فصلنامه اندازه‌گیری تربیتی
دانشگاه علامه طباطبائی
سال هشتم، شماره ۳۲، تابستان ۹۷، ص ۱۶۳ تا ۲۱۹

Quarterly of Educational Measurement
Allameh Tabataba'i University
Vol. 8, No. 32, Summer 2018

تحلیل بُعدیّت و مدل‌سازی ساختار توانایی‌های شناختی به روش نظریّه سؤال-پاسخ چند بُعدی و مبتنی بر شاخص‌های وابستگی موضعی[۱]

محسن یزدان فر[۲]، منیجه شهنی ییلاق[۳]، علیرضا حاجی یخچالی[۴]، سیروس عالی‌پور بیرگانی[۵]

تاریخ دریافت: ۹۷/۰۳/۰۲
تاریخ پذیرش: ۹۷/۰۶/۲۸

چکیده

تحلیل بُعدیّت (dimensionality analysis) یا مشخص کردن تعداد خصیصه‌های مکنون زیربنایی تاثیر گذار بر داده‌های حاصل از اجرای یک مقیاس یا آزمون و بررسی فرضیّه‌های استقلال موضعی (local independence) و یکنوایی (monotonicity)، از پیش‌فرض‌های مهم اعمال یک مدل سؤال-پاسخ بر مجموعه‌ای از داده‌ها می‌باشند. در این پژوهش، پس از توضیح مختصر شاخص‌های وابستگی موضعی (local dependence) و مدل‌های آن، به منظور مدل‌سازی ساختار توانایی‌های شناختی بر اساس داده‌های حاصل از اجرای مقیاس سنجش توانایی‌های شناختی روی یک نمونه ۱۱۰۵ نفری از دانش‌آموزان پایه نهم تحصیلی شهر اهواز که به روش نمونه‌گیری تصادفی طبقه‌ای نسبتی انتخاب شده بودند، پس از بررسی فرضیّه استقلال موضعی از طریق شاخص‌های وابستگی موضعی، فرضیّه یکنوایی و برآورد پارامترهای دو مدل تک بُعدی و دو بُعدی و مقایسه آن‌ها با هم، بُعدیّت ساختار توانایی‌های شناختی مشخص شد. نتایج نشان داد مدل وابستگی‌های مشاهده شده بین سؤال‌ها، مدل وابستگی موضعی اساسی می‌باشد و مقایسه شاخص‌های برازش مدل‌های تک بُعدی و دو بُعدی پاسخ مدرّج شده (graded response) نشان داد که مدل دو بُعدی نسبت به مدل تک بُعدی از برازش بهتری با داده‌ها برخوردار می‌باشد . در ادامه از طریق به-

---

۱. این مقاله برگرفته از پایان‌نامه دکتری و با حمایت ستاد راهبری توسعه علوم و فنّاوری‌های شناختی انجام گرفته است.

۲. دانشجوی دکتری روان‌شناسی تربیتی، دانشگاه شهید چمران اهواز، اهواز، ایران (نویسنده مسئول).
mnyazdan@gmail.com

۳. استاد روان‌شناسی تربیتی، دانشگاه شهید چمران اهواز، اهواز، ایران.

۴. استادیار روان‌شناسی تربیتی، دانشگاه شهید چمران اهواز، اهواز، ایران.

۵. دانشیار روان‌شناسی تربیتی، دانشگاه شهید چمران اهواز، اهواز، ایران.

کارگیری مفهوم بُعدیّت فضای عاملی، نوعی مقایسه بین دو نظریّه کلاسیک (CTT) آزمون و نظریّه سوال-پاسخ چند بُعدی (MIRT) انجام گرفت و نوعی همگرایی در نتایج به‌دست آمد. نتایج تحلیل بُعدیّت نشان داد ساختار توانایی‌های شناختی دو بُعدی می‌باشد؛ بُعد اوّل شناخت غیراجتماعی و بُعد دوّم شناخت اجتماعی نام‌گذاری شد و مشاهده شد این یافته‌ها با نتایج موجود در حوزه علوم عصب‌شناختی همگرایی دارند.

واژگان کلیدی: بُعدیّت، بُعدیّت فضای عاملی، نظریّه سؤال-پاسخ چند بُعدی، شاخص‌های وابستگی موضعی، توانایی‌های شناختی، شناخت اجتماعی، شناخت غیراجتماعی.

مقدمه

امروزه مدل‌های نظریّه سؤال-پاسخ[1] (IRT)، به طور گسترده‌ای در ساخت و تحلیل ساختار آزمون‌های روان‌شناختی مورد استفاده قرار می‌گیرند (سن، کوهن و کیم[2]، ۲۰۱۳). از نظر ادواردز، هاتس و کای[3] (۲۰۱۸)، نظریّه سؤال-پاسخ مجموعه‌ای از مدل‌های اندازه‌گیری متغیر نهفته[4] را شامل می‌شود که تلاش می‌کنند پاسخ‌های آزمودنی‌ها به مجموعه‌ای از سؤال‌های یک آزمون را به عنوان تابعی از پارامترهای مربوط به سؤال‌ها و پارامترهای مربوط به آزمودنی‌ها (اشخاص)، توضیح دهند. انتشار چندین مقاله توسط لُرد[5] در فاصله سال‌های ۱۹۵۲ تا ۱۹۵۳ زمینه انتقال از نظریّه کلاسیک آزمون[6] (CTT) به نظریّه مُدرن آزمون (IRT)، که مهم‌ترین مشخصّه آن قدرت مدل‌سازی پاسخ‌های آزمودنی‌ها یا نمره‌ها در سطح سؤال‌ها (ماده‌ها) می‌باشد را فراهم کرد (هامبلتون، وان در لیندن و ولز[7]، ۲۰۱۰). نظریّه سؤال-پاسخ که در ابتدا نظریّه خصیصه مکنون[8] و بعداً نظریّه منحنی ویژگی‌های سؤال[9]، نامیده شد، نظریّه‌ای است که بین پاسخ یا نمره سؤال با خصیصه یا خصایص مکنون اندازه‌گیری شده توسط آزمون، از طریق توابع سؤال-پاسخ، یک رابطه را ایجاد می‌کند (هامبلتون و همکاران، ۲۰۱۰).

1. item response theory (IRT) models
2. Sen, Cohen & Kim
3. Edwards, Houts & Cai
4. latent variable
5. Lord
6. classic test theory
7. Hambleton, Van der Linden & Wells
8. latent trait theory
9. item characteristic curve theory

بررسی مدل‌های مورد استفاده در تحلیل داده‌ها به کمک نظریّه سؤال-پاسخ، نشان می‌دهد که این مدل‌ها می‌توانند یک مدل نظریّه سؤال-پاسخ تک بُعدی[1] (UIRT)، یا یک مدل نظریّه سؤال-پاسخ چند بُعدی[2] (MIRT) باشند. شکل کلی یک مدل چند بُعدی نظریّه سؤال-پاسخ را می‌توان به صورت ریاضی و بر اساس نمادگذاری ریکاس[3] (۲۰۰۹) با رابطه (۱) نمایش داد:

$$P(U = u \mid \boldsymbol{\theta}) = f(\boldsymbol{\theta}, \boldsymbol{\eta}, u) \qquad (۱)$$

در رابطه (۱)، حروف پررنگ برای نمایش شکل برداری پارامترها و حروف کم رنگ برای نشان دادن شکل غیربرداری یا عددی (اسکالر) پارامترها به کار رفته است. حرف P نشان دهنده احتمال، حرف U نشان دهنده نمره سؤال در آزمون، حرف u نشان دهنده نمره آزمودنی در سؤال؛ که می تواند به صورت دو ارزشی[4] (۰و۱) یا چند ارزشی[5] (۰و۱و۲و۳و۴و...) باشد، علامت $\mid$ نشان دهنده شرطی بودن احتمال، بردار $\boldsymbol{\theta}$ نشان دهنده ویژگی‌های آزمودنی (شخص) و بردار $\boldsymbol{\eta}$ نشان دهنده ویژگی‌های سؤال‌ها می‌باشد. در نظریّه سؤال-پاسخ چند بُعدی، بردار $\boldsymbol{\eta}$، پارامتر تمیز سؤال[6] که معمولاً با a نشان داده می‌شود، پارامتر دشواری سؤال[7] که معمولاً با b نشان داده می‌شود و پارامتر شبه حدس زدن[8] یا به طور خلاصه پارامتر حدس زدن[9] که معمولاً با c نشان داده می‌شود را شامل و به صورت $\boldsymbol{\eta} = (a, b, c)$ نشان داده می‌شود. همچنین در نظریّه سؤال-پاسخ چند بُعدی، بردار $\boldsymbol{\theta}$ به صورت $\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ مشخص می‌شود که در آن عدد m نشان دهنده تعداد ابعاد مورد استفاده برای مدل‌سازی داده‌های سؤال-پاسخ، در یک فضای مختصات m بُعدی می‌باشد (ریکاس، ۲۰۰۹). انتخاب یک مدل مناسب نظریه سؤال-پاسخ (تک بُعدی یا چند بُعدی)، نیازمند تحلیل بُعدیّت[10] ساختار داده‌ها می‌باشد. در صورتی‌که بردار

1. unidimensional item respons theory
2. multidimensional item response theory
3. Reckase
4. dichotomously
5. polytomously
6. item discriminating
7. item difficulty
8. pseudo-guessing parameter
9. guessing parameter
10. dimensionality analysis

$\boldsymbol{\theta}$ تنها داری یک مؤلّفه باشد، به صورت غیربرداری θ نشان داده می‌شود و در این صورت مدل مورد استفاده یک مدل تک بُعدی خواهد بود. در غیر این صورت، به شکل برداری $\boldsymbol{\theta}$ نشان داده می‌شود و مدل مورد استفاده دو بُعدی یا چند بُعدی خواهد بود (ریکاس، ۲۰۰۹). در ادامه مروری بر مفهوم بُعدیّت و مفاهیم وابسته به آن، انجام گرفته است.

۱- مفهوم بُعدیّت و بُعدیّت فضای عاملی

ریکاس (۲۰۰۹)، بحث در مورد مفهوم بُعدیّت را با اشاره به مفهوم پیوستار[1] شروع کرده است. از نظر وی، پیوستار مفهومی شبیه به مفهوم سازه فرضی[2] دارد و در واقع عبارت است از یک مقیاس طولی که می‌توان آزمودنی‌ها را بر اساس آن مرتّب کرد. از نظر ریکاس، پیوستار اگر چه وجود دارد، ولی قابل مشاهده نیست و وجود آن را تنها می‌توان از طریق داده‌های مشاهده شده، یعنی پاسخ‌های داده شده به سؤال‌ها استنتاج کرد. فرض می‌شود تعداد پیوستارهای لازم برای تبیین تفاوت‌ها در جامعه آزمودنی‌ها متناهی است، هر چند ممکن است تعداد آن‌ها زیاد باشد. از نظر ریکاس (۲۰۰۹)، تعداد پیوستارهایی که می‌توان بر اساس پاسخ‌های داده شده به سؤال‌ها مشخص کرد، هم به تعداد ابعاد تغییر پذیری در نمونه افرادی که آزمون روی آن‌ها اجرا شده و هم به تعداد ابعاد حساسیّت پذیری سؤال‌های آزمون وابسته است. به عنوان مثال اگر آزمون به گونه‌ای طراحی شده باشد که تنها نسبت به تفاوت‌های آزمودنی‌ها در یک مهارت شناختی خاص یا نوع خاصی از دانش حساس باشد، تنها قادر خواهد بود تفاوت‌های بین افراد را در این ابعاد خاص انعکاس دهد. به طور مشابه، اگر نمونه آزمودنی‌ها تنها در یک مهارت شناختی خاص یا نوع خاصی از دانش با هم تفاوت داشته باشند، آزمون قادر خواهد بود تنها تفاوتهای ناشی از تغییر پذیری آزمودنی‌ها در این بُعد را انعکاس دهد. از نظر چنگ[3] (۲۰۱۶)، مفهوم بُعدیّت، به طور خلاصه، با تعداد خصیصه های مکنون مرتبط است و منظور از بررسی بُعدیّت این است که چند عامل یا خصیصه مکنون بر پاسخ‌های داده شده به ماده‌های یک مقیاس یا سؤال‌های یک آزمون خاص، تأثیر می‌گذارند.

