشترین میزان تسلط در

۱. دانشجوی دکتری رشته سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی

۲. * استاد رشته سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی، delavarali@yahoo.com

۳. دانشیار رشته سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی

۴. استادیار رشته سنجش و اندازه‌گیری، دانشگاه علامه طباطبائی

صفت دانش واژگانی و کمترین میزان تسلط نیز مربوط به انعطاف‌پذیری در مشابهت بود. بررسی پارامترهای سؤال نیز نشان داد که در سؤال‌های آسان هر خرده مقیاس احتمال حدس بالا و در سؤال‌های دشوار هر خرده مقیاس احتمال لغزش بالا بود.

واژگان کلیدی: هوش، آزمون هوش و کسلر، نظریه کتل - هورن-کارول، مدل‌های تشخیصی شناختی

مقدمه

در میان آزمون‌های روان‌شناسی، آزمون‌های استاندارد هوش از جمله ابزارهایی هستند که بیشترین استفاده را توسط روان‌شناسان دارند. روان‌شناسان مدرسه به‌طور خاص اغلب از آزمون‌های استاندارد هوش به‌عنوان یکی از اجزای اصلی ارزشیابی روان‌شناختی دانش‌آموزان برای تعیین شایستگی افراد برای آموزش‌های خاص استفاده می‌کنند. در میان آزمون‌های استاندارد هوش موجود، آزمون و کسلر چهار نسخه کودکان^۱ دارای بیشترین استفاده است (ریچرسون، واتکینس و بیوجان، ۲۰۱۴).

آزمون‌های هوش در مدارس و موقعیت‌های آموزشی برای کسب اطلاعات درباره توانایی‌های کودکان و بزرگسالان برای ابراز وجود خود، حل مسئله و استدلال و انجام تکالیف مختلف مورد استفاده قرار می‌گیرند. از اطلاعات به‌دست‌آمده از آزمون‌های هوش می‌توان برای طراحی برنامه‌ها برای کودکان با نیازمندی‌های خاص استفاده کرد. همچنین می‌توان اقدام به شناسایی افراد باهوش و هدایت آنان در برنامه‌های پیشرفته آموزشی کرد. این آزمون‌ها نقش ویژه‌ای در آموزش و پرورش کودکان استثنایی بر عهده‌دارند. علاوه بر این، توانایی شناختی کودکان مانند توانایی‌های آن‌ها برای انجام محاسبات ذهنی، توجه و یادآوری آنچه یاد گرفته‌اند،

-
1. Wechsler intelligence scale for children-fourth edition (WISC-IV)
 2. Richerson, Watkins & Beaujean

همیشه منبع اضطرابی برای والدین است برای اینکه بدانند آیا کودکشان به صورت نرمال رشد کرده‌اند یا خیر (ماسون و ویلکاکس^۱، ۲۰۰۹).

آزمون هوش و کسلر و به‌طور خاص آزمون هوش و کسلر برای کودکان نسخه چهار، از آزمون‌هایی است که بیشترین استفاده را در جهان دارد. آزمون‌های هوش و کسلر تقریباً در ۲۰ کشور جهان برای اندازه‌گیری هوش استفاده می‌شود. با توجه به تحلیل‌های عاملی اکتشافی و تأییدی این آزمون از یک ساختار چهار عاملی در همه این کشورها تبعیت می‌کند (رورت، گولی، فاویز، روزیر و لچرف^۲، ۲۰۱۵).

در سال ۲۰۰۸ آزمون و کسلر مورد تجدیدنظر قرار گرفت. هدف از این تجدیدنظر توسعه مبانی نظری و زیربنایی آزمون و افزایش کیفیت ویژگی‌های روان‌سنجی آزمون و سهولت استفاده و کاربرد بالینی آن بود. بزرگ‌ترین تغییر در نسخه جدید حذف مقیاس‌های هوش کلامی و غیرکلامی است. به دلیل کمبود مبانی نظری و روایی این ساختار دوگانه جای خود را به چهار شاخص جدید داده است (ون آکن، هیدن، ولد، هرمانس، کسل و اگر^۳، ۲۰۱۵).

اخیراً محققان به بررسی ساختار آزمون هوش و کسلر چهار بر اساس نظریه کتل-هورن - کارول^۴ پرداخته‌اند، توانایی‌هایی که در این نظریه ارائه شده‌اند نقش مهمی در توسعه آزمون‌های هوش در سه دهه اخیر داشته‌اند و این نظریه بیشتر و بیشتر به‌عنوان راهنمایی در ساخت و توسعه خرده مقیاس‌ها استفاده می‌شود. اگرچه پژوهش‌ها از یک ساختار چهار عاملی در آزمون و کسلر چهار حمایت می‌کند، چندین مطالعه از وجود شواهد روان‌سنجی برای ساختارهای دیگر، از جمله ساختارهای مبتنی بر نظریه مانند نظریه کتل-هورن-کارول را در آزمون و کسلر نشان می‌دهد. اگرچه در استفاده از نظریه کتل-هورن-کارول برای تفسیر نتایج آزمون و کسلر دشواری‌هایی وجود دارد؛ برای مثال الف) خرده مقیاس‌ها صرفاً بر اساس اندازه‌گیری توانایی اشاره در نظریه

-
1. Mason & Wilcox
 2. Reverte, Golay, Favez, Rossier & Lecerf
 3. Van Aken, Heijden, Veld, Hermans, Kessels & Egger
 4. Cattle, Horn & Caroll

کتل-هورن-کارول نیستند، ب) خرده مقیاس‌های همه توانایی‌های نظریه کتل-هورن-کارول را پوشش نمی‌دهند ج) مشخص نیست که آیا خرده مقیاس‌ها همه دامنه و حوزه‌های آن توانایی را اندازه‌گیری می‌کنند یا خیر؟. با وجود این تفسیر آزمون و کسلر بر اساس نظریه کتل-هورن-کارول و قرار دادن نتایج در یک چهارچوب نظری کاری ارزشمند محسوب می‌شود (ون آکن و همکاران، ۲۰۱۵).

نظریه سلسله مراتبی کتل-هورن-کارول که از جمله نظریات جدید در حوزه سنجش و اندازه‌گیری هوش است، یک نظریه جامع است که هم توانمندی‌های شناختی و هم توانمندی‌های تحصیلی را دربر می‌گیرد و هم قادر به تمایز این دو مفهوم از همدیگر هست و در رویکرد سوم مربوط به تعیین اختلالات یادگیری استفاده می‌شود. نظریه کتل-هورن-کارول درباره توانایی‌های شناختی در بین نظریه‌های روان‌سنجی در واقع جامع‌ترین و دارای بیشترین حمایت تجربی از داده برای ساختار توانایی‌های شناختی است. این نظریه در واقع ترکیبی از کارهای ریموند کتل، جان هورن و جان کارول است. از آنجایی که این نظریه یک مجموعه حمایت‌های تجربی در پیشینه پژوهش (رشد، عصب‌شناختی، نتیجه محوری^۱) برخوردار است، به صورت وسیعی به‌عنوان مبنایی برای انتخاب، سازمان‌دهی و تفسیر آزمون‌های هوش و توانایی‌های شناختی استفاده می‌شود. اخیراً این نظریه برای طبقه‌بندی سؤالات هوش، پیشرفت تحصیلی و آزمون‌های روان‌شناسی استفاده می‌شود تا الف) تفسیر عملکردهای شناختی را تسهیل کند ب) یک مبنایی برای ارزشیابی افرادی که مشکوک به ناتوانایی‌های یادگیری هستند، فراهم آورد (رینولدز، ونست، فلچر-یانزن^۲، ۲۰۱۳).

تئوری سه لایه^۳ پیشرفت‌های به‌دست‌آمده در روان‌شناسی شناختی را نشان می‌دهد زیرا که این پیشرفت‌ها تفسیر (تبیین) یافته‌های به‌دست‌آمده از تحلیل‌های عاملی درباره وظایف شناختی

1. outcome - criterion
2. Reynolds, Vannest & Fletcher-Janzen
3. stratum

را آسان‌تر می‌کند. همچنین تحقیقات روان‌شناسی شناختی باعث شد تا بر روی وظایف شناختی مختلفی که قبلاً به‌طور گسترده‌ای به‌وسیله روان‌سنجان نادیده گرفته می‌شد، تأکید شود. تئوری سه لایه توانایی‌های شناختی در واقع توسعه‌ی تئوری‌های قبلی است. این تئوری مشخص می‌کند که کدام‌یک از انواع تفاوت‌های فردی در توانایی‌های شناختی وجود دارد و چطور آن تفاوت‌ها با همدیگر ارتباط دارند. این نظریه یک نقشه از همه توانایی‌های شناختی شناخته‌شده فراهم می‌آورد که به‌عنوان راهنمایی برای تحقیق و اجرا می‌توان از آن استفاده کرد. در این نظریه فرض بر این است که تفاوت‌های فردی مجزایی در توانایی‌های شناختی وجود دارد و روابط بین آن‌ها را می‌توان با طبقه‌بندی در سه لایه متفاوت نشان داد: لایه اول، توانایی‌های باریک^۱ که حدود ۷۰ توانایی شناختی می‌باشند، لایه دوم، توانایی‌های گسترده^۲ با ۱۶ توانایی شناختی گسترده، لایه سوم از تنها عامل توانایی عمومی یا کلی^۳ تشکیل شده است (فلانگان و هریسون^۴، ۲۰۱۲).