1. continuum
2. hypothetical construct
3. Cheng

در ارتباط با مفهوم بُعدیّت، نکته‌ای که باید به آن توجّه کرد، تفاوت قائل شدن (یا نشدن) بین دو مفهوم بُعد[1] و عامل[2] می‌باشد. مفهوم عامل بیشتر در نظریّه کلاسیک آزمون و مفهوم بُعد بیشتر در نظریّه سؤال-پاسخ مطرح شده است. ارتباط عامل و بُعد را می‌توان در گذر از نظریّه کلاسیک به نظریّه جدید آزمون و با استفاده از مفهوم بُعدیّت فضای عاملی[3]، که هم کرلینجر[4] (۱۹۸۶، ترجمه شریفی و نجفی زند، ۱۳۸۸) در بحث خود از تحلیل عاملی کلاسیک و هم گیبونز[5] و کای (۲۰۱۸)، در بحث خود از تحلیل بُعدیّت به آن اشاره کرده‌اند، توضیح داد. کرلینجر(۱۹۸۶، ترجمه شریفی و نجفی زند، ۱۳۸۸) هنگام بحث در مورد تعداد عامل‌ها، از مفهوم بُعدیّت فضای عاملی و تعداد محورها استفاده کرده است. اگر چه او این مفهوم را توضیح نداده است و صرفاً داخل پرانتز به آن اشاره‌ای گذرا داشته است، اما بر اساس آن‌چه در ادامه بحث خود بیان کرده است، بُعدیّت فضای عاملی را بر اساس نتایج به‌دست آمده از تحلیل عاملی مرتبه دوّم توضیح داده است. به عبارت دیگر، هنگامی که عامل‌های به‌دست آمده از تحلیل عاملی مرتبه اوّل را مجدداً تحلیل عاملی کنیم، تعداد عوامل به‌دست آمده در تحلیل عاملی مرتبه دوّم (با به‌کارگیری روش چرخش مناسب)، تعداد محورها یا همان تعداد ابعاد فضای عاملی را نشان می‌دهند. در این حالت ممکن است چندین عامل مرتبه اوّل روی یک بُعد قرار گیرند و می‌توانیم تمایز مورد نظر بین مفهوم بُعد و عامل را از این طریق درک کنیم. بر این اساس، در حالتی که هر عامل روی یک محور قرار گیرد، تعداد ابعاد با تعداد عامل‌ها برابر خواهد بود، اما در حالتی که چندین عامل روی یک محور قرار گیرند، تعداد ابعاد کمتر از تعداد عامل‌ها خواهد بود. همچنین گیبونز و کای (۲۰۱۸)، در بحث خود از تحلیل بُعدیّت و انتقال از تحلیل عاملی به تحلیل بُعدیّت، به مدل‌های تحلیل سؤال-عامل چند بُعدی[6] و در ادامه به بُعدیّت فضای عاملی که آن را با D نشان می‌دهند، اشاره کرده‌اند. از نظر آن‌ها، در صورتی که بُعدیّت فضای عاملی برابر D باشد، مدل آماری مفروض برای p اُمین نمره مشاهد شده y یعنی y_p، ترکیبی خطی از D بُعد به همراه مؤلّفه خطای ϵ_j به صورت آمده در رابطه (۲) خواهد بود.

1. dimension
2. factor
3. dimensionality of the factor space
4. Kerlinger
5. Gibbons
6. multidimensional item factor analysis models

$$y_p = \alpha_{p1}\theta_1 + \alpha_{p2}\theta_2 + \cdots + \alpha_{pD}\theta_D + \epsilon_j \qquad (۲)$$

در رابطه (۲) فرض می‌شود بردار متغیرهای مکنون ($\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{\mathrm{D}})$، زیربنای تبیین تفاوت‌های فردی بین آزمودنی‌ها می‌باشد. اگر چه نه کرلینجر(۱۹۸۶، ترجمه شریفی و نجفی زند، ۱۳۸۸) و نه گیبونز و کای (۲۰۱۸)، تفاوت مفاهیم بُعد و عامل را مشخص نکرده‌اند، اما مقایسه بردار ویژگی‌های شخص یعنی ($\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{\mathrm{m}})$ در رابطه (۱) که در آن m نشان دهنده تعداد ابعاد و بردار متغیرهای مکنون ($\boldsymbol{\theta} = (\theta_1, \theta_2, \ldots, \theta_{\mathrm{D}})$ در رابطه (۲) که در آن D بُعدیّت فضای عاملی می‌باشد، نشان می‌دهد بُعدیّت فضای عاملی (D) در نظریّه کلاسیک، همان تعداد ابعاد (m) در نظریّه سؤال-پاسخ چند بُعدی می‌باشد. بر اساس این استنتاج، در این پژوهش منظور ما از بُعدیّت (تعداد ابعاد)، همان بُعدیّت فضای عاملی یا تعداد محورهای مختصات برای مدل‌سازی داده‌ها خواهد بود. در صورتی که روی هر محور مختصات (بُعد)، یک عامل قرار گیرد، تعداد ابعاد برابر با تعداد عامل‌ها خواهد بود و در صورتی که روی یک بُعد (یک محور مختصات) بیش از یک عامل قرار گیرد، تعداد ابعاد کمتر از تعداد عوامل خواهد بود. این برداشت از مفهوم بُعدیّت، با تعریف استوت[1] (۱۹۹۰) از بُعدیّت ضروری[2]، یعنی حداقل ابعاد لازم برای برقراری دو پیش‌فرض استقلال موضعی و یکنوایی سؤال‌ها نیز مطابقت دارد.

از نظر اسوتینا و لوی[3] (۲۰۱۴)، پژوهش‌ها در دهه اخیر نشان داده‌اند که به‌کارگیری یک مدل نظریه سؤال-پاسخ تک بُعدی (UIRT) برای داده‌هایی که ساختار آن چند بُعدی است، یعنی متغیرهای مکنون چندگانه لایه‌های زیرین داده‌ها را تشکیل می‌دهند، می‌تواند تفسیر و کاربرد نتایج به‌دست آمده از تحلیل داده‌های سؤال-پاسخ را فاقد اعتبار کند و از دقت شاخص‌های برآورد شده توسط مدل بکاهد. بر این اساس، ارزیابی بُعدیّت ساختار داده‌ها بر سایر جنبه‌های برازش داده-مدل[4] اولویّت دارد. به خصوص ارزیابی پیش فرض بُعدیّت[5]، زمانی که مدل‌های سؤال-پاسخ چند بُعدی برای برازش داده‌ها انتخاب می شوند، از اهمیّت بیشتری برخوردار می‌شود.

1. Stout
2. essential dimensionality.
3. Svetina & Levy
4. data-model fit
5. dimensionality assumption

۲- بُعدیّت، استقلال موضعی و یکنوایی

علاوه بر بررسی پیش‌فرض بُعدیّت، ریکاس (۲۰۰۹)، برقراری دو پیش فرض یکنوایی[1] (M) و استقلال موضعی[2] (LI) را برای اعمال یک مدل سؤال-پاسخ (تک بُعدی یا چند بُعدی) بر داده‌ها، ضروری و اساسی می‌داند. این پیش فرض که احتمال دادن پاسخ صحیح به یک سؤال از آزمون که به صورت صحیح یا غلط نمره گذاری شده است (در داده‌های دو ارزشی) و یا احتمال انتخاب یکی از مقوله‌های پاسخ (در داده‌های رتبه‌ای چند ارزشی)، با افزایش مقادیر مؤلّفه‌های بردار $\boldsymbol{\theta}$ افزایش می‌یابد، پیش فرض یکنوایی نامیده می‌شود. علاوه بر این، فرض می‌شود که پاسخ‌های داده شده به هر یک از سؤال‌های آزمون، رویدادی مستقل است؛ به این معنا که پاسخ داده شده توسط یک آزمودنی به یک سؤال، تاثیری روی پاسخ داده شده توسط آزمودنی دیگر به همین سؤال ندارد و همچنین پاسخ داده شده به یک سؤال توسط یک آزمودنی، تاثیری روی روش پاسخ یا گرایش او به انتخاب یکی از مقوله‌های پاسخ سؤالی دیگر، ایجاد نمی‌کند. معنای نمادین این فرض آن است که پاسخ هر آزمودنی به هر سؤال، منحصراً وابسته به مقادیر مؤلّفه های بردار ویژگی های آزمودنی ($\boldsymbol{\theta}$) و مقادیر پارامترهای بردار ویژگی‌های سؤال ($\boldsymbol{\eta}$) می‌باشد. معنای عملی این فرض آن است که در حین برگزاری آزمون تلاش شود تا آزمودنی‌ها اطلاعات مربوط به پاسخ‌ها را بین خود رد و بدل نکنند (از یکدیگر کمک نگیرند) و آزمون طوری طراحی شده باشد که اطلاعات موجود در یک سؤال، شانس پاسخ درست دادن به سؤالی دیگر یا انتخاب یکی از مقوله‌های پاسخ برای سؤالی دیگر را افزایش یا کاهش ندهد. مجموعاً، برقراری پیش‌فرض استقلال پاسخ‌ها در کل سؤال‌های آزمون و برای کل آزمودنی‌ها، پیش‌فرض استقلال موضعی (LI) نامیده می‌شود (ریکاس، ۲۰۰۹).

از نظر ادواردز و همکاران (۲۰۱۸)، نوعی سردرگمی شدید در ارتباط با مفاهیم بُعدیّت و استقلال موضعی وجود دارد. این سردرگمی‌ها را می‌توان در این سؤال خلاصه کرد که آیا تک بُعدی بودن پیش‌فرض لازم برای استقلال موضعی است یا بر عکس این استقلال موضعی است که تعداد ابعاد لازم برای مدل‌سازی داده ها را مشخص می‌کند. ادواردز و همکاران، پیش‌فرض تک بُعدی بودن داده‌ها را قویّاً با پیش‌فرض استقلال موضعی داده‌ها

1. monotonicity
2. local independence

مرتبط دانسته‌اند. از نظر آن‌ها، استقلال موضعی قوی[1] (SLI)، پرکاربردترین نسخه استقلال موضعی می‌باشد و حاوی این فرض است که احتمال مشترک مشاهده شده برای همه سؤال-پاسخ‌ها، برابر با حاصلضرب احتمال مشاهده هر یک از سؤال-پاسخ‌ها می‌باشد. به صورت نمادین و با نماد گذاری ریکاس (۲۰۰۹)، احتمال بردار $\boldsymbol{u}$ از سؤال-پاسخ‌ها برای فردی با بردار خصیصه مکنون $\boldsymbol{\theta}$ در آزمونی با n سؤال در رابطه (۳)[2] صدق می کند:

$$P\left(U_1 = u_1, U_2 = u_2, \dots, U_n = u_n \mid \boldsymbol{\theta}\right) = \prod_{i=1}^{n} P(U_i = u_i \mid \boldsymbol{\theta}) \qquad (3)$$

همچنین نسخه ضعیف‌تری از استقلال موضعی نیز وجود دارد (مک‌دونالد، ۱۹۶۷ و ۱۹۸۱)، که مبتنی بر کوواریانس شرطی می باشد و بر اساس آن، آزمونی با n سؤال نسبت به خصیصه مکنون θ استقلال موضعی ضعیف[3] (WLI) دارد، هرگاه میانگین کوواریانس شرطی برای همه زوج سؤال‌های این آزمون برابر با صفر باشد. به عبارت دیگر و به صورت نمادین لازم است شرط $E\left(COV(U_i, U_j) \mid \theta\right) = 0$ برقرار باشد (ریکاس، ۲۰۰۹).