در لایه سوم طبقه‌بندی نظریه کتل _ هورن _ کارول عامل هوش کلی که همان عامل g هست قرار دارد. در لایه دوم توانایی‌های گسترده قرار دارند که شامل هوش سیال (Gf)، هوش متبلور (Gc)، حافظه کوتاه‌مدت (Gsm)، پردازش تصویری (Gv)، پردازش صوتی (Ga)، حافظه بلندمدت (GLr)، سرعت پردازش (Gs)، تصمیم‌گیری و واکنش سریع (Gt)، خواندن و نوشتن (Grw)، دانش کمی (Gq)، دانش کلی (GKn)، توانایی حسی (Gh)، توانایی حرکتی (Gk)، توانایی کشف اطلاعات معنی‌دار (Go)، توانایی حسی حرکتی (Gp)، سرعت روانی - حرکتی (Gps). هر کدام از این عامل‌های گسترده خود نیز عامل‌های باریک‌تری دارند که لایه اول را تشکیل می‌دهند. عامل‌های باریک هوش سیال شامل استقرا، قیاس و استدلال کمی هستند، هوش متبلور شامل پنج عامل باریک مانند اطلاعات کلامی عمومی، رشد زبانی، توانایی شنیداری هست. حافظه کوتاه‌مدت شامل دو عامل مدت نگه‌داری و ظرفیت حافظه فعال هست. پردازش

-
1. narrow ability
 2. broad ability
 3. g ability
 4. Flanagan & Harrison

تصویری شامل یازده عامل مانند تصویرسازی، حافظه تصویری و غیره هست. پردازش صوتی شامل هشت عامل مانند کدگذاری فونتیک، حافظه الگوهای صوتی هست. حافظه بلندمدت شامل دوازده عامل مانند حافظه معنی‌دار، خلاقیت، سیالی کلمات، سیالی شکلی هست. سرعت پردازش شامل شش عامل مانند سرعت ادراکی، سرعت خواندن هست. تصمیم‌گیری و واکنش سریع شامل پنج عامل مانند زمان واکنش، زمان واکنش به گزینه‌ها، سرعت مقایسه ذهنی هست. عامل خواندن و نوشتن شامل هفت عامل مانند رمزگشای خواندن، درک خواندن، سرعت خواندن و غیره هست. عامل دانش کمی شامل دو عامل دانش ریاضی و پیشرفت ریاضی هست. دانش کلی شامل هشت عامل مانند مهارت در زبان‌های خارجه، پیشرفت جغرافی، دانش فرهنگی و غیره هست. توانایی حسی حرکتی شامل هشت عامل مانند آمادگی عضلانی (دست و سر)، دقت کنترل هست؛ و عامل سرعت روانی - حرکتی شامل چهار عامل مانند سرعت حرکت بازو، سرعت نوشتن هست (اشنایدر و مک‌گرو، ۲۰۱۲).

از سویی دیگر، از مشکلات نمره‌های حاصل از مدل‌های سنتی روان‌سنجی، ازجمله نظریه کلاسیک آزمون^۲ و نظریه سؤال پاسخ^۳ این است که اطلاعات شناختی مستتر در آزمون را منعکس نمی‌سازد. امبرستون می‌گوید که اگرچه مدل‌های نظریه سؤال پاسخ از مزایای زیادی نسبت به روش‌های پیشین برخوردارند، اما ارتباط اندکی با موضوعات و مسائل مورد توجه در نظریه‌های شناختی درباره فرایندها، راهبردها و ساختارهای دانش که زیربنای حل سؤال قرار می‌گیرند، دارند. در واقع یکی از محدودیت‌های مدل‌های سنتی روان‌سنجی این است که عملکرد آزمون‌شوندگان در آزمون به صورت یک نمره منفرد، مانند نمره خام یا برآوردهای توانایی (تتا)، ارائه می‌شود که اطلاعاتی در خصوص اینکه دانش آموزان واقعاً چه دانش یا مهارت‌های را دارا هستند فراهم نمی‌کنند. این محدودیت اساساً ناشی از لحاظ نکردن اطلاعات شناختی در مدل‌های

-
1. Schneider and Mc grew
 2. classical test theory
 3. item response theory

روان‌سنجی است. در راستای کمک به استفاده از سنجش برای بهبود و ارتقا یادگیری، پژوهشگران و متخصصان حوزه سنجش آموزشی با اتکا به پژوهش‌های صورت گرفته در حوزه روان‌شناسی یادگیری و با بهره‌گیری از مدل‌های آماری، رویکرد سنجش تشخیصی شناختی^۱ را به‌منظور انجام استنباط‌های راجع به ساختارها و فرایندهایی که زیربنای عملکرد در حیطه موردنظر را تشکیل می‌دهند، پرورش یافت (مینایی، ۱۳۹۱).

در طول دهه گذشته مدل‌های تشخیصی شناختی به‌طور گسترده‌ای مورد مطالعه قرار گرفته‌اند و کاربردهای آن‌ها در داده‌ها آموزشی افزایش یافته است. یکی از هدف‌های CDM طبقه‌بندی افراد بر اساس الگوی پاسخ‌هایشان با توجه به تعداد مشخصی خرده توانایی‌هایی است که توانایی کل یک حیطه را که باید اندازه‌گیری شود، تشکیل می‌دهد. ارزش‌های انفرادی افراد در هر خرده مقیاس یک طبقه‌بندی چندبعدی را می‌سازد که اطلاعات تشخیصی بیشتری را در مقایسه با نمره فراهم می‌آورد و می‌تواند به‌عنوان مبنایی برای بازخورد و حمایت استفاده شود (جورج و رویتیچ^۲، ۲۰۱۴).

به‌عبارت‌دیگر، پارادایم‌های مبتنی بر نمره کل (نظریات کلاسیک اندازه‌گیری) در مقابل دیدگاه توصیف عمیق از پیشرفت دانش آموزان که نیازمند فراهم آوردن اطلاعات تشخیصی است، به چالش کشیده شده‌اند. برای حل چنین چالشی، محققان اندازه‌گیری اقدام به ترکیب روان‌شناسی شناختی و نظریه‌های روان‌سنجی کردند. هدف چنین تعاملی گسترش روش ساخت آزمونی است که برای دانش‌آموزان بازخوردهای تشخیصی مناسب درباره سطح تسلط آن‌ها در صفت‌های شناختی فراهم آورد تا دانش‌آموزان نقص‌هایشان را در مهارت‌هایی که به تسلط نرسیده‌اند، جبران کنند (راپ و تمپلین^۳، ۲۰۰۸).

-
1. cognitive diagnostic model
 2. George & Robitzsch
 3. Rupp & Templin

مدل‌های تشخیصی شناختی از جمله مدل‌های روان‌سنجی هستند که برای مشخص کردن توانایی تسلط آزمودنی در مهارت‌های اساسی بر اساس یک ماتریس از پیش تعیین شده توسعه پیدا کرده‌اند. آزمون‌های تشخیصی شناختی برای شناسایی ترکیبی از مهارت‌ها که آزمودنی می‌تواند آن‌ها را داشته باشد یا نداشته باشد، به کار می‌روند. نتایج این مدل‌ها اطلاعات بازخوردی دقیقی را برای آزمودنی فراهم می‌آورد، بنابراین آن‌ها می‌توانند در تسلطشان درباره مهارت‌های شناختی مختلف استنباط‌هایی داشته باشند. برخلاف مدل‌های سؤال پاسخ که بر روی تحلیل سؤال و نسبت دادن نمره به آزمودنی تمرکز می‌کنند، مدل‌های تشخیصی شناختی بیشتر بر فراهم آوردن بازخوردهای اطلاعاتی و تشخیصی در مهارت‌های خاص تأکید می‌کنند که آزمودنی باید در آن‌ها پیشرفت کند. به‌طور مشخص، آن‌ها نیمرخ صفت آزمودنی را که نشان‌دهنده تسلط یا عدم تسلط آزمودنی در هر مهارت است را نشان می‌دهند. در مدل‌های تشخیصی شناختی، نیمرخ صفت آزمودنی (که گاهی اوقات الگوی مهارت برآورد شده نامیده می‌شود) متغیرهای مکنون طبقه‌ای هستند (متغیرهای گسسته)؛ بنابراین در فرایند نمره‌گذاری، مدل‌های تشخیصی شناختی پاسخ‌های آزمودنی را طبقه‌های مکنون گسسته طبقه‌بندی می‌کنند (ترکیبی از ۰ها و ۱ها که ۰ به معنی عدم تسلط فرد در یک مهارت و ۱ به معنی تسلط فرد در آن مهارت است). تقریباً همه مدل‌های تشخیصی شناختی نیازمند تعیین یک ماتریس کیو $J*K$ هستند که K نشان‌دهنده تعداد مهارت‌های است که به‌وسیله آزمون اندازه‌گیری می‌شوند و J نشان‌دهنده تعداد سؤال‌های یک آزمون است (تاتسوکا، ۱۹۹۵؛ به نقل از سوآ، ۲۰۱۳).

در سنجش تشخیصی شناختی، علاوه بر برآورد توانایی فرد در سازه یا سازه‌های مورداندازه‌گیری، نیمرخ برای هر فرد فراهم می‌آید که تسلط یا عدم تسلط وی را در مجموعه‌ای از مهارت‌های اساسی یا صفات شناختی خرد از پیش تعیین شده که برای پاسخگویی به سؤال‌ها لازم است، مشخص می‌سازد. به عبارت دقیق‌تر، در این مدل‌ها با توجه به تسلط یا عدم

1. Tatsuoka
2. Yu-Lan Su

تسلط فرد بر مهارت‌ها و صفاتی که برای پاسخگویی درست به سؤال موردنیاز هستند، احتمال پاسخ درست فرد به سؤال نیز برآورد می‌شود. در این مدل‌ها می‌توان بر اساس وضعیت تسلط افراد در هر یک از مهارت‌ها و صفات زیربنایی موردنیاز سؤال‌ها، آن‌ها را به دو گروه مسلط و غیر مسلط و در صورت لزوم به ۳ گروه مسلط، غیر مسلط و نامعین طبقه‌بندی نمود و از این طریق مشخص کرد فرد در کدام یک از مهارت‌ها و صفات زیربنایی مشکل دارد (مینایی، ۱۳۹۱). پژوهش‌های زیادی در رابطه با کاربرد مدل‌های شناختی در حوزه‌های مختلف انجام شده است.

تمپلین و ایوی^۱ (۲۰۰۶) به تحلیل آزمون ماتریس‌های پیش‌رونده ریون^۲ با استفاده از سنجش مهارت‌ها پرداختند. آن‌ها با استفاده از مدل‌های شناختی به شناسایی مؤلفه‌های زیربنایی آزمون ریون پرداختند و مهارت‌های زیربنایی زیر را استخراج کردند: یکسانی، پیشرفت، جمع کردن و تفریق کردن شکل‌ها، توزیع‌های سه‌تایی، توزیع‌های دوتایی. در این تحقیق هم از مدل دینا استفاده شد و نتایج نشان داد که این مدل اطلاعات کامل‌تری از روش‌ها نمره کل به دست می‌دهد.

لی^۳ (۲۰۱۱) مدل‌های شناختی را برای تحلیل آزمون خواندن^۴ استفاده کرد. وی با استفاده از مدل فیوژن چهار خرده‌مقیاس زیربنایی برای این آزمون شامل: لغات، علم نحو، استخراج اطلاعات عینی، فهم اطلاعات ضمنی و هم‌چنین نیمرخ هر فرد را نیز به دست آورد. وی^۵ (۲۰۱۳) به اجرای سنجش تشخیصی شناختی در آزمون‌های موسسه‌ای پرداخت. وی در پژوهش خود به مقایسه پنج مدل شناختی در آزمون‌های خواندن تافل، شنیداری تافل^۶ و گرامر

-
1. Templin & Ivie
 2. Raven's progressive matrices (RPM)
 3. Lee
 4. Michigan english language assessment battery (MELAB)
 5. Yi
 6. test of english as a foreign language (TOEFL)

آزمون صلاحیت در مهارت‌های زبان^۱ پرداخت. نتایج نشان داد که مدل جبرانی مدل یکپارچه با پارامتر پردازی مجدد^۲ بهترین برازش را با داده‌ها دارد.

گارسیا، دیاز و دلاتوره^۳ (۲۰۱۴) به بررسی کاربرد مدل‌های تشخیصی شناختی برای شناسایی مهارت‌های آزمون‌های قضاوت موقعیتی^۴ پرداختند. آن‌ها از مدل جی‌دینا استفاده کردند. نتایج نشان داد که مدل‌های تشخیصی شناختی ابزار مناسبی برای اندازه‌گیری توانایی شغلی از طریق آزمون قضاوت موقعیتی هستند. مدل‌های شناختی توانستند مهارت‌های موردنیاز را به‌خوبی اندازه‌گیری کنند و برآورد دقیقی از این مهارت‌ها نیز داشته باشند.

دلاتوره، ون در آرک و روسی^۵ (۲۰۱۵) به ارائه یک مدل تشخیصی شناختی کلی برای تشخیص اختلالات روانی با استفاده از داده‌های به‌دست‌آمده از آزمون بالینی چند محوری میلیون^۶ پرداختند. نتایج تحلیل داده‌های آن‌ها نشان داد که استفاده از مدل‌های تشخیصی شناختی برای تحلیل پرسشنامه‌های شخصیت منجر به نتایج دقیق‌تر برای تشخیص اختلالات، بررسی ویژگی‌های سؤالات و بررسی ساختار زیر بنایی هر اختلال می‌شود.

همچنین در داخل کشور مینایی (۱۳۹۱) در پایان‌نامه خود تحت عنوان مدل پردازی تشخیصی شناختی سؤال‌های ریاضی مطالعه بین‌المللی روند آموزش ریاضیات و علوم^۷ ۲۰۰۷ در دانش آموزان پایه هشتم ایران با استفاده از مدل یکپارچه با پارامتر پردازی مجدد^۸ و مقایسه مهارت‌های ریاضی دانش آموزان دختر و پسر به این نتیجه رسید که وضعیت دانش آموزان ایرانی در صفات

-
1. Examination for the Certificate of Proficiency in English (ECPE)
 2. Compensatory RUM
 3. Garcia, Diaz and Torre
 4. situational judgment tests
 5. de la Torre, van der Ark and Rossi
 6. Million Clinical Multiaxial Inventory (MCMI)
 7. Trends in International Mathematics and Science Study (TIMSS)
 8. Reparametrized unified model

و مهارت‌های ریاضی خوب نیست، همچنین نتایج نشان داد که جنسیت تأثیر زیادی بر روی عملکرد افراد ندارد.

افضلی (۱۳۹۳) در پژوهش خود به تدوین و ارزشیابی مدل تشخیصی شناختی ریاضیات پایه اول دبیرستان با استفاده از روش سلسله مراتبی صفات^۱ پرداخت. تجزیه و تحلیل داده‌ها با به کارگیری مدل دینای سلسله مراتبی در مدل‌سازی تشخیصی شناختی ریاضیات نشان داد که هشت صفت زیربنایی تبیین‌کننده عملکرد دانش آموزان پایه اول است. همچنین نتایج پژوهش نشان داد که آزمودنی‌ها در هیچ کدام از هشت مهارت به تسلط نرسیده‌اند.

کبیری (۱۳۹۳) در پژوهش خود به مطالعه کاربرد سنجش تشخیصی شناختی به منظور تعیین مهارت‌های کسب‌شده علوم تجربی در دانش آموزان پایه تحصیلی هشتم ایران بر اساس داده‌های مطالعه بین‌المللی روند آموزش ریاضیات و علوم ۲۰۱۱ پرداخت. نتایج نشان داد که مدل هشت خصیصه‌ای برای تبیین عملکرد یادگیری دانش آموزان در جامعه مورد مطالعه بهترین مدل است. همچنین محسن پور (۱۳۹۳) در پژوهش خود به طراحی، ساخت و اعتبار یابی آزمون تشخیصی شناختی سواد ریاضی در پایان دوره آموزش عمومی در جهت تدوین الگوی کیفیت آن پرداخت. وی در پژوهش خود از مدل دینا^۲ استفاده کرد و نتایج نشان داد که هراندازه سطوح صلاحیت شناختی در سؤال‌ها افزایش یابد، عملکرد دانش آموزان به همان اندازه پایین تر خواهد بود.

مقدم (۱۳۹۴) به بررسی کاربرد مدل‌های تشخیصی شناختی به منظور تعیین مهارت‌های زیر بنایی عملکردی داوطلبان در آزمون ورودی زبان انگلیسی عمومی دوره دکتری پرداخت. وی با استفاده از روش‌های کیفی سه مهارت را برای آزمون قواعد و شش مهارت احتمالی آزمون درک مطلب را شناسایی کرد. در بخش کمی نیز مدل غیر جبرانی فیوژن، با رویکرد بیز و روش

1. Attribute Hierarchal Model (AHM)
2. Deterministic inputs, noisy 'and' gate

زنجیره‌های مارکوف مونته کارلو^۱ را استفاده کرد. نتایج این مطالعه نشان داد که استفاده از مدل‌های تشخیصی شناختی نسبت به روش‌های مرسوم در تحلیل داده‌های آزمون زبان بهتر هستند.

می‌توان گفت که وقتی آزمونی بر اساس نظریه CHC ساخته و تفسیر می‌شود، ماهیت توانای شناختی و تحصیلی که به وسیله این آزمون اندازه‌گیری می‌شود، بهتر درک می‌شوند؛ بنابراین وقتی که به ارزیابی و تفسیر پیشرفت تحصیلی پرداخته می‌شود، استفاده از این نظریه ضروری به نظر می‌رسد تا ماهیت توانایی‌ها و ضعف‌های دانش‌آموزان را به درستی روشن سازد. همچنین استفاده از آزمون‌های مبتنی بر نظریه CHC به درمانگران این امکان را می‌دهد که اندازه‌گیری دقیق و کافی و عمیق از مشکل داشته باشند و همچنین این فرصت را مهیا می‌سازد تا به بررسی دلایل بنیادی مشکلات یادگیری برای مثال در کودکان با اختلال یادگیری پردازند. با توجه به کاربرد زیاد آزمون هوش و کسلر تاکنون پژوهشی در مورد فرایندهای شناختی زیر بنایی این آزمون انجام نشده است. این پژوهش به دنبال پاسخگویی به سؤال زیر است:

- ۱) مجموعه مهارت‌ها و صفات زیر بنایی شناختی برای آزمون هوش و کسلر کودکان نسخه چهار (۲۰۰۳) بر اساس توانایی‌های باریک نظریه کتل-هورن _ کارول کدام‌اند؟
- ۲) دانش‌آموزان در چه صفات و مهارت‌هایی به تسلط رسیده‌اند و در چه صفات و مهارت‌هایی به تسلط نرسیده‌اند؟

روش پژوهش

از آنجاکه در این تحقیق از داده‌های آزمون هوش و کسلر نسخه چهار که در سازمان آموزش و پرورش استثنایی گردآوری شده بود، استفاده شد، بنابراین این تحلیل از نوع تحلیل‌های ثانویه هست. جامعه آماری این پژوهش از کلیه دانش‌آموزان ابتدایی سراسر کشور بودند. نمونه آماری

این پژوهش برابر با ۱۲۲۲ نفر که از این تعداد ۶۲۲ نفر پسر و ۶۰۰ نفر دختر می‌باشند. این افراد با استفاده از روش نمونه‌گیری تصادفی خوشه‌ای انتخاب شدند.

ایزار اندازه‌گیری در این پژوهش آزمون و کسلر کودکان نسخه چهار هست. این آزمون برای دامنه سنی ۶ تا ۱۶ استفاده می‌شود. از ۱۵ خرده‌مقیاس تشکیل شده است، ۱۰ مقیاس اصلی و ۵ مقیاس مکمل، هر مقیاس دارای میانگین ۱۰ و انحراف استاندارد ۳ هست. این آزمون دارای یک نمره هوش بهر کلی^۱ برای کل مقیاس و چهار شاخص نمره‌گذاری: درک کلامی^۲، استدلال مفهومی^۳، حافظه فعال^۴، سرعت پردازش^۵ هست. میانگین نمره کل برابر ۱۰۰ و انحراف استاندارد آن ۱۵ هست.

شواهد زیادی در مورد روایی آزمون هوش و کسلر وجود دارد. نتایج تحلیل عاملی‌های مختلف مدل چهار عاملی را برای این آزمون نشان می‌دهد. همچنین آزمون هوش و کسلر همبستگی بالایی با سایر آزمون‌هایی که هوش را اندازه‌گیری می‌کنند، دارد که این نشان از روایی همگرایی این آزمون دارد. برای مثال همبستگی آن با آزمون هوش سیال-متبلور کافمن برای کودکان^۶ برابر ۰/۸۹ و با آزمون پردازش ذهنی کافمن برای کودکان^۷ برابر ۰/۸۸ و همچنین با مقیاس توانایی‌های افتراقی^۸ برابر ۰/۸۴ هست که همبستگی‌های بالایی محسوب می‌شود و نشان از روایی بالای آن دارد (والستروم، بروژ، ژو و ویس، ۲۰۱۲).