از نظر برخی پژوهشگران این حوزه، لازم است که یک مدل تک بُعدی باشد تا بتواند شرط استقلال موضعی را برآورده کند. اما از نظر برخی دیگر، مفهوم استقلال موضعی مفهومی بنیادی تر از مفهوم بُعدیّت است و در واقع این تک بُعدی بودن است که می‌تواند از استقلال موضعی نتیجه شود. در این دیدگاه، در صورتی که حداقل ابعاد لازم θ برای برقرار بودن استقلال موضعی سؤال ها برابر یک باشد، آنگاه می‌توانیم بگوییم که آزمون تک بُعدی است. بر این اساس، استوت (۱۹۹۰)، بُعدیّت ضروری یک آزمون را برابر با حداقل ابعاد لازم برای برقراری دو پیش فرض استقلال موضعی و یکنوایی سؤال‌ها تعریف کرده است. همچنین از نظر ادواردز و همکاران (۲۰۱۸)، برای یک مدل چند بُعدی که ابعاد آن به صورت صحیح مشخص شده باشد، لازم است استقلال موضعی در سطح سؤال-ها برقرار باشد.

رویکرد پژوهش حاضر در ارتباط با مفاهیم بُعدیّت، استقلال موضعی و یکنوایی، مبتنی بر تعریف استوت (۱۹۹۰) است که بر اساس آن بُعدیّت ضروری یک آزمون عبارت است

1. strong local independence

۲. در رابطه (۳) نماد Π نشان دهنده عمل ضرب می‌باشد.

3. weakly locally independent

از حداقل ابعاد مورد نیاز برای بردار خصیصه مکنون $\boldsymbol{\theta}$، به نحوی که مدل در نظر گرفته شده برای داده‌ها قادر باشد هر دو پیش‌فرض استقلال موضعی و یکنوایی را برقرار سازد.

۳- بُعدیّت و وابستگی موضعی

وابستگی موضعی[1] (LD)، اصطلاحاً به نقطه مقابل استقلال موضعی (LI)، یعنی حالتی که بین سؤال‌ها استقلال موضعی برقرار نباشد، اشاره دارد. ین[2] (۱۹۹۳)، سه علت را برای رخ دادن وابستگی موضعی بین سؤال‌های یک آزمون بیان کرده است. از نظر ین (۱۹۹۳)، یک علت می‌تواند وجود عبارت‌های شبیه به هم در متن این‌گونه از سؤال‌ها باشد. این نوع از وابستگی، وابستگی موضعی ناشی از عبارت شبیه[3] نامیده شده است. مثلاً در یک آزمون درک مطلب، سؤال‌هایی که واژه‌های یکسانی را در متن خود دارند، ممکن است وابستگی موضعی از خود نشان دهند. از نظر چن و تیسن[4] (۱۹۹۷)، در این حالت وابستگی موضعی ناشی از متنی است که منبع سؤال‌ها به شمار می‌رود و این نوع از وابستگی نشان دهنده این واقعیت است که سؤال‌ها در معنای روان‌شناختی خود تک بُعدی نیستند. علت دیگری که ین (۱۹۹۳) برای مشاهده وابستگی موضعی بین سؤال‌ها بیان کرده است، مکان قرار گرفتن سؤال‌ها در آزمون می‌باشد. این نوع از وابستگی، وابستگی موضعی ناشی از مکان نامیده شده است. از نظر چن و تیسن (۱۹۹۷)، وابستگی موضعی ناشی از مکان زمانی اتفاق می‌افتد که سرعت یک عامل تاثیر گذار در پاسخ‌گویی به سؤال‌ها باشد. بر این اساس، سؤال‌هایی که در انتهای آزمون قرار می‌گیرند، ممکن است وقت کافی برای پاسخ‌گویی به آن‌ها وجود نداشته باشد و توسط آزمودنی‌ها نادیده گرفته شوند. به همین دلیل ممکن است بین آن‌ها وابستگی موضعی مشاهده شود. در این حالت، اگر تعدادی از آزمودنی‌ها فرصت کافی برای پرداختن به سؤال i را نداشته باشند، برای سؤال i+۱ و به همین ترتیب سؤال‌های بعدی نیز فرصت پاسخ‌گویی نخواهند داشت و بین این سؤال‌ها ممکن است وابستگی موضعی مشاهده شود. علت سوّم از نظر ین (۱۹۹۳)، وجود سؤال‌های یکسان (مشابه)[5] در آزمون می‌باشد. در این حالت احتمال پاسخ دادن آزمودنی‌ها به دو سؤال یکسان (مشابه)، برابر با احتمال پاسخ دادن آن‌ها به یکی از این دو سؤال خواهد بود. این نوع از وابستگی،

1. local dependence
2. Yen
3. same passage
4. Chen & Thissen
5. identical items

وابستگی موضعی کامل[1] (PLD) نامیده شده است و از نظر چن و تیسن (۱۹۹۷) به‌ندرت مشاهده می‌شود. در رابطه (۴) این نوع از وابستگی به صورت نمادین نشان داده شده است.

$$P(U_i = u_i, U_j = u_j \mid \boldsymbol{\theta}) = P(U_i = u_i \mid \boldsymbol{\theta}) \qquad (۴)$$

از نظر چن و تیسن (۱۹۹۷)، سؤال‌هایی که با یکدیگر وابستگی موضعی دارند، سؤال‌هایی زائد به حساب می‌آیند. زیرا این نوع از سؤال‌ها اطلاعات اندکی را برای مدل سؤال-پاسخ در نظر گرفته شده، فراهم می‌کنند. به عبارت دیگر، آزمون‌های حاوی این گونه از سؤال‌ها، اطلاعات کمتری نسبت به آن‌چه که در ظاهر نشان می‌دهند، ارائه می‌کنند. به همین دلیل وجود وابستگی موضعی بین سؤال‌ها، برآورد پارامترها در مدل‌های سؤال-پاسخ را تحت تاثیر قرار می‌دهد. همچنین، برآورد پارامترهای مدل‌های سؤال-پاسخ، در حالتی که سؤال‌ها با یکدیگر وابستگی موضعی دارند، نسبت به حالتی که دارای استقلال موضعی هستند، ممکن است دارای اختلاف زیادی باشد و در چنین حالتی مقیاس‌بندی افراد توسط مدل، می‌تواند کاملاً غیر دقیق و نامعتبر باشد.

۳-۱. شاخص‌های نمایش دهنده وابستگی موضعی

با توجّه به اهمیت تشخیص سؤال‌های دارای وابستگی موضعی در تعیین تعداد دقیق ابعاد لازم برای یک مدل سؤال-پاسخ، لازم است شاخص‌هایی برای مشخص کردن سؤال‌هایی که با یکدیگر وابستگی موضعی دارند، به همراه تفسیرهای مرتبط با آن‌ها ارائه شود. هاتس و ادواردز (۲۰۱۳)، بر اساس چهار قاعده، شامل: (الف) شاخص در عمل به طور گسترده مورد استفاده قرار گرفته باشد؛ (ب) به طور مکرر در ادبیّات پژوهشی به آن استناد شده باشد؛ (ج) روش محاسبه آن (توسط نرم افزارها یا برنامه‌ها) در دسترس باشد و (د) مقادیر آن برای هر زوج از سؤال‌های آزمون در مقابل کل آزمون وجود داشته باشد، چندین شاخص را معرفی کرده‌اند. این شاخص‌ها عبارتند از شاخص Q_3 ین (۱۹۸۴ و ۱۹۹۳)، تبدیل r به z فیشر برای شاخص Q_3[2]، آماره نسبت درستنمایی[3] G^2 یا معادل آن برای نمونه‌های بزرگ یعنی شاخص X^2 پیرسون (ارائه شده توسط چن و تیسن، ۱۹۹۷) و

1. perfectly locally dependent
2. Fisher's r-to-z transformed Q3
3. likelihood ratio statistic

شاخص شیب جک نایف[1] (*JSI*) ادواردز و کای (۲۰۰۸)، که همگی تحت عنوان شاخص‌های مبتنی بر نظریّه سؤال-پاسخ ارائه شده‌اند. با توجّه به این‌که شاخص‌های X^2 پیرسون و Q_3 ین به لحاظ زمان ارائه شدن نسبت به سایر شاخص‌ها تقدم دارند و پژوهش-های بیشتری به کمک آن‌ها انجام شده است، بحث‌های بعدی با توجّه به این دو شاخص خواهد بود. به همین دلیل در ادامه این دو شاخص به طور مختصر معرفی شده‌اند.

۳-۱-۱. شاخص X^2 پیرسون

از نظر چن و تیسن (۱۹۹۷)، روش بیشینه درستنمایی حاشیه ای[2] (MML)، روشی است که معمولاً بیشترین کاربرد را در محاسبه پارامترهای مدل‌های سؤال-پاسخ دارد. این روش داده‌های سؤال-پاسخ را با k^m جدول که در آن k تعداد مقوله‌های پاسخ و m نشان دهنده تعداد سؤال‌ها می‌باشد، برازش می‌دهد. یکی از اطلاعات جانبی به‌دست آمده از روش برآورد بیشینه درستنمایی حاشیه‌ای، اطلاعات مربوط به فراوانی‌های مورد انتظار پیش‌بینی شده برای هر الگوی پاسخ، توسط مدل سؤال-پاسخ اعمال شده می‌باشد. در این حالت بر اساس فراوانی‌های مورد انتظار[3] (پیش‌بینی شده توسط مدل سؤال-پاسخ) برای هر زوج از سؤال‌ها که با E_{pq} نشان داده می‌شوند و فراوانی‌های مشاهده شده[4] (به‌دست آمده از داده‌های خام) برای هر زوج از سؤال‌ها که با O_{pq} نشان داده می‌شوند، شاخص X^2 پیرسون برای هر زوج سؤال p و q با i و j مقوله پاسخ بر اساس رابطه (۷) محاسبه می شود.

$$X^2 = \sum_i \sum_j \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \qquad (۷)$$

از نظر چن و تیسن (۱۹۹۷)، آماره X^2 از توزیع χ^2 (خی دو) با درجه آزادی برابر با تعداد خانه‌های جدول (مثلا برای دو سؤال با دو مقوله پاسخ این عدد برابر با ۴ است) منهای تعداد پارامترهای لُگاخطی[5] مدل، برخوردار می باشد. از نظر کای، دوتویت[6] و تیسن (۲۰۱۱)، در صورتی که قدر مطلق مقدار استاندارد شده (تقریبی) شاخص X^2 ($|X^2|$)

1. Jackknife Slope Index
2. maximum marginal likelihood
3. expected frequency
4. observed frequency
5. loglinear parameters
6. Du Toit

بزرگتر از ۱۰ باشد، نشان دهنده احتمال وجود وابستگی موضعی بین دو سؤال می‌باشد که توسط مدل تک بُعدی سؤال-پاسخ در نظر گرفته نشده است؛ مقادیر بین ۵ تا ۱۰ به عنوان مقادیر متوسط و مشکوک در نظر گرفته می‌شوند و مقادیر کوچکتر از ۵ ناچیز و کم اهمیت پنداشته می شوند.

۳-۱-۲. شاخص Q_3 ین

شاخص Q_3 ین (۱۹۸۴)، پاسخ های مورد انتظار و مشاهده شده برای هر آزمودنی را مقایسه و انحراف بین این مقادیر را از طریق رابطه $d_{ja} = x_{ja} - E_{ja}$ محاسبه می‌کند. در این رابطه، x_{ja} برابر با نمره مشاهده شده a اُمین آزمودنی به سؤال j و E_{ja} نمره واقعی[1] مورد انتظار آزمودنی در سؤال j می باشد و معنای آن این است که E_{ja} ممکن است دارای مقداری غیرصحیح[2] باشد. بر این اساس شاخص وابستگی موضعی Q_{3ij} از حاصلضرب همبستگی گشتاوری پیرسون بین انحراف نمره‌های سؤال های i وj محاسبه می‌شود، یعنی $Q_{3ij} = r_{d_i d_j}$ (هاتس و ادواردز، ۲۰۱۳). شاخص Q_3 در واقع بیان کننده نوعی همبستگی بین باقی‌مانده‌ها در یک مدل سؤال-پاسخ می‌باشد. از نظر چن و تیسن (۱۹۹۷) و تیت[3] (۲۰۰۳)، نقطه برش مرسوم برای مقایسه مقدار شاخص Q_3 برابر با ۰/۲ می‌باشد و مقادیر بزرگتر از این مقدار، نشان دهنده وابستگی موضعی بین سؤال‌ها می‌باشد. هر چند چن و تیسن (۱۹۹۷) بیان کرده‌اند که توان این مقدار (۰/۲) پایین است و بهتر است از روش‌های شبیه‌سازی با حجم نمونه‌های متفاوت و آزمون‌های دارای طول متفاوت به همراه بررسی‌های تجربی، برای در نظر گرفتن نقطه برش مناسب، استفاده شود. ریس[4] (۱۹۹۵) نشان داده است شاخص Q_3 میزان وابستگی موضعی را برای سؤال‌هایی که اندازه وابستگی موضعی آن‌ها متوسط تا زیاد شبیه‌سازی شده بودند را کم‌تر از مقدار در نظر گرفته شده برآورد می‌کند. بنابراین برای مقادیری که نزدیک به این مقدار (۰/۲) می‌باشند، باید با احتیاط عمل کرد.

1. true score
2. noninteger
3. Tate
4. Reese

٤- بُعدیّت و مدل‌های وابستگی موضعی

چن و تیسن (۱۹۹۷)، دو مدل برای توضیح وابستگی‌های موضعی مشاهده شده بین سؤال‌ها ارائه کرده‌اند. این دو مدل، مدل وابستگی موضعی اساسی[1] (ULD) و مدل وابستگی موضعی سطحی[2] (SLD)، نامیده شده‌اند.

در مدل وابستگی موضعی اساسی فرض می‌شود که یک خصیصه مکنون مجزّا وجود دارد که بین هر یک از سؤال‌هایی که با یکدیگر وابستگی موضعی دارند، مشترک است، ولی با سایر سؤال‌های موجود در آزمون اشتراکی ندارد. در این مدل، بردار خصیصه مکنون ($\boldsymbol{\theta}$) مرتبط با سؤال i، به صورت ترکیبی خطی از وزن‌ها (wt_{i} و $wt_{\mathrm{ki\,سؤال}}$) و خصیصه‌های مکنون($\theta_{۱}^{*}$ و θ_{k}^{*}) تعیین می شود که در آن k نشان دهنده تعداد خصیصه‌های مکنون (تعداد ابعاد) می‌باشد. شکل ماتریسی ارائه شده توسط چن و تیسن (۱۹۹۷) برای چهار سؤال فرضی، که در آن فرض شده است سؤال‌های ۱ و ۲ و همچنین سؤال‌های ۳ و ۴ دارای وابستگی موضعی با یکدیگر می‌باشند را می‌توان به صورت رابطه خطی (۵) نشان داد.

$$\begin{aligned}
\boldsymbol{\theta}_{\text{سؤال ۱}} &= wt_{۱}\theta_{۱}^{*} + wt_{۲۱\,\text{سؤال}}\theta_{۲}^{*} \\
\boldsymbol{\theta}_{\text{سؤال ۲}} &= wt_{۲}\theta_{۱}^{*} + wt_{۲۲\,\text{سؤال}}\theta_{۲}^{*} \\
\boldsymbol{\theta}_{\text{سؤال ۳}} &= wt_{۳}\theta_{۱}^{*} + wt_{۳۳\,\text{سؤال}}\theta_{۳}^{*} \\
\boldsymbol{\theta}_{\text{سؤال ۴}} &= wt_{۴}\theta_{۱}^{*} + wt_{۳۴\,\text{سؤال}}\theta_{۳}^{*}
\end{aligned} \qquad (۵)$$

در رابطه (۵)، wt_{i} ها وزن هر یک از سؤال‌ها را روی خصیصه مکنون یا بُعد یک ($\theta_{۱}^{*}$) نشان می‌دهند و $wt_{\mathrm{ki\,سؤال}}$ وزن سؤال i را روی k اُمین خصیصه مکنون (θ_{k}^{*}) نشان می‌دهد. در صورتی که همه $wt_{\mathrm{ki\,سؤال}}$ ها برابر صفر باشند، همه wt_{i} ها برابر یک خواهند بود و $\boldsymbol{\theta}$ هر سؤال برابر $\theta_{۱}^{*}$ خواهد شد. برقراری این شرایط استقلال موضعی سؤال‌ها روی یک بُعد را نشان می‌دهد و در این حالت می گوییم آزمون تک بُعدی است. در حالتی که همه wt_{i} ها برابر صفر باشند، آنگاه باید همه $wt_{\mathrm{ki\,سؤال}}$ ها برابر یک باشند و در این حالت آزمون یک آزمون چند بُعدی با ابعاد غیرهمبسته خواهد بود. در این حالت مثلاً سؤال‌های ۱ و ۲ روی بُعد $\theta_{۲}^{*}$ و سؤال‌های ۳ و۴ روی بُعد $\theta_{۳}^{*}$ قرار می گیرند و آزمون با این چهار سؤال فرضی

1. underlying local dependence
2. surface local dependence

یک آزمون دو بُعدی با ابعاد غیرهمبسته خواهد بود. از نظر چن و تیسن (۱۹۹۷)، آزمون‌هایی با این نوع از ساختار باید به صورت آزمون‌های چندگانه در نظر گرفته شوند و لازم است تحلیل‌های سؤال-پاسخ برای سؤال‌های هر بُعد به صورت مجزّا انجام شود. بر اساس همین مدل مذکور، i اُمین و j اُمین سؤال‌های یک آزمون زمانی وابسته موضعی هستند که wt_{i} ، wt_{j} ، wt_{ki} سؤال و wt_{kj} سؤال همگی غیر صفر باشند. اگر این روند در مورد سایر سؤال‌های آزمون نیز برقرار باشد، در این صورت گوییم آزمون یک آزمون چند بُعدی با ابعاد همبسته می‌باشد. از نظر ین (۱۹۸۴)، هنگامی که پارامترهای یک مدل چند بُعدی با ابعاد همبسته توسط یک مدل تک بُعدی برآورد شوند، در این صورت ترکیبی از ابعاد زیربنایی به عنوان تک متغیر مکنون (تک بُعد) در نظر گرفته می‌شود و در این حالت وابستگی‌های مثبت و منفی بین زوج سؤال‌ها اتفاق می‌افتد. وابستگی مثبت بین زوج سؤال‌های آزمون‌هایی اتفاق می‌افتد که یک خصیصه مکنون دوّم مشابه با خصیصه مکنون اوّل را اندازه‌گیری می‌کنند و وابستگی موضعی منفی بین زوج سؤال‌های آزمون‌هایی اتفاق می‌افتد که یک خصیصه مکنون غیر مشابه با خصیصه مکنون اوّل را اندازه می‌گیرند.

دوّمین مدل وابستگی موضعی یعنی مدل وابستگی سطحی (SLD)، بیشتر قابل اطلاق به آزمون‌هایی است که تعداد سؤال‌های آن‌ها زیاد است و یا در آن‌ها سؤال‌های مشابه وجود دارد. از نظر ادواردز و همکاران (۲۰۱۸)، مدل وابستگی سطحی زمانی مشاهده می‌شود که سؤال‌ها از نظر محتوا یا مکان قرار گرفتن در آزمون، آن‌چنان به یکدیگر شبیه هستند که آزمودنی‌ها به آن‌ها به صورت یکسان پاسخ می‌دهند. در مدل وابستگی سطحی، احتمال این که پاسخ آزمودنی به سؤال i برابر با پاسخ او به سؤال j باشد، با نماد π_{LD} نشان داده می‌شود. برای دو سؤال دو ارزشی، مدل وابستگی سطحی با رویکرد بیان شده توسط چن و تیسن (۱۹۹۷)، به صورت ذکر شده در رابطه (۶) تعریف می‌شود. با احتمال π_{LD}:

$$\text{پاسخ به سؤال ۲} = \begin{cases} ۱ & \text{اگر } X_{۱} = ۱ \\ ۰ & \text{اگر } X_{۱} = ۰ \end{cases} \qquad (۶)$$

در این مدل، π_{LD} نشان دهنده احتمال یکسان بودن پاسخ آزمودنی به دوّمین سؤال با پاسخ او به اوّلین سؤال، بدون در نظر گرفته شدن مقادیر پارامترهای بردار $\boldsymbol{\theta}$ آزمودنی می‌باشد. اگر مقدار π_{LD} برابر صفر باشد، نشان می‌دهد که پاسخ آزمودنی به سؤال دوّم، ارتباطی با پاسخ او به سؤال اوّل ندارد. در این حالت بین دو سؤال وابستگی موضعی وجود

ندارد و مدل سؤال-پاسخ، یک مدل عادی و درست می‌باشد. اگر مقدار π_{LD} برابر یک باشد، نشان می‌دهد که پاسخ آزمودنی به سؤال دوّم، وابسته به پاسخ او به سؤال اوّل است. در این حالت دو سؤال وابسته موضعی کامل (PLD) می‌باشند و مدل سؤال-پاسخ، در صورت اعمال، از شرایط عادی برخوردار نمی‌باشد.

از نظر ادواردز و همکاران (۲۰۱۸)، هنگامی که بین دو سؤال i و j یک آزمون وابستگی سطحی (SLD) وجود داشته باشد، فرآیندهای زیربنایی تاثیرگذار بر پاسخ‌های آزمودنی‌ها (مقادیر پارامترهای بردار $\boldsymbol{\theta}$)، نخواهند توانست فرآیندهای موثر بر پاسخ آزمودنی‌ها به این دو سؤال را تعیین کنند. به عبارت دیگر در چنین حالتی، فرآیندهای در نظر گرفته شده توسط مدل سؤال-پاسخ برای پاسخ آزمودنی‌ها به سؤال j ، بر اساس پاسخ‌های آن‌ها به سؤال i تعیین می‌شوند، بدون این‌که سطح توانایی آن‌ها در خصیصه مکنون مورد نظر برای پاسخ به سؤال j توسط مدل در نظر گرفته شود.