همچنین مقدار ضریب پایایی به‌دست آمده برای تمامی خرده‌مقیاس‌ها به‌جز دو خرده‌مقیاس برابر ۰/۸ تا ۰/۹ است. برای خرده‌مقیاس‌های تکمیل خط‌زنی و نماد یابی برابر ۰/۷۹ است.

-
1. full scale IQ (FSIQ)
 2. verbal comprehension index (VCI)
 3. perceptual reasoning index (PRI)
 4. working memory index (WMI)
 5. processing speed index (PSI)
 6. kaufman assessment battery for children- fluid- crystalized index
 7. Kaufman Assessment Battery for Children- mental processing index
 8. Differential Ability Scales- General conceptual ability
 9. Wahlstrom, Breaux, Zhu, Weiss

همچنین ضریب باز آزمایی برای نمره کل آزمون برابر ۰/۹۲ برای چهار شاخص اصلی برابر ۰/۸۴ تا ۰/۹۳ و برای خرده مقیاس‌ها برابر دامنه‌ای از ۰/۷۴ تا ۰/۹۲ است (والستروم، بروژ، ژو و ویس، ۲۰۱۲).

در ایران هم صادقی، ربیعی، عابدی (۱۳۹۰) به بررسی روایی و پایایی آزمون در کودکان ۶ تا ۱۶ سال استان چهارمحال و بختیاری پرداختند نتایج نشان داد که بین ویرایش چهارم آزمون و کسلر و مقیاس تجدیدنظر شده هوش و کسلر و آزمون ریون همبستگی معنادار وجود دارد. برای بررسی پایایی آزمون از روش‌های باز آزمایی و دونیمه کردن استفاده شد. پایایی به دست آمده با استفاده از روش باز آزمایی برای نمره کلی هوش برابر ۰/۹۱ و برای شاخص‌ها از ۰/۸۸ تا ۰/۸۱ بود که مقدارهای بالایی حساب می‌شوند. برای خرده مقیاس‌ها نیز بالاترین پایایی مربوط به خرده مقیاس واژگان با ۰/۹۴ و پایین‌ترین آن مربوط به مفاهیم تصویری با ۰/۶۵ صدم بود. در روش دو نیمه‌سازی نیز پایایی نمره کل هوش برابر ۰/۹۵ و برای شاخص‌ها از ۰/۹۱ تا ۰/۸۳ بود. در خرده مقیاس‌ها بالاترین پایایی برای واژگان و استدلال تصویری ۰/۸۶ و کمترین پایایی مربوط به فراخنای ارقام با ۰/۷۱ هست.

برای تحلیل داده‌ها در این پژوهش از مدل تشخیصی شناختی جی دینا استفاده شد. مدل تشخیصی شناختی جی دینا از جمله مدل‌های کلی شناختی هست که توانایی انجام تحلیل ماتریس‌های با صفات زیاد و تعداد سؤالات بالا را دارد. از آنجاکه در این پژوهش نیز هم تعداد صفات (۹ صفت) و هم تعداد سؤالات (۲۱۹) بالا بود، مدل جی دینا برای تحلیل داده‌ها انتخاب گردید. برای تحلیل داده‌ها در مدل‌های شناختی به صورت زیر عمل شد:

مرحله اول: تعریف مهارت‌ها و صفات زیربنایی شناختی. برای تعریف صفات از منابع مختلف مانند نظریه حیطه محتوایی، تحلیل محتوای سؤال، تفکر با صدای بلند، نظر متخصصان و پژوهش‌های مرتبط می‌توان استفاده کرد. در این پژوهش با استفاده از توانایی‌های لایه اول نظریه CHC و نظر متخصصان به شناسایی اینکه هر سؤال چه مهارت و توانایی را اندازه‌گیری می‌کند،

پرداخته شد. در این مرحله از دو نفر از متخصصان حوزه هوش و روان‌شناسی شناختی برای شناسایی صفات زیر بنایی استفاده شد.

مرحله دوم: ساخت ماتریس کیو. ماتریس Q نشان‌دهنده ساختار بارگیری در مدل‌های شناختی است. هر سطر این ماتریس، یک فرضیه درباره صلاحیت‌های موردنیاز برای کسب پاسخ درست به یک سؤال است و دارای مرتبه $m \times k$ است که در آن n، تعداد سؤال‌ها و k تعداد صلاحیت‌های موردنیاز برای سؤال‌هاست. تمام درایه‌های این ماتریس، اعداد صفر و یک هستند. برای یک درایه خاص ماتریس Q در سطر n م و ستون k م، عدد یک نشانگر آن است که سؤال n آم صفت یا صلاحیت k آم را اندازه‌گیری می‌کند و عدد صفر نشان می‌دهد که سؤال‌ها، صلاحیت یا صفت موردنظر را اندازه نمی‌گیرد.

مرحله سوم: برآورد مدل. هدف اولیه اصلی تحلیل داده‌ها با استفاده از CDM برآورد نیمرخ صفتی برای فراگیران به صورت فردی بر اساس عملکرد آن‌ها در آزمون است. برای این کار باید یک مدل تشخیصی شناختی انتخاب شود، برای این پژوهش مدل غیر جبرانی جی‌دینا مورد استفاده قرار گرفت. در مدل جی‌دینا، احتمال پاسخ درست آزمون‌شونده i آم با الگوی صلاحیت α ؛ در سؤال j آم، به وسیله معادله زیر تعریف می‌شود:

$$P(X_j = 1 | \alpha_{ij}^*) = P(\alpha_{ij}^*)$$

مرحله چهارم: گزارش نمره و بازخورد تشخیصی. در این مرحله نمرات افراد و میزان تسلط یا عدم تسلط افراد در هر یک از مؤلفه‌ها نیز مشخص شد.

لازم به ذکر است در این پژوهش از نرم‌افزار آماری R برای انجام تحلیل‌های مربوط به مدل شناختی استفاده گردید.

سؤال اول: مجموعه مهارت‌ها و صفات زیر بنایی شناختی برای آزمون هوش و کسلر کودکان

نسخه چهار (۲۰۰۳) بر اساس توانایی‌های باریک نظریه کتل _ هورن _ کارول کدام‌اند؟

برای پاسخ‌دهی به این پرسش اول ماتریس کیو ساخته شد. مرحله اول در ساخت ماتریس کیو تعریف صفات و مهارت‌های زیربنایی سؤال‌های آزمون است. در این پژوهش با توجه به اینکه شناسایی صفات بر اساس توانایی‌های باریک نظریه کتل - هورن - کارول بود، از صفات شناسایی شده در این نظریه استفاده شد. در این نظریه نزدیک به ۷۰ توانایی باریک شناسایی شده‌اند. در این مرحله هم‌چنین از نظر ۲ نفر متخصص برای مشخص کردن اینکه آزمون هوش و کسلر کودکان کدام‌یک از این توانایی‌ها را اندازه‌گیری می‌کند، استفاده شد. در نهایت با توجه به مبانی نظری و نظر متخصصان ۱۰ توانایی باریک به‌عنوان توانایی‌ها و صفات زیربنایی خرده‌مقیاس‌های اصلی آزمون هوش و کسلر کودکان نسخه چهار شناسایی شدند. این توانایی‌ها عبارت بودند از: تصویرسازی^۱، استقرا^۲، حافظه فعال^۳، اطلاعات عمومی کلامی^۴، دانش واژگانی^۵، انعطاف‌پذیری در مشابهت^۶، فراخوانی حافظه^۷، رشد زبانی^۸، استدلال عمومی زنجیره‌ای^۹، سرعت چرخش^{۱۰}. در جدول زیر ارتباط هر یک از این توانایی‌ها با خرده‌مقیاس‌های اصلی آزمون هوش و کسلر نشان داده شده است.

-
1. visualization
 2. induction
 3. working memory
 4. general verbal information
 5. lexical knowledge
 6. flexibility of closure
 7. memory span
 8. Language development
 9. general sequential reasoning
 10. speeded rotation

جدول ۱. ماتریس توانایی‌های باریک و خرده مقیاس‌های آزمون هوش و کسلر

SR	RG	LD	Ms	Cf	VL	Ko	Mw	I	Vz
سرعت	استدلال	رشد	فراختای	انعطاف‌پذیری	دانش	اطلاعات	حافظه	استقرا	تصویرسازی
چرخش	عمومی	زبانی	حافظه		واژگانی	عمومی	فعال		
		زنجیره‌ای							
*									* طراحی مکعب
				*	*			*	شباهت‌ها
			*			*	*		فراختای ارقام
								*	مفاهیم تصویری
					*				واژگان
							*		توالی حرف و عدد
	*							*	استدلال تصویری
			*			*			درک مطلب

مرحله دوم ساخت ماتریس کیو. در ماتریس کیو در مدل‌های شناختی هر سؤال در ردیف و صفات در ستون‌ها قرار می‌گیرند. درایه‌های ماتریس از اعداد ۰ و ۱ تشکیل می‌شود. ۰ به این معنی که آن صفت برای آن سؤال لازم نیست و ۱ به این معنی که برای پاسخ‌دهی صحیح به سؤال آن صفت لازم است. از آنجا که تعداد سؤال‌های آزمون و کسلر و تعداد صفات مورد نیاز نیز بسیار بالا هست، بنابراین اندازه ماتریس بسیار بزرگ هست. در ادامه خلاصه‌ای از ماتریس مورد نظر آورده می‌شود.