با توجّه به آن‌چه در مورد دو مدل وابستگی اساسی (ULD) و وابستگی سطحی (SLD) بیان شد، هنگامی که بخواهیم این دو مدل را در ارتباط با تعداد متغیرهای مکنون، یا بُعدیّت ساختار یک آزمون با یکدیگر مقایسه کنیم، می‌توانیم بیان کنیم که در واقع وابستگی اساسی حاکی از وجود متغیرهای مکنونی است که در مدل سؤال-پاسخ اعمال شده بر داده‌ها لحاظ نشده‌اند، در حالی که وابستگی سطحی حتی نسبت به متغیرهای مکنون موجود و لحاظ شده در مدل نیز بی‌تفاوت است. از نظر ادواردز و همکاران (۲۰۱۸)، مدل وابستگی اساسی نشان می‌دهد که در درون هر مجموعه از سؤال‌های دارای وابستگی موضعی، یک خصیصه مکنون زیربنایی وجود دارد که سؤال‌های دارای وابستگی موضعی با آن اشتراک دارند، ولی سایر سؤال‌های آزمون با آن اشتراکی ندارند. بر خلاف مدل وابستگی سطحی که درآن خصیصه‌های مکنون نقشی در نحوه پاسخ‌گویی آزمودنی‌ها به سؤال‌ها ندارند، در مدل وابستگی اساسی، خصیصه‌های مکنون چندگانه‌ای پاسخ آزمودنی‌ها به سؤال‌ها را تعیین و تحت تاثیر قرار می‌دهند، ولی برخی از این خصیصه‌ها توسط مدل نادیده گرفته می‌شوند.

۵- بُعدیّت ساختار توانایی‌های شناختی

از نظر دوکاس[1] (۲۰۰۴، به نقل از نجاتی، ۱۳۹۲) و شتلوُرث[2] (۲۰۱۰، به نقل از نجاتی، ۱۳۹۲)، توانایی‌های شناختی[3] عبارتند از فرآیندهای عصبی درگیر در اکتساب، پردازش، نگهداری و کاربست اطلاعات که به منظور حل مشکلات بوم شناختی و هدایت محیط-های اجتماعی پیچیده، تحول یافته‌اند.

ایده تعیین بُعدیّت، به‌طور گسترده در روانشناسی (کارول[4]، ۱۹۹۳)، اندازه‌گیری (ریکاس، ۲۰۰۹) و علوم عصب‌شناختی[5] (اتول، جیانگ، آبدی، پینارد، دانلوپ و پارنت[6]، ۲۰۰۷) مورد استفاده قرار گرفته است. در گذشته، معمول‌ترین راه مطالعه بُعدیّت ساختار توانایی‌های شناختی[7]، استفاده از روش تحلیل عاملی[8] بوده است. در طول این سال‌ها، پژوهشگران از طریق تحلیل روابط متقابل کارکردهای شناختی متفاوت در جمعیّت‌های مختلف، به این نتیجه مشترک رسیده‌اند که ساختار توانایی‌های شناختی به صورت سلسله-مراتبی[9] می‌باشد و این ساختار رابطه بین توانایی‌های عمومی و خاص[10] را نشان می‌دهد. در این ساختار سلسله‌مراتبی، توانایی‌های عمومی که آن را عامل g می‌نامند، در بالاترین سطح قرار می‌گیرد و سایر توانایی‌های شناختی در سطوح دوّم و به همین ترتیب توانایی‌های جزئی‌تر در سطوح پایین‌تر قرار می‌گیرند. با این وجود و به دلیل این‌که بُعدیّت چنین ساختاری توسط تحلیل عاملی تجزیه و تحلیل و شناسایی شده است، بسیاری بر این عقیده هستند که ابعاد مشخص شده برای این ساختار می‌تواند مصنوعی و غیرواقعی و حاصل یک مدل عاملی آماری و نه حاصل ارتباط مستقیم با فرآیندهای شناختی باشد (چنگ، ۲۰۱۶).

علاوه بر این، به استفاده از روش‌های آماری مورد استفاده برای تعیین بُعدیّت (مثلاً تحلیل عاملی و برخی روش‌های کلاسیک رایج) نیز انتقاداتی وارد شده است. مثلاً،

1. Dukas
2. Shettleworth
3. cognitive abilities
4. Carroll
5. neuroscience
6. Otoole, Jiang, Abdi, Penard, Dunlop & Parent
7. structure of cognitive abilities
8. factor analysis
9. hierarchical
10. general and special abilities

امبرتسون[1] و رایس[2] (۲۰۰۰، ترجمه شریفی، فرزاد، حبیبی عسگرآباد و ایزانلو، ۱۳۸۸) بیان کرده‌اند که علی‌رغم رویکردهای اکتشافی و روش‌های آماری متعددی که برای بررسی بُعدیّت یک مقیاس یا آزمون پیشنهاد شده است، پژوهش‌ها در این حوزه با بد فهمی‌هایی مواجه بوده است و تعدادی از روش‌هایی که ظاهراً به بررسی بُعدیّت آزمون می‌پردازند، در اصل نسبت به این موضوع بی ارتباط هستند. هتی[3] (۱۹۸۴، ۱۹۸۵، به نقل از امبرتسون و رایس،۲۰۰۰، ترجمه شریفی و همکاران، ۱۳۸۸) به برخی از این روش‌ها مانند «همسانی الگوهای پاسخ‌دهی آزمودنی[4]»، «اعتبار نمره‌های آزمون[5]»، نتایج آماری حاصل از «تحلیل مؤلّفه‌های اصلی[6]» (مانند تعداد ارزش‌های ویژه[7] بزرگتر از ۱ و یا نسبت ارزش ویژه عامل اوّل به ارزش ویژه عامل دوّم)، «نتایج تحلیل عاملی خطی و غیرخطی[8]»، «اطلاعات آماری حاصل از برازش مدل‌های خصیصه مکنون» (مانند تحلیل باقی مانده‌ها[9] پس از برازش یک مدل خاص) اشاره کرده است. مطالعه هتی (۱۹۸۴ و ۱۹۸۵) نشان داده است که بیشتر این رویکردها نسبت به سنجش بُعدیّت آزمون، یا بی‌ارتباط هستند و یا نارسایی شدید دارند. از جمله مشکلات عمده این روش‌ها می‌توان به عدم توانایی در تمیز معتبر بین یک مجموعه از داده‌های تک بُعدی و یک مجموعه از داده‌های دو بُعدی یا بیشتر؛ فاقد توزیع نمونه-گیری مشخص بودن و همچنین معدود بودن روش‌های خاص قضاوت در مورد تعداد ابعاد آزمون برای پژوهشگران، اشاره کرد (به نقل از امبرتسون و رایس،۲۰۰۰، ترجمه شریفی و همکاران، ۱۳۸۸).

یک روش متفاوت تعیین بُعدیّت، مسبوق به یافته‌های مطالعات زوج-تکلیف[10] روان‌شناختی یا تجزیه و تحلیل خطای شناختی[11] می‌باشد. در مطالعات زوج-تکلیفی، پژوهشگران تلاش می‌کنند تا کشف کنند کدام زوج از تکالیف می‌توانند توسط یک

1. Embretson
2. Reise
3. Hattie
4. the consistency of examinee answer patterns
5. test score reliability
6. principal components analyses
7. eigenvalues
8. linear and nonlinear factor analyses
9. residual analyses
10. dual-task
11. cognitive error analysis

شخص در یک زمان انجام شوند. پژوهشگران در این گونه مطالعات، به دنبال مشخص کردن این موضوع هستند که کدام زوج-تکلیف‌های در نظر گرفته شده، فرآیندهای کارکردی یکسانی را در مغز انسان به اشتراک می‌گذارند. تکالیفی که چنین خاصیتی داشته باشند، روی بُعد یکسانی از توانایی‌های شناختی قرار می‌گیرند (پاشلر[1]، ۱۹۹۴).

روش دیگر و تا حدودی متفاوت برای مطالعه بُعدیّت ساختار توانایی‌های شناختی، استفاده از روش‌های عصب-روان‌شناختی[2] (نوروپسیکولوژی) برای مطالعه اتصال‌های مغزی در طی فرآیندهای شناختی به کمک تصویر برداری مغزی[3] می‌باشد. از نظر ناگلیری و اترو[4] (۲۰۱۱)، این روش نیز یک ساختار سلسله‌مراتبی را برای ساختار توانایی‌های شناختی نشان داده است.

برخی رویکردها به تعیین بُعدیّت نیز به صورت ترکیبی می‌باشند. به عنوان مثال، پیشرفت‌های اخیر در زمینه پژوهش‌های مبتنی بر تصویر برداری مغزی، تلاش برای تعیین این‌که کدام دو تکلیف، نواحی یکسانی از مغز را به فعالیت وادار می‌کنند، گسترش داده و بنابراین هر دو رویکرد استفاده از زوج-تکلیف‌ها و تصویر برداری مغزی، روش یکسانی را به کار می‌گیرند (هابارد، پیازا، پینل و دیهین[5]، ۲۰۰۵).

بین روش‌های روان‌سنجی و تصویر برداری مغزی، تفاوت‌های آشکاری وجود دارد. این موضوع می‌تواند به درک بیشتر بُعدیّت ساختار توانایی‌های شناختی کمک کند. به عنوان مثال، هاریس، هیرش-پاسک و نیوکامب[6] (۲۰۱۳)، مطالعات مربوط به تکالیف چرخش ذهنی[7] و تاکردن ذهنی[8] را در هر دو روش روان‌سنجی و تصویر برداری مغزی مورد مقایسه و مقابله قرار دادند. آن‌ها دریافتند که اگر چه روش تحلیل عاملی مبتنی بر روان‌سنجی این دو تکلیف را به صورت مجزّا در نظر می‌گیرد، اما بررسی فعالیت‌های ذهنی به کمک تصویر برداری مغزی نشان دهنده این موضوع بوده است که این دو تکلیف در نواحی یکسانی از مغز انجام می‌گیرند.

1. Pashler
2. neuropsychological
3. brain imaging
4. Naglieri & Otero
5. Hubbard, Piazza, Pinel & Dehaene
6. Harris, Hirsh-Pasek & Newcombe
7. mental rotation
8. mental folding

از نظر چنگ (۲۰۱۶)، هر چند رویکردهای مختلف شرح داده شده، در سلسله مراتبی بودن ساختار توانایی‌های شناختی توافق دارند، اما پژوهشگران اغلب در تعریف تقسیمات این توانایی‌های شناختی با یکدیگر هم‌رأی نمی‌باشند. بیشترین توافق در این تقسیم بندی از قرار گرفتن توانایی‌ها به صورت یک ساختار واحد یا به صورت سطوحی در درون سلسله مراتب وجود دارد و هیچ‌کدام از وضوح کافی برخوردار نمی‌باشند. به همین دلیل پژوهشگران همزمان با این که تلاش می‌کنند فرآیندهای فکری شرکت کنندگان را از طریق تحلیل الگوهای خطای آن‌ها کشف کنند، این الگوها را با مدل‌های آماری، از قبیل مدل‌های مبتنی بر نظریّه سؤال-پاسخ، نیز ارتباط می‌دهند. در این مدل‌ها با توجّه به این‌که تعامل بین پارامترهای شخص و پارامترهای سؤال‌ها به طور همزمان در نظر گرفته می‌شود، اطلاعات بیشتر و کامل‌تری مبتنی بر تعامل همزمان توانایی شخص و ویژگی سؤال‌ها، تولید می‌شود (تاتسوکا[1]، ۱۹۹۰). این موضوع که مدل‌های سؤال-پاسخ اطلاعات بیشتر و کامل-تری را نسبت به مدل‌های کلاسیک در نظر می‌گیرند، نقطه قوّت روش تحلیل عاملی مبتنی بر نظریّه سؤال-پاسخ نسبت به تحلیل عاملی کلاسیک می‌باشد.

از نظر گیبونز و کای (۲۰۱۸)، روش‌های کلاسیک تحلیل عاملی را به دلیل این‌که ماتریس همبستگی گشتاوری ضربی داده‌ها به طور مستقیم قابل محاسبه نمی‌باشد، نمی‌توان به صورت مستقیم برای داده‌های سؤال-پاسخ مورد استفاده قرار داد. از نظر آن‌ها، آن‌چه در عمل و در تحلیل عاملی کلاسیک بر اساس این ماتریس انجام می‌گیرد، صرفاً می‌تواند شامل تقریبی از ضرایب همبستگی تتراکوریک برای داده‌های دو ارزشی، و در صورتی که یکی از سؤال‌ها دارای بیش از دو ارزش باشد، مبتنی بر ضرایب همبستگی پلی‌کوریک برای تمام زوج سؤال‌ها، و بر اساس فراوانی‌های مشترک مشاهده شده برای مقوله‌های پاسخ سؤال‌ها باشد. در صورتی که یکی از این فراوانی‌ها برابر صفر باشد، محاسبات مربوط به این ضریب همبستگی‌ها مختل می‌شود و به همین دلیل لازم است مقداری اختیاری (مثلاً ۰/۵) جایگزین شود تا این ماتریس کامل شود. اما در تحلیل عاملی مبتنی بر نظریّه سؤال-پاسخ، تمام اطلاعات مربوط به مقوله‌های پاسخ مورد استفاده قرار می‌گیرند و صرفاً وابسته به شاخص‌های زوجی از قبیل ضرایب همبستگی تتراکوریک و پلی‌کوریک نمی باشد. به همین دلیل تحلیل عاملی مبتنی بر نظریّه سؤال-پاسخ به عنوان تحلیل عاملی با اطلاعات

1. Tatsuoka

کامل[1] شناخته می‌شود. برخلاف تحلیل عاملی کلاسیک که در آن ماتریس همبستگی سؤال‌ها، به دلیل این که فرآیندهای پاسخ قابل مشاهده نیستند، به طور غیر مستقیم مشخص می‌شود، تحلیل عاملی مبتنی بر نظریّه سؤال-پاسخ، به طور مستقیم با استفاده از احتمال‌های مشاهده شده برای هر مقوله پاسخ به عنوان تابعی از پارامترهای مربوط به خصیصه‌های مکنون آزمودنی‌ها و پارامترهای سؤال‌ها، انجام می‌گیرد. از نظر گیبونز و کای (۲۰۱۸)، تحلیل عاملی کلاسیک می‌تواند به عنوان یک تحلیل مقدماتی و راهی سریع برای یافتن کران بالایی از تعداد عوامل پذیرفتنی مورد استفاده قرار گیرد و سپس روش‌های دقیق‌تر مبتنی بر نظریّه سؤال-پاسخ مورد استفاده قرار گیرد. همچنین از نظر گیبونز و کای، روش‌های سنتی تعیین بُعدیّت مبتنی بر مقادیر ویژه، بیشتر گرایش به تعیین عامل‌هایی دارند که نشان دهنده خصیصه‌های مکنون زیربنایی نیستند، اما شاخص‌های نیکویی برازش در نظریّه سؤال-پاسخ، که مدل‌های آماری آشیانه‌ای[2] و غیر آشیانه‌ای را با هم مقایسه می‌کنند، می‌توانند به طور کاملاً مؤثّری بُعدیّت یک آزمون خاص را ارزیابی و تعیین کنند.

٦- مروری بر پیشینه پژوهش مرتبط با بُعدیّت در ایران

نگاهی به پژوهش‌های انجام شده در ایران و جستجوی کلید واژه‌های وابستگی موضعی در موتورهای جستجو نشان داد پژوهشی که در آن شاخص‌های وابستگی موضعی معرّفی یا به منظور بررسی بُعدیّت ساختار خاصی از داده‌ها یا توانایی‌ها (شناختی یا غیرشناختی) به-کار گرفته شده باشد، وجود ندارد. اگر چه فرّخی، حرفه‌دوست، محمدجانی، جبّاری و رعایی (۱۳۹۳)، شاخص چن و تیسن (۱۹۹۷) را برای بررسی استقلال موضعی بین سؤال‌ها مورد استفاده قرار داده‌اند، اما پیش‌فرض تک بعدی بودن را با استفاده از روش‌های تحلیل مؤلّفه‌های اصلی، تحلیل عاملی تاییدی و نسبت مقدار ویژه مؤلّفه اوّل به مقدار ویژه مؤلّفه دوّم بررسی کرده‌اند (این نسبت برابر با ۴/۴۸ بوده است).

همان‌گونه که در بخش‌های قبل اشاره شد، پیش‌فرض‌های استقلال موضعی و یکنوایی، دو شرط اساسی در به‌کارگیری مدل‌های سؤال-پاسخ می‌باشند، با این وجود، در پژوهش‌هایی که در کشور با استفاده از نظریّه سؤال-پاسخ انجام گرفته است، کمتر به بررسی این دو پیش‌فرض اساسی پرداخته شده است. در میان پژوهش‌های انجام شده در

1. full information factor analysis
2. nested

ارتباط با بررسی و تعیین بُعدیّت ساختار یک آزمون، ایزانلو، بازرگان، فرزاد، صادقی و کاووسی (۱۳۹۳)، به منظور مشخص کردن ابعاد متعامد خوشه‌های آزمون ۵۵ سؤالی ریاضی (دو ارزشی صحیح-غلط)، پژوهشی را با به‌کارگیری روش‌های متفاوت تحلیل بُعدیّت از قبیل روش‌های ناپارامتریک تعبیه شده در نرم افزار آزمون بُعدیّت[1] (DIMTEST)، نرم افزار ارزیابی بُعدیّت به منظور شمارش خصیصه‌های مکنون مشمول[2] (DETECT)، تحلیل خوشه‌ای[3] بر اساس زاویه بین بردار سؤال‌ها، تحلیل عاملی سؤال با اطلاعات کامل، تحلیل موازی[4]، روش برآورد بیشینه پسین[5] (MAP) مبتنی بر قضیه بیزی[6]، تحلیل عامل تاییدی[7] (CFA) و روش تجربی بهره‌گیری از نظر متخصصان انجام داده‌اند. ایزانلو و همکاران (۱۳۹۳) در پژوهش خود به شاخص‌های استقلال موضعی از قبیل شاخص Q_3، شاخص X^2 و شاخص G^2 اشاره و بیان کرده‌اند به دلیل در دسترس نبودن نرم افزارهای مورد نیاز، استفاده از این شاخص‌ها رایج نمی‌باشد.

یونسی، دلاور و فلسفی‌نژاد (۱۳۸۹) در بررسی ویژگی‌های روان‌سنجی سؤال‌های تخصصی آزمون فراگیر رشته روان‌شناسی تربیتی دانشگاه پیام نور، برای بررسی فرضیّه تک بُعدی بودن آزمون، از روش پیشنهادی هامبلتون[8] (۱۹۸۹)، یعنی تحلیل عاملی با رویکرد غیرخطی با استفاده از نرم‌افزار تحلیل عاملی سؤال‌ها و نمره‌گذاری آزمون[9] (TESTFACT)، که در آن تشکیل ماتریس ضرایب همبستگی بر اساس همبستگی تتراکوریک به جای همبستگی پیرسون صورت می‌گیرد، استفاده کرده‌اند. از نظر والر[10] (۲۰۰۰، به نقل از یونسی و همکاران، ۱۳۸۹)، تحلیل عاملی خطی هنگامی که با داده‌های چند ارزشی به کار گرفته شود، نتایج تورش داری به‌دست می‌دهد. یونسی و همکاران (۱۳۸۹) به منظور تعیین تعداد ابعاد آزمون، از نسبت حاصل از ارزش ویژه عامل اوّل به

1. dimensionality test
2. dimensionality evaluation to enumerate contributing traits
3. cluster analysis
4. parallel analysis
5. maximum a posteriori estimation
6. Bayes' theorem
7. confirmatory factor analysis
8. Hambleton
9. Test scoring and item factor analysis
10. Waller

ارزش ویژه عامل دوّم، با در نظر گرفتن این اصل که «اگر ارزش ویژه عامل اوّل دو تا سه برابر ارزش ویژه عامل دوّم باشد، می‌توان نتیجه گرفت که آزمون تک بُعدی است و یک عامل غالب، عملکرد افراد را تبیین می‌کند» و نمودار اسکری، استفاده کرده‌اند. یونسی و همکاران (۱۳۸۹)، فرضیّه استقلال موضعی را مورد بررسی قرار نداده‌اند، زیرا از نظر آن‌ها با استناد به لُرد (۱۹۸۰) و هامبلتون (۱۹۸۹)، استقلال موضعی از تک بُعدی بودن پیروی می‌کند و فرضیّه‌ای جدا به حساب نمی‌آید.

به علاوه، مینایی (۱۳۹۳) در بررسی برازش سؤال‌های آزمون مهارت‌های دیداری-حرکتی[1] (TVMS-R) گاردنر[2] (۱۹۹۱)، با استفاده از روش تحلیل مؤلّفه‌های اصلی (PCA) روی باقی مانده‌ها[3] و آزمون t (اسمیت[4]، ۲۰۰۲، به نقل از مینایی، ۱۳۹۳)، نشان داده است که ۶۴ سؤال از ۱۴۲ سؤال این آزمون با مدل راش[5] برازش ندارند و آزمون مذکور بر خلاف ادعای سازنده آن تک بُعدی نمی‌باشد. از نظر مینایی (۱۳۹۳)، به دلیل ضعف آماره‌های سنتی برازش، عاقلانه‌تر این است که جهت بررسی حالت‌های گوناگون نقض استقلال موضعی، از آزمون‌های آماری که مشخصاً برای این منظور طراحی شده‌اند، استفاده نمود. برای مثال می‌توان از همبستگی‌های بین پس مانده‌های سؤال‌ها استفاده نمود که در این صورت همبستگی‌های بالا ممکن است نشان دهنده چند بُعدی بودن یا وابستگی سؤال‌ها با یکدیگر باشد. مینایی (۱۳۹۳) با استناد به مقاله اسمیت (۲۰۰۴) بیان کرده است که آماره‌های برازش از حساسیت لازم برای تشخیص چند بُعدی بودن داده‌ها برخوردار نمی‌باشند و منطقی است که در کنار آماره‌های برازش، از روش تحلیل مؤلّفه‌های اصلی روی پس مانده‌ها نیز استفاده گردد.

در پژوهشی دیگر، احمدی ده‌قطب‌الدینی، خدایی، فرزاد، مقدم زاد و کبیری (۱۳۹۶)، با هدف بررسی بُعدیّت و کارکرد افتراقی سؤال-آزمون مبتنی بر متن پرلز[6] ۲۰۱۱، جهت تحلیل بُعدیّت داده‌ها با استناد به دی مارس[7] (۲۰۱۳)، از مدل تک بُعدی پاسخ مدرّج شده[8]

1. Test of Visual Motor Skills - Revised
2. Gardner
3. principal component analysis of residuals
4. Smith
5. Rasch model
6. progress in international reading literacy study (PIRLS)
7. DeMars
8. graded response model

(GRM) جهت بررسی تک بُعدی بودن داده‌ها و از مدل دو عاملی چند بُعدی نظریّه سؤال-پاسخ[1]، جهت تعیین اثر متن بر انحراف از تک بُعدی بودن داده‌ها استفاده کرده‌اند. نتایج بررسی بُعدیّت توسط آن‌ها نشان داده است که مدل دو عاملی نسبت به مدل پاسخ مدرّج شده، برازش بهتری با داده‌ها دارد. به عبارت دیگر احمدی ده‌قطب‌الدینی و همکاران (۱۳۹۶)، در پژوهش خود از مقایسه شاخص‌های برازندگی دو مدل مذکور برای بررسی بُعدیّت داده‌ها استفاده کرده‌اند.