جدول ۲. ماتریس کیو سؤالات و توانایی های باریک

Vz	SR	I	MW	Ko	vi	cf	Ms	LD	RG
تصویرسازی	سرعت چرخش	استقرا	حافظه فعال	اطلاعات عمومی کلامی	دانش واژگانی	انعطاف پذیری در مشابهت	فراختای حافظه	رشد زبانی	استدلال عمومی زنجیره ای
۱	۱	۰	۰	۰	۰	۰	۰	۰	۰
۱	۱	۰	۰	۰	۰	۰	۰	۰	۰
۰	۰	۱	۰	۰	۱	۱	۰	۰	۰
۰	۰	۱	۰	۰	۱	۱	۰	۰	۰
۰	۰	۰	۱	۰	۰	۰	۱	۰	۰
۰	۰	۰	۱	۰	۰	۰	۱	۰	۰
۰	۰	۱	۰	۱	۰	۰	۰	۰	۰
۰	۰	۱	۰	۱	۰	۰	۰	۰	۰
۰	۰	۰	۰	۰	۱	۰	۰	۰	۰
۰	۰	۰	۰	۰	۱	۰	۰	۰	۰
۰	۰	۱	۰	۰	۰	۰	۰	۰	۱
۰	۰	۱	۰	۰	۰	۰	۰	۰	۱
۰	۰	۰	۰	۱	۰	۰	۰	۱	۰
۰	۰	۰	۰	۱	۰	۰	۰	۱	۰

برای مثال برای پاسخدهی صحیح به سؤال های ۱ و ۲ که مربوط به طراحی مکعب می باشند، صفات تصویرسازی و سرعت چرخش لازم هستند، برای سؤال های ۳ و ۴ که مربوط به شباهت می باشند، صفات استقرا، دانش واژگانی و انعطاف پذیری؛ برای سؤال های ۵ و ۶ که مربوط به فراختای ارقام می باشند، صفات حافظه فعال و فراختای حافظه؛ برای سؤال های ۷ و ۸ که مربوط به مفاهیم تصویری می باشند، استقرا و اطلاعات عمومی کلامی؛ برای سؤال های ۹ و ۱۰ که مربوط به واژگان می باشند، صفت دانش واژگانی لازم است. برای سؤال های ۱۱ و ۱۲ که مربوط به توالی حرف و عدد می باشند، صفت حافظه فعال؛ برای سؤال های ۱۳ و ۱۴ که مربوط به استدلال تصویری می باشند، صفات استقرا و استدلال عمومی زنجیره ای؛ و در نهایت برای سؤال های ۱۵ و

۱۶ که مربوط به درک مطلب می‌باشند، صفات رشد زبانی و اطلاعات عمومی کلامی لازم می‌باشند.

برای بررسی روایی ماتریس کیو از دو روش استفاده شد. در روش اول از میزان توافق نظر متخصصان استفاده کردیم. بر این اساس از ضریب کاپا برای بررسی توافق بین نظر دو نفر از متخصصان که به بررسی ماتریس پرداخته بودند استفاده کردیم. میزان ضریب کاپا به دست آمده برابر $0/78$ بود که نشان از توافق بالا و قابل قبول دارد. در مرحله دوم از روش دلتا توره و چپو (۲۰۱۶) که یک روش آماری برای بررسی روایی ماتریس کیو در مدل جی دینا هست، استفاده کردیم. این روش میزان درستی ماتریس کیو و اصلاحاتی را که مدل می‌تواند انجام دهد مشخص می‌کند. برای بررسی ماتریس کیو در این پژوهش نیز از این روش استفاده شد و نتایج آن در جدول ۳ ارائه شده است:

جدول ۳. روایی ماتریس کیو

بدون تغییر	اصلاح شده	ورودی ماتریس
$0/98$	$0/2$	ماتریس صحیح

مقدار ستون بدون تغییر برابر $0/98$ است که نشان می‌دهد ۹۸ درصد از خانه‌های ماتریس بدون تغییر باقی مانده‌اند و تنها ۲ درصد از خانه‌ها تغییر کرده‌اند. بررسی موارد اصلاحی پیشنهادی توسط روش نشان داد که این موارد بیشتر ناشی از خطای آماری روش دلتا توره و چپو است که ناشی از بالا بودن تعداد خانه‌های ماتریس کیو ورودی ($n=1971$) است. با توجه به موارد فوق می‌توان گفت که ماتریس به دست آمده از روایی مناسبی برخوردار است.

سؤال دوم: دانش آموزان در چه صفات و مهارت‌هایی به تسلط رسیده‌اند و در چه صفات و مهارت‌هایی به تسلط نرسیده‌اند؟

برای پاسخ به پرسش بالا تحلیل داده‌ها بر اساس مدل‌های تشخیصی شناختی انجام شد.

تحلیل داده‌ها بر اساس مدل عمومی و کلی جی دینا انجام شد؛ اما از آنجا که مدل جی دینا یک مدل کلی هست پارامترهای که برآورد می‌کند زیاد می‌باشند و فرایند تحلیل را پیچیده می‌کند. مدل جی دینا خود مدل کلی هست و احتمال خطا در آن زیاد هست؛ بنابراین مدل‌های محدودتر که از خود این مدل کلی به دست می‌آیند مورد بررسی قرار می‌گیرند. پنج مدل $DINA-DINO-ACDM-LLM-RRUM$ با محدود کردن پارامترهای از مدل جی دینا به دست می‌آیند. از بین مدل‌های ذکر شده یک مدل انتخاب می‌شود. انتخاب مدل‌های محدود بر اساس آماره والد صورت می‌گیرد. آماره والد به بررسی تفاوت مدل‌های محدود شده با مدل جی دینا می‌پردازد. مدلی که در آن آماره والد معنی دار نشود انتخاب می‌شود. با توجه به اطلاعات به دست آمده بر اساس آماره والد^۴ مدل LLM انتخاب شد و تحلیل بر اساس این مدل انجام شد.

جدول ۴. مقایسه شاخص‌های برازش مدل

شاخص	مدل جی دینا	مدل LLM
شاخص اطلاعات آکایک AIC	۱۴۵۴۵۰/۱۷۳	۱۴۵۷۶۵
شاخص اطلاعات بیزی BIC	۱۵۲۱۳۶/۸۶۴	۱۵۱۳۹۹/۴

جدول بالا شاخص‌های برازش برای مدل جی دینا و مدل محدود LLM را نشان می‌دهد. شاخص‌های اطلاعات آکایک^۵ برای مدل جی دینا برابر ۱۴۵۴۵۰/۱۷۳ هست و مقدار شاخص اطلاعات بیزی^۶ برابر ۱۵۲۱۳۶/۸۶۴ است. همچنین جدول بالا شاخص برازش به دست آمده برای مدل LLM را نشان می‌دهد. شاخص اطلاعات آکایک برابر ۱۴۵۷۶۵ و شاخص اطلاعات بیزی برابر ۱۵۱۳۹۹/۴ هست. در مقایسه شاخص‌های برازش هر مدلی که شاخص‌های برازش

1. Additive CDM
2. log-linear model
3. reduced reparametrized unified model
4. The Wald Test
5. Akaike Information Criteria
6. Bayesian Information Criteria

کوچک‌تری باشد، مدل مناسب‌تری است. شاخص اطلاعات آکایک برای مدل جی دینا کوچک‌تر از مدل LLM و شاخص اطلاعات بیزی مدل LLM کوچک‌تر از مدل جی دینا است، اما چون مدل‌های محدود تعداد پارامترهای کم‌تر و خطای کمتری دارند، مدل LLM برای تحلیل داده‌های این پژوهش از بین مدل‌های محدود در مدل کلی جی دینا انتخاب شد. پس از مشخص شدن مدل، به بررسی میزان تسلط آزمودنی‌ها در هر یک از صفات مشخص شده در مرحله قبل پرداخته شد.

جدول ۵. درصد تسلط آزمودنی‌ها در هر یک از توانایی‌های باریک

درصد تسلط	تعریف صفت	صفت
۰/۴۳۳۹	توانایی درک الگوهای پیچیده و اینکه به‌طور ذهنی شبیه‌سازی کند که چه اتفاقی برای آن‌ها می‌افتد وقتی که تبدیل می‌شوند (چرخش، تغییر در اندازه).	تصویرسازی (Vz)
۰/۳۷۲۳	توانایی مشاهده یک پدیده و کشف اصول و قوانین زیربنایی که رفتار آن پدیده را تبیین می‌کند. افرادی که استدلال استقرایی قوی دارند نظم و الگوهای را در موقعیت می‌بینند که برای دیگران ممکن است غیرقابل پیش‌بینی باشد.	استقرا (I)
۰/۳۶۶۷	توانایی جهت‌دهی توجه برای انجام دست‌کاری‌های ساده، ترکیب و تبدیل اطلاعات در حافظه اولیه، درحالی‌که از محرک‌های مزاحم دوری می‌کند و درگیر فرایند جستجوی کنترل‌شده برای اطلاعات در حافظه بلندمدت هست.	حافظه فعال (Mw)
۰/۴۵۷۹	دانشی که در یک فرهنگ ضروری، عملی و ارزشمند تلقی می‌شود و هرکسی باید آن را بداند. این عامل به‌وسیله مقدار دانشی که در حیطه‌های مختلف کسب‌شده، سنجیده می‌شود.	اطلاعات عمومی کلامی (Ko)
۰/۵۶۵۳	دانش تعریف لغات و مفاهیمی که زیربنای آن‌ها را تشکیل می‌دهد. این عامل بر تعریف لغات به‌تنهایی تأکید دارد. دانش فرهنگ لغتی است به‌گونه‌ای.	دانش واژگانی (VI)
۰/۳۵۸۴	توانایی تشخیص یک‌شکل یا الگو که در یک الگوی پیچیده‌تر ترکیب شده است، یا الگوی که به‌هم‌ریخته است، وقتی از قبل الگو را می‌شناسیم.	انعطاف‌پذیری در مشابهت (Cf)
۰/۴۶۴	توانایی کدگذاری اطلاعات، نگه‌داری آن در حافظه اولیه و بازتولید فوری اطلاعات به همان ترتیبی که ارائه شده بود.	فراخوانی حافظه (Ms)
۰/۳۶۴۷	فهم کلی از زبان در سطح لغات، اصطلاحات و جملات.	رشد زبانی (LD)

درصد تسلط	تعریف صفت	صفت
۰/۴۲۸۲	توانایی استدلال منطقی، استفاده از الگوهای شناخته‌شده. این توانایی همچنین با استدلال قیاسی و کاربرد قوانین شناخته می‌شود. درحالی‌که استقرا توانای استفاده حقایق شناخته‌شده برای کشف اصول جدید است، استدلال عمومی توانایی استفاده از اصول شناخته‌شده و موجود در یک یا چند مرحله منطقی برای کشف حقایق جدید و حل مشکلات است. یک مثال برای استدلال عمومی قاضی یا هیئت‌منصفه‌ای است که بر اساس حقایق ارائه‌شده و قوانین مربوط تصمیم‌گیری می‌کند.	استدلال عمومی زنجیره‌ای (RG)

جدول شماره ۵ احتمال تسلط در هر یک از مؤلفه‌های به‌دست آمده را نشان می‌دهد. میزان تسلط برای مؤلفه‌ها به ترتیب دانش واژگانی با احتمال (۰/۵۶۵۳)، فراخنای حافظه (۰/۴۶۴)، اطلاعات عمومی کلامی (۰/۴۵۷۹)، تصویرسازی (۰/۴۳۳۹)، اطلاعات عمومی زنجیره‌ای (۰/۴۲۸۲)، استقرا (۰/۳۷۲۳)، حافظه فعال (۰/۳۶۶۷)، رشد زبانی (۰/۳۶۴۷)، انعطاف‌پذیری در مشابَهت (۰/۳۵۸۴) است.