بُعدیّت ساختار توانایی‌های شناختی در بین پژوهش‌های انجام شده در کشور، کمتر مورد بررسی قرار گرفته است. نجاتی (۱۳۹۲)، پرسشنامه سنجش توانایی‌های شناختی را طراحی و به لحاظ خصوصیات روان‌سنجی مورد بررسی قرار داده است. در نتیجه به-کارگیری روش تحلیل عاملی اکتشافی توسط نجاتی (۱۳۹۲)، هفت عامل شامل حافظه[2]، کنترل مهاری و توجّه انتخابی[3]، تصمیم‌گیری[4]، برنامه‌ریزی[5]، توجّه پایدار[6]، شناخت اجتماعی[7] و انعطاف‌پذیری شناختی[8] از این پرسشنامه استخراج شده است. این هفت عامل در مجموع قادر بوده‌اند ۴۹/۹۶٪ از واریانس داده‌ها را تبیین کنند.

با توجّه به این‌که (الف) ساختار پرسشنامه سنجش توانایی‌های شناختی (نجاتی، ۱۳۹۲)، تاکنون به لحاظ بُعدیّت و همچنین سایر ویژگی‌های روان‌سنجی آن، به خصوص روایی آن برای سنجش توانایی‌های شناختی، کمتر مورد بررسی قرار گرفته است؛ (ب) یک جنبه مهم رواسازی سازه، توضیح تعداد و ماهیّت ابعاد زیربنایی یک آزمون است و مدل‌های نظریه سؤال-پاسخ چند بُعدی (MIRT) می‌توانند ابعاد چندگانه نهفته در پاسخ‌های سؤال‌های یک آزمون را مشخص کنند (امبرتسون و رایس،۲۰۰۰، ترجمه شریفی و همکاران، ۱۳۸۸)؛ (ج) یکی از کاربردهای مهم مدل‌های سؤال-پاسخ چند بُعدی، تحلیل ساختار سؤال-پاسخ‌هایی است که حاصل از اجرای مجموعه‌ای از سؤال‌ها روی نمونه‌ای از آزمودنی‌ها می‌باشد (ریکاس، ۲۰۰۹)؛ (د) ضعف‌هایی که هتی (۱۹۸۴، ۱۹۸۵، به نقل از

1. bi-factor multidimensional IRT model
2. memory
3. inhibitory control & selective attention
4. decision making
5. planning
6. sustain attention
7. social cognition
8. cognitive flexibility

امبرتسون و رایس،۲۰۰۰، ترجمه شریفی و همکاران، ۱۳۸۸) برای بیشتر روش‌های کلاسیک تحلیل بُعدیّت بیان کرده است و پیش از این نیز ذکر شد و (ه) مبتنی بودن تحلیل عاملی به روش نظریّه سؤال-پاسخ بر اطلاعات کامل حاصل از سؤال‌ها و آزمودنی‌ها به جای صرفاً اطلاعات حاصل از همبستگی‌های تتراکوریک یا پلی‌کوریک بین زوج سؤال‌ها (گیبونز و کای، ۲۰۱۸)، این پژوهش قصد دارد تا با به‌کارگیری شاخص‌های وابستگی موضعی Q_3 ین و X^2 پیرسون؛ بررسی پیش‌فرض یکنوایی سؤال‌ها؛ بررسی مدل‌های وابستگی موضعی سؤال‌ها؛ مقایسه شاخص‌های برازش مدل‌ها در نظریّه سؤال-پاسخ؛ تحلیل بُعدیّت فضای عاملی سؤال‌ها و سپس مقایسه روش کلاسیک با روش مبتنی بر نظریّه سؤال-پاسخ چند بُعدی (MIRT)، تعداد ابعاد ساختار توانایی‌های شناختی را بر اساس داده‌های حاصل از اجرای پرسشنامه سنجش توانایی های شناختی نجاتی (۱۳۹۲)، مشخص و مدلی برای ساختار آن ارائه کند.

روش پژوهش

پژوهش حاضر به لحاظ روش، یک پژوهش توصیفی-تحلیلی و از نوع مدل‌سازی رابطه بین سؤال‌ها و پاسخ‌ها به کمک نظریّه سؤال-پاسخ (چند بُعدی) و همچنین مدل‌سازی معادلات ساختاری می‌باشد.

جامعه آماری این پژوهش مشتمل بر کلیه دانش‌آموزان پسر و دختر پایه نهم تحصیلی دوره اوّل متوسطه شهر اهواز در سال تحصیلی ۱۳۹۶-۱۳۹۵، می‌باشد. با توجّه به آمار اخذ شده از اداره کل آموزش و پرورش شهر اهواز، تعداد کل این دانش‌آموزان برابر ۱۷۲۲۶ نفر (۹۱۴۰ پسر و ۸۰۸۶ دختر) بود. از این جامعه نمونه‌ای به صورت تصادفی با حجم ۱۱۰۵ نفر (۵۷۸ پسر و ۵۲۷ نفر دختر) و به نسبت تعداد دبیرستان‌ها و جمعیت دانش‌آموزی از نواحی چهارگانه آموزش و پرورش شهر اهواز، به روش نمونه گیری تصادفی طبقه ای نسبتی[1]، از بین ۳۲ دبیرستان انتخاب شد. جیانگ، وانگ و ویس[2] (۲۰۱۶) معتقدند یک نمونه با حجم ۵۰۰ نفر برای برآورد دقیق پارامترها در مدل‌های نظریّه سؤال-پاسخ چند بُعدی (MIRT)، کفایت می‌کند.

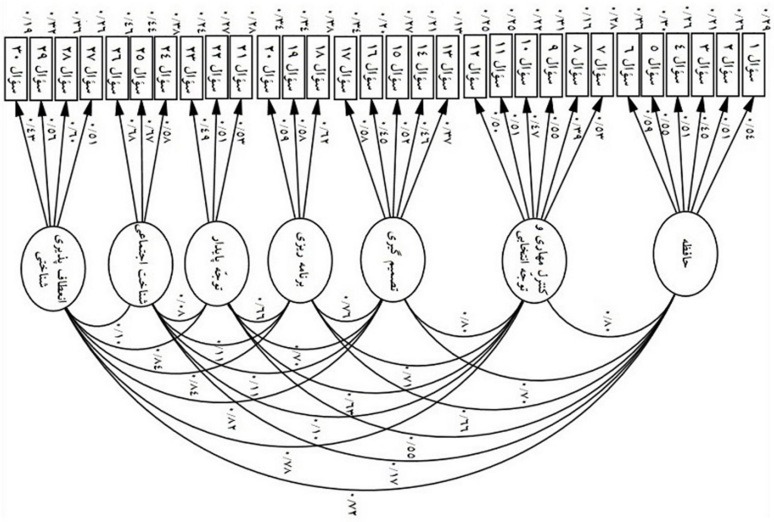
1. proportional stratified
2. Jiang, Wang & Weiss

در این پژوهش، به منظور تحلیل بُعدیّت ساختار توانایی‌های شناختی، از پرسشنامه سنجش توانایی‌های شناختی که توسط نجاتی (۱۳۹۲) طراحی و به لحاظ خصوصیات روان‌سنجی مورد بررسی قرار گرفته است، استفاده شده است. این پرسشنامه شامل ۳۰ سؤال (گویه) و هر سؤال شامل پنج ارزش یا درجه (تقریباً هرگز = ۱ تا تقریباً همیشه = ۵)، می‌باشد. نمره آزمودنی‌ها در این پرسشنامه بین ۳۰ تا ۱۵۰ در تغییر می‌باشد. با به‌کارگیری تکنیک تحلیل عامل اکتشافی توسط نجاتی، هفت عامل (خرده مقیاس) شامل حافظه (سؤال های ۱ تا ۶)، کنترل مهاری و توجّه انتخابی(سؤال های ۷ تا ۱۲)، تصمیم‌گیری (سؤال های ۱۳ تا ۱۷)، برنامه ریزی (سؤال های ۱۸ تا ۲۰)، توجّه پایدار (سؤال های ۲۱ تا ۲۳)، شناخت اجتماعی (سؤال های ۲۴ تا ۲۶) و انعطاف‌پذیری شناختی (سؤال های ۲۷ تا ۳۰) از این مقیاس استخراج شده است. نجاتی پایایی این پرسشنامه را با استفاده از روش آلفای کرونباخ محاسبه و ضریب آلفای کرونباخ را برای کل پرسشنامه ۰/۸۳ گزارش کرده است. علاوه بر این، ضریب همبستگی پیرسون برای بررسی پایایی این پرسشنامه، به روش بازآزمایی، برابر با ۰/۸۶ ($p<$ ۰/۰۰۰۱) گزارش شده است. روایی همزمان این پرسشنامه، با استفاده از روش محاسبه همبستگی بین معدّل تحصیلی و خرده مقیاس‌ها توسط نجاتی مورد بررسی قرار گرفته است و به جز شناخت اجتماعی سایر زیر مقیاس‌ها ضرایب همبستگی معنی‌داری با معدّل تحصیلی داشته‌اند. همچنین، روایی تمیزی این پرسشنامه از طریق به‌کارگیری آزمون تی مستقل روی دوگروه (یک گروه دارای آسیب تحلیلی مغز و یک گروه سالم) نشان داده است که این پرسشنامه از روایی لازم برای تمیز مشکلات شناختی ($t =$ ۴/۴۹۳، p< ۰/۰۰۰۱) برخوردار می‌باشد (نجاتی،۱۳۹۲).

در این پژوهش، به منظور بررسی پایایی پرسشنامه سنجش توانایی‌های شناختی، ضرایب پایایی آلفای کرونباخ و تنصیف اعمال شده در فرمول تصحیح شده اسپیرمن-براون، محاسبه گردید. ضریب آلفای کرونباخ برای کل مقیاس برابر با ۰/۸۶ (پسرها ۰/۸۶ و دخترها ۰/۸۴) و در روش دونیمه کردن در فرمول تصحیح شده اسپیرمن-براون برابر ۰/۸۸ (پسرها ۰/۸۸ و دخترها ۰/۸۸) به‌دست آمد و نشان داد که در مجموع همسانی درونی سؤال‌های پرسشنامه سنجش توانایی‌های شناختی، رضایت بخش می‌باشد. همچنین، بررسی پایایی این پرسشنامه به روش بازآزمایی روی ۱۳۸ دانش‌آموز و به فاصله ۲۰ روز پس از آزمون اوّلیه صورت گرفت. ضریب همبستگی پیرسون بین نمره کل این پرسشنامه طی دو

مرحله اجرا برابر با ۰/۹۲ (پسرها ۰/۸۸، دخترها ۰/۹۷) به‌دست آمد. اندازه این ضرایب نشان دادند که پایایی داده‌ها رضایت بخش می‌باشد.

همچنین در این پژوهش، به منظور بررسی روایی پرسشنامه سنجش توانایی‌های شناختی به روش کلاسیک، پس از بررسی پیش فرض نرمال چند متغیره بودن توزیع داده‌ها، از روش بررسی روایی سازه با به‌کارگیری تحلیل عامل تاییدی (CFA) در نرم افزار AMOS و با اعمال روش برآورد بیشینه درستنمایی[1] (MLE) استفاده شد. بررسی ضرایب رگرسیون استاندارد نشده و اندازه مقادیر بحرانی[2] (C.R) نشان داد که بارهای عاملی همگی به طور معنی‌داری مخالف صفر می‌باشند ($p < ۰/۰۰۰۱$). این موضوع نشان داد که ماده‌ها به طور معنی داری با عامل‌های فرض شده مرتبط هستند (به لحاظ آماری). همچنین نتایج تحلیل عامل تاییدی نشان داد که مقدار بارهای عاملی برای ماده‌های پرسشنامه سنجش توانایی-های شناختی روی عامل‌های هفت‌گانه در این پژوهش، همگی بزرگ‌تر از ۰/۳۷ می‌باشند. شکل ۱، نمودار مسیر، ضرایب استاندارد شده و واریانس تبیین شده سؤال‌های پرسشنامه سنجش توانایی‌های شناختی را نشان می دهد.