هر مدل شناختی دارای پارامترهای می‌باشد که باید برآورد شوند. پارامترهای مدل‌های شناختی متفاوت می‌باشند. مدل شناختی جی‌دینا دو پارامتر حدس و لغزش دارد. پارامتر حدس به احتمال پاسخگویی صحیح به یک سؤال بدون تسلط به مهارت‌های لازم برای پاسخگویی به آن سؤال اشاره دارد و پارامتر لغزش به احتمال ارتکاب به خطا زمانی که مهارت‌های لازم برای پاسخگویی به یک سؤال را داشته باشند، اشاره دارد. در جدول مقدار هر کدام از این دو پارامتر حدس و لغزش با خطای استاندارد آنها ارائه گردیده‌است. با توجه به حجم بالای جدول فقط پارامترهای مربوط به سؤالات ابتدایی و انتهایی هر خرده‌مقیاس گزارش می‌شود.

جدول ۶. پارامترهای حدس و لغزش برای سؤالات آزمون هوش و کسلر

خطای استاندارد لغزش	خطای استاندارد حدس	لغزش	حدس	
۰/۰۰۰۵	۰/۰۰۳۸	۰/۰۰۰۱	۰/۹۹۴۳	۱Bd
۰/۰۰۰۵	۰/۰۰۰۹	۰/۰۰۰۱	۰/۹۷۴	۲Bd
۰/۰۰۰۵	۰/۰۱۵۴	۰/۰۰۰۱	۰/۹۰۹۸	۳Bd
۰/۰۰۸۷	۰/۰۲۶۱	۰/۰۱۳۷	۰/۶۸۷۸	۴Bd
۰/۰۱۴۸	۰/۰۲۸۱	۰/۰۵۳۸	۰/۴۵۵۳	۵Bd
۰/۰۲۹۳	۰/۰۰۰۷	۰/۸۴۱۳	۰/۰۰۰۱	۱۱Bd
۰/۰۲۲۶	۰/۰۰۰۷	۰/۹۰۷۸	۰/۰۰۰۱	۱۲Bd
۰/۰۱۹۵	۰/۰۰۰۷	۰/۹۴۰۸	۰/۰۰۰۱	۱۳Bd
۰/۰۰۹۲	۰/۰۰۰۷	۰/۹۸۳۶	۰/۰۰۰۱	۱۴Bd
۰/۰۰۲۵	۰/۰۳۷۹	۰/۰۰۳۲	۰/۶۲۴۳	۱S
۰/۰۰۲۳	۰/۰۳۵۵	۰/۰۰۳۲	۰/۷۱۴۲	۲S
۰/۰۰۰۸	۰/۰۲۷۵	۰/۰۰۰۸	۰/۸۲۹۷	۳S
۰/۰۰۲۴	۰/۰۳۰۸	۰/۰۰۲۸	۰/۷۸۲۵	۴S
۰/۰۰۰۲	۰/۰۳۳۴	۰/۰۰۳۵	۰/۶۷۹۵	۵S
۰/۰۰۱۸	۰/۰۳۶۳	۰/۰۰۳۴	۰/۴۵۰۴	۶S
۰/۰۷۲۴	۰/۰۰۴۶	۰/۵۶۹۸	۰/۰۰۸۸	۱۹S
۰/۰۵۹۵	۰/۰۰۴۶	۰/۸۹۸۴	۰/۰۰۰۱	۲۰S
۰/۰۷۰۱	۰/۰۰۳۸	۰/۸۹۱۸	۰/۰۰۰۱	۲۱S
۰/۰۷۸۳	۰/۰۰۷۲	۰/۷۹۸۸	۰/۰۰۰۱	۲۲S
۰/۰۳۲۷	۰/۰۱۱	۰/۹۷۹۵	۰/۰۰۰۱	۲۳S
۰/۰۹۶۳	۰/۰۸۷۸	۰/۰۰۲۱	۰/۹۹۸۶	۱۲۱Dsf
۰/۰۱۴۷	۰/۰۲۰۲	۰/۰۰۰۲	۰/۹۴۶۶	۱۲۲Dsf
۰/۰۰۰۶	۰/۰۱۸۵	۰/۰۰۰۳	۰/۹۵۰۳	۲۲۲Dsf
۰/۰۰۰۷	۰/۰۴۲۱	۰/۰۰۱	۰/۶۰۶۲	۱۲۳Dsf
۰/۰۰۲۶	۰/۰۴۱۷	۰/۰۰۰۶	۰/۳۹۰۸	۲۲۳Dsf
۰/۱۲۱۵	۰/۰۰۱۵	۰/۵۲۸۶	۰/۰۰۰۱	۱۲۴Dsf
۰/۰۶۸۷	۰/۰۰۰۳	۰/۷۹۱۲	۰/۰۰۰۲	۲۲۴Dsf

خطای استاندارد لغزش	خطای استاندارد حدس	لغزش	حدس	
۰/۱۱۸	۰/۰۰۰۶	۰/۸۷۵۵	۰/۰۰۰۱	۱t۸Dsf
۰/۰۹۶۶	۰/۰۰۰۵	۰/۹۱۸۸	۰/۰۰۰۱	۲t۸Dsf
۰/۰۰۰۱	۰/۰۴۶۳	۰/۰۰۰۱	۰/۸۳۸۶	۱t۱Dsb
۰/۰۰۰۳	۰/۰۴۷۸	۰/۰۰۰۱	۰/۸۱۸۷	۲t۱Dsb
۰/۰۰۱۴	۰/۰۶۱۵	۰/۰۰۰۱	۰/۷۴۵۳	۱t۲Dsb
۰/۰۰۰۴	۰/۰۶۱	۰/۰۰۰۴	۰/۷۳۴۶	۲t۲Dsb
۰/۰۱۵۲	۰/۰۳۰۷	۰/۰۵۰۷	۰/۱۳۲۵	۱t۳Dsb
۰/۰۱	۰/۰۲۸۳	۰/۰۳۰۲	۰/۱۱۹۸	۲t۳Dsb
۰/۱۹۷۳	۰/۰۰۱۲	۰/۹۹۵۷	۰/۰۰۰۱	۱tvDsb
۰/۲۲۴۴	۰/۰۰۲۱	۰/۹۹۵۱	۰/۰۰۰۱	۲tvDsb
۰/۰۴۹۲	۰/۰۰۰۸	۰/۹۷۹۴	۰/۰۰۰۱	۱t۸Dsb
۰/۰۰۰۹	۰/۰۱۷۴	۰/۰۰۰۴	۰/۹۲۸۸	۱Pc
۰/۰۰۴۳	۰/۰۱۹۷	۰/۰۰۰۱	۰/۹۰۷۸	۲Pc
۰/۰۰۱۸	۰/۰۲۱۷	۰/۰۰۰۱	۰/۹۱۳۳	۳Pc
۰/۰۰۰۸	۰/۰۲۱۲	۰/۰۰۰۱	۰/۸۸۷۵	۴Pc
۰/۰۰۲۲	۰/۰۳۱۵	۰/۰۰۳۲	۰/۶۸۵۱	۵Pc
۰/۰۳۵۵	۰/۰۰۸۴	۰/۹۰۳۲	۰/۰۰۰۱	۲۴Pc
۰/۰۴۶۱	۰/۰۰۳۸	۰/۸۴۴۲	۰/۰۰۳۸	۲۵Pc
۰/۰۲۲۲	۰/۰۰۱۸	۰/۹۵۴۳	۰/۰۰۰۷	۲۶Pc
۰/۰۱۳۶	۰/۰۰۳۴	۰/۹۷۹۲	۰/۰۰۲۷	۲۷Pc
۰/۰۱۹۹	۰/۰۰۱۷	۰/۹۷۲۹	۰/۰۰۱۴	۲۸Pc
۰/۰۰۰۶	۰/۰۰۰۴	۰/۰۰۰۱	۰/۹۹۵۶	۲۷
۰/۰۰۰۸	۰/۰۰۳۸	۰/۰۰۰۱	۰/۹۹۴۱	۳۷
۰/۰۰۲۷	۰/۰۰۵۷	۰/۰۰۱۹	۰/۹۸۹۵	۴۷
۰/۰۰۵۱	۰/۰۱۷۵	۰/۰۰۷۴	۰/۸۹۴۵	۵۷
۰/۰۲۸۲	۰/۰۰۱۸	۰/۸۳۲	۰/۰۰۱۵	۳۰۷
۰/۰۴۱۳	۰/۰۰۹۵	۰/۴۲۶۴	۰/۰۲۲۸	۳۱۷
۰/۰۳۸۹	۰/۰۰۲۱	۰/۶۵۵۱	۰/۰۰۱۶	۳۲۷

خطای استاندارد لغزش	خطای استاندارد حدس	لغزش	حدس	
۰/۰۰۵۷	۰/۰۰۱۱	۰/۹۹۲۴	۰/۰۰۰۱	۳۳۷
۰/۰۰۶۹	۰/۰۰۰۸	۰/۹۸۸۷	۰/۰۰۰۱	۳۴۷
۰/۰۳۹۶	۰/۰۱۰۲	۰/۶۱۹	۰/۰۲۸۵	۳۵۷
۰/۰۰۶۶	۰/۰۰۰۸	۰/۹۹۱۸	۰/۰۰۰۱	۳۶۷
۰/۰۰۲۴	۰/۰۵۳۵	۰/۰۰۲۹	۰/۷۷۱	۱۱Lns
۰/۰۰۰۷	۰/۰۶۷۷	۰/۰۰۰۱	۰/۷۵۸	۱۲Lns
۰/۰۰۷۱	۰/۰۵۳۸	۰/۰۱۷۵	۰/۷۳۷۳	۱۳Lns
۰/۰۰۳۵	۰/۰۶۶۷	۰/۰۰۲۸	۰/۷۴۴۶	۲۱Lns
۰/۰۰۳۴	۰/۰۵۴۶	۰/۰۰۰۴	۰/۷۰۱	۲۲Lns
۰/۰۰۳۹	۰/۰۵۳	۰/۰۰۴۳	۰/۶۹۷۷	۲۳Lns
۰/۰۱۵۶	۰/۰۰۰۷	۰/۹۶۵۸	۰/۰۰۰۱	۸۱Lns
۰/۰۱۳۹	۰/۰۰۰۷	۰/۹۷۹۶	۰/۰۰۰۱	۸۲Lns
۰/۰۰۶۶	۰/۰۰۰۷	۰/۹۹۱۱	۰/۰۰۰۱	۸۳Lns
۰/۰۱۱۶	۰/۰۰۰۷	۰/۹۸۸۴	۰/۰۰۰۱	۹۱Lns
۰/۱۵۷۲	۰/۰۰۳۲	۰/۹۹۷۱	۰/۰۰۰۱	۹۲Lns
۰/۰۷۰۹	۰/۰۰۱۵	۰/۹۹۸	۰/۰۰۰۱	۹۳Lns
۰/۱۵۷۳	۰/۰۰۳۲	۰/۹۹۷۱	۰/۰۰۰۱	۱۰۱Lns
۰/۰۷۰۹	۰/۰۰۱۵	۰/۹۹۸	۰/۰۰۰۱	۱۰۲Lns
۰/۰۷۰۹	۰/۰۰۱۵	۰/۹۹۸	۰/۰۰۰۱	۱۰۳Lns
۰/۰۰۴۵	۰/۰۰۳۵	۰/۰۰۰۲	۰/۹۹۷۹	۱Mr
۰/۰۱۵۸	۰/۰۰۳۴	۰/۰۰۰۴	۰/۹۹۷۸	۲Mr
۰/۰۱۲۱	۰/۰۱۰۷	۰/۰۰۰۱	۰/۹۸۹۵	۳Mr
۰/۰۰۷۳	۰/۰۰۱۶	۰/۰۰۰۱	۰/۹۴۲۲	۴Mr
۰/۰۰۳۳	۰/۰۱۶۹	۰/۰۰۰۱	۰/۹۳۸۲	۵Mr
۰/۰۶۳۱	۰/۰۰۱۷	۰/۶۹۵۲	۰/۰۰۱۸	۳۰Mr
۰/۰۵۸۸	۰/۰۰۵۲	۰/۷۲۷	۰/۰۰۷۲	۳۱Mr
۰/۰۴۹۶	۰/۰۰۵۳	۰/۸۵۸۵	۰/۰۰۶۲	۳۲Mr
۰/۰۴۴	۰/۰۰۴۹	۰/۹۱۱۸	۰/۰۰۰۵	۳۳Mr

خطای استاندارد لغزش	خطای استاندارد حدس	لغزش	حدس	
۰/۰۳۹۵	۰/۰۰۱	۰/۹۲۴۷	۰/۰۰۰۶	۳۴Mr
۰/۰۴۱۲	۰/۰۰۴۸	۰/۹۰۱۹	۰/۰۰۲۱	۳۵Mr
۰/۰۰۰۹	۰/۰۱۴۲	۰/۰۰۰۱	۰/۹۷۶۶	۱C
۰/۰۰۰۵	۰/۰۳۳۴	۰/۰۰۸۵	۰/۷۰۲۸	۲C
۰/۰۰۲۵	۰/۰۳۱۲	۰/۰۰۴۳	۰/۷۴۲۳	۳C
۰/۰۰۰۶	۰/۰۳۴۴	۰/۰۱۳۲	۰/۶۷۴۹	۴C
۰/۰۰۵۲	۰/۰۲۹۸	۰/۰۱۲۳	۰/۷۶۹۷	۵C
۰/۰۵۴۲	۰/۰۰۱۵	۰/۷۳۸	۰/۰۰۲	۱۷C
۰/۰۴۴	۰/۰۰۱۸	۰/۸۷۶۵	۰/۰۰۰۱	۱۸C
۰/۰۵۴۱	۰/۰۰۸۸	۰/۸۸۹۷	۰/۰۰۰۱	۱۹C
۰/۰۳۵۱	۰/۰۰۸۹	۰/۹۲۷۱	۰/۰۰۰۱	۲۰C
۰/۰۳۴۳	۰/۰۰۲۵	۰/۹۵۸۶	۰/۰۰۰۱	۲۱C

جدول ۶ پارامترهای لغزش^۱ و حدس^۲ را برای هر سؤال بر اساس مدل LLM را نشان می‌دهند. در خرده مقیاس طراحی مکعب بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به سؤال ۱۴ هست، برای خرده مقیاس شباهت‌ها بالاترین احتمال حدس مربوط به سؤال ۳ و بالاترین احتمال لغزش مربوط به سؤال ۲۳ هست، در خرده مقیاس فراخنای ارقام، بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به سؤال ۷ هست، در خرده مقیاس مفاهیم تصویری، بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به سؤال ۲۷، برای خرده مقیاس واژگان، بالاترین احتمال حدس مربوط به سؤال ۲ و بالاترین احتمال لغزش مربوط به سؤال ۳۶ هست، برای توالی حرف و عدد، بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به سؤال ۱۰ هست. در خرده مقیاس استدلال تصویری، بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به

1. Slip
2. Guess

سؤال ۳۴ هست و نهایتاً برای خرده مقیاس درک مطلب بالاترین احتمال حدس مربوط به سؤال ۱ و بالاترین احتمال لغزش مربوط به سؤال ۲۱ هست.

نتیجه‌گیری

هدف از پژوهش حاضر شناسایی صفات زیربنایی آزمون هوش و کسلر چهار بر اساس توانایی‌های باریک نظریه کتل - هورن - کارول با استفاده از مدل‌های تشخیصی شناختی بود. نظریه کتل - هورن - کارول حدود ۷۰ توانایی باریک را شناسایی کرده است. در این پژوهش نیز به مطالعه و ارتباط این توانایی‌های باریک با سؤالات آزمون هوش و کسلر کودکان نسخه چهار پرداخته شد و پس از مطالعه مبانی نظری و استفاده از نظر متخصصان در این زمینه حدود ۹ توانایی باریک برای سؤالات خرده مقیاس‌های اصلی آزمون هوش و کسلر شناسایی شد و ماتریس کیو بر اساس این ۹ توانایی و صفت تشکیل شد.

از بین صفات شناسایی شده برای آزمون و کسلر، صفت تصویرسازی برای پاسخ‌دهی به سؤالات خرده مقیاس طراحی مکعب، صفت استقرا برای پاسخ‌دهی سؤالات خرده مقیاس‌های شباهت‌ها، مفاهیم تصویری و استدلال تصویری، صفت حافظه فعال برای خرده مقیاس‌های فراخنای ارقام و توالی حرف و عدد، صفت اطلاعات عمومی کلامی برای خرده مقیاس‌های مفاهیم تصویری و درک مطلب، دانش واژگانی برای خرده مقیاس‌های شباهت‌ها و واژگان، صفت انعطاف‌پذیری در مشابهت برای خرده مقیاس شباهت‌ها، صفت فراخنای حافظه برای خرده مقیاس فراخنای ارقام، صفت رشد زبانی برای خرده مقیاس درک مطلب و درنهایت صفت استدلال عمومی زنجیره‌ای برای خرده مقیاس استدلال تصویری شناسایی شدند.

پس از شناسایی صفات اولیه و تأیید روایی ماتریس کیو، تحلیل داده‌ها با استفاده از مدل عمومی جی‌دینا انجام شد، شاخص‌های برازش مدل با داده‌ها مانند شاخص اطلاعات آکاپک و شاخص اطلاعات بیزی نشان از برازش خوب دارد. از آنجا که مدل جی‌دینا یک مدل کلی و

عمومی هست، باید از مدل‌های محدود در این مدل کلی استفاده کرد. نتایج آزمون والد نشان داد که از بین مدل‌های محدود، مدل LLM بهترین مدل هست و تحلیل بر اساس این مدل انجام شد. شاخص‌های برازش محاسبه شده نیز برای مدل LLM نشان از برازش خوب این مدل با داده‌ها داشت. برازش خوب مدل شناختی با داده‌ها برای شناسایی صفت‌های زیربنایی آزمون هوش و کسلر همسو با نتایج تحقیقات قبلی (تمپلین و ایوی، ۲۰۰۶؛ لی، ۲۰۱۱؛ وای، ۲۰۱۲؛ گارسیا، دیاز، دلا توره، ۲۰۱۴؛ مینایی، ۱۳۹۱؛ کبیری، ۱۳۹۳؛ مقدم، ۱۳۹۵) هست که در آن نشان داده شده بود که مدل‌های شناختی به خوبی می‌توانند صفت‌های زیربنایی آزمون‌ها را شناسایی کنند.

سپس وضعیت تسلط دانش آموزان ایرانی در هر یک از صفات شناسایی شده بررسی شد. نتایج میزان تسلط دانش آموزان در جدول شماره ۵ آورده شده است. بالاترین میزان تسلط برای صفت دانش واژگانی و پس از آن فراخنای حافظه هست و پایین‌ترین میزان تسلط نیز مربوط به صفت انعطاف‌پذیری در مشابهت هست. صفت دانش واژگانی بر دانش تعریف لغات و مفاهیمی که زیربنای آن را تشکیل می‌دهند، تأکید دارد و به نوعی دانش فرهنگ لغتی به حساب می‌آید. فراخنای حافظه به توانایی کدگذاری اطلاعات، نگهداری آن‌ها در حافظه اولیه و بازتولید اطلاعات به همان ترتیبی که ارائه شده است، اشاره دارد. انعطاف‌پذیری در مشابهت به توانایی تشخیص یک الگو یا یک شکل که در یک الگوی پیچیده‌تر ترکیب شده‌اند، یا تشخیص یک الگو از بین الگوهای به هم ریخته‌ای که قبلاً می‌شناختیم، اشاره دارند. دانش واژگانی خود به عامل گسترده هوش متبلور مربوط می‌شود که به نوعی یک عامل اکتسابی در نظر گرفته می‌شود. اینکه دانش آموزان در این عامل بالاترین میزان تسلط را دارند شاید به دلیل تأکیدی باشد که در آموزش ما بر روی حفظیات انجام می‌شود، باشد. در همین رابطه دومین صفتی که بالاترین میزان تسلط را دانش آموزان در آن دارند، فراخنای حافظه هست که خود مربوط به عامل گسترده‌تر حافظه کوتاه مدت هست. از سوی دیگر، پایین‌ترین میزان تسلط مربوط به صفت انعطاف‌پذیری در

مشابهت هست، صفتی که زیرمجموعه عامل گسترده پردازش تصویری قرار می‌گیرد، تسلط پایین دانش‌آموزان در این صفت را می‌توان به توجه کمتر به فرایند حل مسئله در امر آموزش دانست و همچنین به دلیل سن کم دانش‌آموزان و عدم رشد کافی فرایند استدلال انتزاعی در آن‌ها دانست.

بنابراین با شناسایی صفات و مهارت‌های زیربنایی شناختی که افراد در آن توانایی و یا ضعف دارند، می‌توان با تقویت این مؤلفه‌ها و صفات کودکان را برای مراحل بعدی فعالیت‌های آموزشی وزندگی روزمره آماده کرد و پیش‌بینی که بر اساس نتایج این آزمون‌ها انجام می‌شود، دقیق‌تر و کامل‌تر باشد. علاوه بر این با شناسایی این صفات زیربنایی و نقاط قوت و ضعف کودکان در آن می‌توان به تبیینی برای مشکلات و اختلالات یادگیری و ناتوانایی‌های شناختی آنان رسید. از آنجایی که یکی از کاربردهای آزمون هوش و کسلر در تشخیص و غربالگری کودکان دارای اختلال یادگیری هست، شناسایی دقیق صفت و مهارت‌های که این کودکان در آن مشکل دارند، کمک بسیار زیادی به تشخیص درست نوع اختلال یادگیری، طراحی برنامه‌های آموزشی و مداخله‌ای مناسب برای این کودکان داشته باشد. درنهایت می‌توان گفت توجه به آزمون‌های هوش و شناسایی مهارت‌های زیربنایی این آزمون‌ها می‌تواند نقش مهم و ویژه‌ای چه از نظر کاربردهای آموزشی و چه از نظر رشد فردی و اجتماعی داشته باشد.

از نظر توجه به مبانی نظری و نظریه‌های زیربنایی برای ساخت آزمون، توجه به نظریه هوش سلسله مراتبی کتل-هورن-کارول در ساخت آزمون و ارتباط دادن آزمون هوش و کسلر به این نظریه می‌تواند در فراهم کردن تبیین‌های دقیق‌تر برای نتایج به‌دست آمده و فراهم آوردن چارچوب‌های نظری کمک شایانی کند. امروزه تقریباً اکثر آزمون‌های هوش بر اساس نظریه کتل-هورن-کارول ساخته و یا سعی می‌کنند که در ساخت آزمون از این نظریه استفاده کنند. آزمون هوش و کسلر چهار نیز به‌عنوان یکی از آزمون‌های توانمند در سنجش هوش نباید از این امر مستثنا باشد و از این نظریه جدید در هوش استفاده کند.

در مورد پارامترهای سؤال می توان گفت در اکثر خرده مقیاس احتمال حدس (احتمال پاسخگویی صحیح به یک سؤال بدون تسلط به مهارت های لازم برای پاسخگویی به آن سؤال) در سؤالات آسان (سؤالات ابتدایی هر خرده مقیاس) بالا و هر چه به سمت سؤالات دشوارتر پیش می رویم احتمال حدس پایین می آید. در مورد پارامتر لغزش (احتمال ارتکاب به خطا زمانی که مهارت های لازم برای پاسخگویی به یک سؤال را داشته باشند) در سؤالات آسان هر خرده مقیاس پایین و در سؤالات دشوار هر خرده مقیاس بالا می رود. از محدودیت های این پژوهش به تعداد زیاد سؤالات و تعداد زیاد صفات زیر بنایی نظریه کتل -هورن -کارول اشاره کرد.

منابع

- افضلی، افشین. (۱۳۹۳). تدوین و ارزشیابی مدل تشخیصی شناختی ریاضیات پایه اول دبیرستان با استفاده از روش سلسله مراتبی صفات AHM. پایان نامه دکتری، دانشگاه علامه طباطبایی. کبیری، مسعود. (۱۳۹۳). کاربرد سنجش شناختی تشخیصی به منظور تعیین مهارت های کسب شده علوم تجربی در دانش آموزان پایه تحصیلی هشتم ایران بر اساس داده های تیمز ۲۰۱۱. پایان نامه دکتری، دانشگاه تهران.
- صادقی، احمد، ربیعی، محمد، عابدی، محمدرضا. (۱۳۹۰). روا سازی و اعتبار یابی چهارمین ویرایش مقیاس هوش و کسلر. *مجله روانشناسی تحولی*، دوره ۷، شماره ۲۸.
- محسن پور، مریم. (۱۳۹۳). طراحی، ساخت و اعتبار یابی آزمون شناختی تشخیصی سواد ریاضی در پایان دوره آموزشی عمومی جهت تدوین الگوی ارتقای کیفیت آن. پایان نامه دکتری، دانشگاه تهران.
- مینایی، اصغر. (۱۳۹۱). مدل پردازی تشخیصی شناختی سؤال های ریاضی تیمز ۲۰۰۷ در دانش آموزان پایه هشتم با استفاده از مدل یکپارچه با پارامتر پردازی مجدد و مقایسه مهارت های ریاضی دانش آموزان دختر و پسر. پایان نامه دکتری، دانشگاه علامه طباطبایی.

مقدم، اعظم. (۱۳۹۴). کاربرد مدل‌های تشخیصی شناختی به منظور تعیین مهارت‌های زیربنایی عملکردی داوطلبان در آزمون ورودی زبان انگلیسی دوره دکتری. پایان‌نامه دکتری، دانشگاه علامه طباطبائی.

- De la Torre, J., van der Ark, L. A., & Rossi, G. (2015). Analysis of Clinical Data from Cognitive Diagnosis Modeling Framework. *Measurement and Evaluation in Counseling and Development*, 0748175615569110.
- De la Torre, J. & Chiu, C. Y. (2016). A General method of empirical q matrix validation. *Psychometrika* - vol. 81, no. 2, 253-273
- Flanagan, D. P., & Harrison, P. L. (Eds.). (2012). *Contemporary intellectual assessment: Theories, tests, and issues*. Guilford Press.
- García, P. E., Olea Díaz, J., & Torre, J. D. L. (2014). Application of cognitive diagnosis models to competency-based situational judgment tests. *Psicothema*.
- George, A. C., & Robitzsch, A. (2014). Multiple group cognitive diagnosis models, with an emphasis on differential item functioning. *Psychological Test and Assessment Modeling*, 56(4), 405-432.
- Li, H. (2011). A cognitive diagnostic analysis of the MELAB reading test. *Spain Working Fellow Papers in Second or Foreign Language Assessment*, 9, 17-46.
- Mason, E & Wilcox, K(2009). Intelligence: an overview. www.education.com
- Reverte, I., Golay, P., Favez, N., Rossier, J., & Lecerf, T. (2015). Testing for multi group invariance of the WISC-IV structure across France and Switzerland: Standard and CHC models. *Learning and individual differences*, 40, 127-133.
- Reynolds, C. R. J. Vannest, J.K & Fletcher-Janzen, E. (2013). *Encyclopedia of special education: A reference for the education of children, adolescents, and adults with disabilities and other exceptional individuals* (Vol. 3). John Wiley & Sons.
- Richerson, L. P., Watkins, M. W., & Beaujean, A. A. (2014). Longitudinal Invariance of the Wechsler Intelligence Scale for Children-Fourth Edition in a Referral Sample. *Journal of Psychoeducational Assessment*, 0734282914538802.
- Rupp A. & Templin L. (2008). Unique characteristic of diagnostic models: A comprehensive review of the current state of art, measurement: interdisciplinary research and perspective. 6:4, 219-262.

- Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. *Contemporary intellectual assessment: Theories, tests, and, (3rd)*, 99-144.
- Su, Y. L. (2013). Cognitive diagnostic analysis using hierarchically structured skills.
- Templin, J., & Ivie, J. L. (2006). 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*.
- Van Aken, L, van der Heijden, T, van der Veld, W .Hermans, L, Kessels, R, and I. M. Egger, J. (2015). Representation of the Cattell–Horn–Carroll Theory of Cognitive Abilities in the Factor Structure of the Dutch-Language Version of the WAIS-IV. Assessment 1–9, DOI: 10.1177/1073191115607973 asm.sagepub.com
- Wahlstrom, D., Breaux, K. C., Zhu, J., & Weiss, L. G. (2012). The Wechsler preschool and primary scale of intelligence–third edition, the Wechsler intelligence scale for children–fourth edition, and the Wechsler individual achievement test–. *Contemporary intellectual assessment: Theories, tests, and issues*, 224-248. Guilford Press.
- Yi, Y. (2013). *Implementing a cognitive diagnostic assessment in an institutional test: a new networking model in language testing and experiment with a new psychometric model and task type* .Doctoral dissertation, University of Illinois at Urbana-Champaign.