شکل ۱. نمودار مسیر، ضرایب استاندارد و واریانس تبیین شده پرسشنامه توانایی‌های شناختی

1. Maximum Likelihood Estimation
2. critical value

به منظور بررسی برازش داده‌های پژوهش با مدل فرضی (هفت عاملی)، شاخص‌های برازش داده‌ها مورد بررسی قرار گرفتند. مقدار شاخص مجذور خی (χ^2) برابر با ۸۵۳/۱۹۸ با ۳۸۴ درجه آزادی ($p < ۰/۰۰۰۱$) و مقدار شاخص مجذور خی هنجار شده[1] (χ^2/df) برابر با ۲/۲۲ به‌دست آمد. با توجّه به این‌که مقدار شاخص مجذور خی هنجار شده کوچک‌تر از ۳ بود، مقدار این شاخص قابل قبول ارزیابی گردید. مقدار شاخص ریشه میانگین مجذورات خطای تقریب[2] (RMSEA) برابر با ۰/۰۳۳، و فاصله اطمینان ۹۰٪ یی آن برابر با (۰/۰۳۶ و ۰/۰۳۰) بود. با توجّه به این‌که مقدار این شاخص کوچک‌تر از ۰/۰۵ بود، در محدوده برازش خوب قرار داشت. مقدار شاخص برازش تطبیقی[3] (CFI) برابر با ۰/۹۲ و مقدار شاخص تاکر- لویس[4] (TLI) برابر با ۰/۹۱، نشان دادند که نیکویی برازش داده‌ها قابل قبول می‌باشد. در مجموع، شاخص‌های کلاسیک برازش مدل، کفایت مدل فرضی را تایید کردند. همچنین مقدار شاخص معیار اطلاعات آکائیک[5] (AIC) برابر با ۱۰۱۵/۹۸ و مقدار شاخص معیار اطلاعات بیزی شوارتز[6] (BIC) برابر با ۱۴۲۰/۸۱۳ به‌دست آمد.

به منظور بررسی روایی ملاکی آزمون (روایی پیش‌بین)، از اندازه ضریب همبستگی بین نمره کل پرسشنامه سنجش توانایی‌های شناختی با میانگین نمره‌های ریاضی پایه هفتم و پایه هشتم و میانگین معدّل‌های کل پایه هفتم و پایه هشتم دانش‌آموزان (به عنوان دو ملاک) استفاده شد. ضرایب همبستگی بین نمره کل پرسشنامه سنجش توانایی‌های شناختی و دو ملاک مذکور بیشتر از ۰/۴۵ به‌دست آمد (۰/۴۶ با میانگین نمره ریاضی پایه هفتم و هشتم و ۰/۴۹ با میانگین معدّل پایه هفتم و هشتم). با توجّه به این موضوع که هر چه این ضرایب بیشتر باشند، آزمون دارای روایی بیشتری است، روایی ملاکی (پیش‌بین) آزمون توانایی‌های شناختی در حد قابل قبول ارزیابی شد.

در این پژوهش و به منظور تحلیل داده‌ها از نسخه ۲۴ نرم‌افزار تحلیل ساختارهای گشتاوری[7] (AMOS)؛ نسخه ۲۵ نرم‌افزار بسته آماری برای علوم اجتماعی[8] (SPSS)؛ نسخه

1. Normed Chi-square
2. Root Mean Square Error of Approximation
3. Comparative Fit Index
4. Tucker-Lewis Index
5. Akaike Information Criterion
6. Schwarz's Bayesian Information Criterion
7. Analysis of Moment Structures
8. Statistical Package for the Social Sciences

۳/۴/۴ نرم‌افزار R؛ نسخه ۱/۲۹ بسته نرم‌افزاری نظریه سؤال-پاسخ چند بُعدی برای محیط R[1] (mirt)؛ نسخه ۲/۴/۳ منوی R-Factor در نسخه ۲۵ نرم‌افزار SPSS و نسخه ۴/۲ نرم-افزار نظریه سؤال-پاسخ برای نتایج گزارش شده توسط بیمار[2] (IRTPRO)، استفاده شده است.

نتایج

با استناد به نظر هاتس و ادواردز[3] (۲۰۱۳) و با توجّه به این‌که مقیاس پاسخ‌ها در این پژوهش از نوع لیکرت[4] و مبتنی بر خودگزارشی[5] بود، برای مدل مورد استفاده در تحلیل داده‌ها، پارامتر حدس لحاظ نگردید. در این پژوهش و به منظور تحلیل داده‌ها، مدل‌های تک بُعدی پاسخ مدرّج شده (GRM) سِیم‌جیما[6] (۱۹۶۹) و مدل تعمیم یافته آن برای حالت‌های چند بُعدی یعنی مدل چند بُعدی پاسخ مدرّج شده[7] (MGRM) موراکی و کارلسون[8] (۱۹۹۵)، مورد استفاده قرار گرفتند. سِیم‌جیما (۱۹۶۹ و ۱۹۷۲) مدل پاسخ مدرّج شده را به عنوان یک مدل سؤال-پاسخ که با سؤال‌های دارای دو یا بیش از دو مقوله پاسخ (مدرّج یا ترتیبی)، قابلیت برازش را دارد، معرفی کرده است. در این مدل، برای یک سؤال یا ماده با C مقوله پاسخ، یک پارامتر تمیز و C-۱ پارامتر آستانه[9] محاسبه می‌شود. در این مدل احتمال این‌که آزمودنی، p اُمین مقوله را از بین پاسخ‌ها انتخاب کند، بر اساس نماد گذاری هاتس و ادواردز (۲۰۱۳) از رابطه ۸ به دست می‌آید:

$$P(u_{\mathrm{i}} = p \mid \theta) = \frac{1}{1 + \mathrm{e}^{(-a_{\mathrm{j}}(\theta - b_{\mathrm{jp}}))}} - \frac{1}{1 + \mathrm{e}^{(-a_{\mathrm{j}}(\theta - b_{\mathrm{j(p+1)}}))}} = P^*(p) - P^*(p+1). \quad (۸)$$

در رابطه ۸، a_{j} پارامتر تمیز و b_{jp} مقدار پارامتر آستانه برای مقوله j می‌باشد. برای سهولت در محاسبات، فرض می‌شود که $P^*(0) = 1$ و $P^*(c+1) = 0$ (هاتس و ادواردز، ۲۰۱۳).

1. a multidimensional item response theory package for the R environment
2. Item Response Theory for Patient-Reported Outcomes
3. Houts & Edwards
4. likert
5. self-report
6. Samejima
7. multidimensional graded response model
8. Muraki & Carlson
9. threshold parameter

در ادامه، پیش فرض‌های استقلال موضعی و یکنوایی سؤال‌ها مورد بررسی و بحث قرار گرفته است.

الف (۱). بررسی پیش‌فرض استقلال موضعی سؤال‌ها روی یک بُعد

به منظور بررسی پیش فرض استقلال موضعی بین سؤال‌های پرسشنامه سنجش توانایی‌های شناختی، مقادیر شاخص X^2 پیرسون با اعمال مدل پاسخ مدرّج شده سِیم‌جیما و در نظر گرفتن یک بُعد برای کل سؤال‌ها، با استفاده از نرم افزار IRTPRO محاسبه گردید. برای این منظور با به‌کارگیری الگوریتم بیشینه سازی انتظار[1] (EM) باک و ایتکین[2] (۱۹۸۱) (پس از ۶۷ تکرار[3])، فراوانی‌های مورد انتظار برای هر یک از مقوله‌های پاسخ سؤال‌ها برآورد شدند. سپس با به‌کارگیری روش برآورد بیشینه درستنمایی (MLE)، مقادیر شاخص X^2 محاسبه شدند. این مقادیر در جدول ۱ آمده است.

جدول ۱. مقادیر شاخص X^2 سؤال‌های پرسشنامه سنجش توانایی‌های شناختی روی یک بُعد

سؤال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵
۲	۱۰/۱														
۳	۴/۴	۱/۲													
۴	۱/۴	۱/۳	۵/۵												
۵	۱/۱	−۰/۵	۶/۵	۶/۱											
۶	−۰/۶	۱/۱	۱/۷	۳/۳	۴										
۷	۱/۴	۴/۱	۳/۵	۴/۷	۳/۶	۲/۵									
۸	۳	۱/۴	۱/۷	−۰/۶	۳	۰	۶/۱								
۹	۱/۲	۱/۸	۱/۵	۳/۹	۳/۶	۰/۲	۲/۴	۱/۸							
۱۰	−۰/۱	۳/۲	−۱	۱/۱	۰/۹	۲/۹	۰/۶	۱/۲	۰/۷						
۱۱	۵/۹	۱/۱	۱/۱	۳/۷	۱/۱	۲	۲/۴	۴/۸	۰	۲					
۱۲	۲/۳	۲/۱	−۰/۹	۱/۲	۰/۹	−۱	۵/۱	۱/۴	۱/۸	۲/۵	۸/۳				
۱۳	۱/۸	۰/۴	۱/۷	۱	−۱/۷	−۰/۱	−۰/۳	۲/۴	−۰/۱	۱/۳	۰/۳	۰/۲			
۱۴	۲/۷	۱/۶	۱/۲	۳/۳	۲/۲	−۰/۵	۳/۹	۳/۹	۰/۵	۳/۹	۱/۲	۲/۸	۱۰/۱		
۱۵	۱/۲	۰/۵	۰/۵	۳/۱	۰	۰/۹	۰/۷	۱/۳	۰/۸	۱	۰	۳/۸	۰/۶	۴/۲	
۱۶	۱	۰/۸	−۰/۶	۳/۴	۱/۴	۲/۸	۳	۱/۵	۰/۱	۰/۱	۰/۵	۲/۸	۰	۲/۷	۲/۳
۱۷	۱/۸	۱/۲	۱/۴	۵/۱	۲/۵	۱/۱	۲/۴	۱/۷	۴/۵	۳/۱	۴/۷	۵/۸	۲/۳	۰/۷	۴/۷
۱۸	۱/۴	۰/۳	−۰/۲	۵/۳	۱/۸	۱/۱	۱/۳	۴	۲/۳	۰/۸	۳/۱	۲/۴	۱/۹	۰/۷	۲/۷
۱۹	۰/۹	−۱/۱	۱/۳	۲	۲/۹	۱/۳	−۰/۵	۲/۹	۴/۸	۰/۵	۴/۳	۲	۲/۸	۶/۳	۵/۴

1. expectation-maximization algorithm
2. Bock & Aitkin
3. iteration

۲۰	-۰/۱	-۰/۵	۰/۲	۱/۳	۰/۲	۰/۱	۳/۱	۲/۶	۰/۸	۳/۳	۳/۵	۲/۱	۴/۲	۱	۱/۹
۲۱	-۰/۷	-۰/۱	-۰/۴	-۰/۴	۳/۶	۲/۴	۱/۷	۰/۵	۱/۵	۱/۳	۴/۸	۰/۵	۱/۲	۲/۹	۰/۲
۲۲	۰/۲	۰/۶	۲/۳	۱/۵	۰/۵	۱/۲	۲/۷	۱	۳/۱	۰/۶	۴/۸	۰/۵	۰/۶	-۰/۱	۳/۷
۲۳	۱/۳	-۰/۴	۰/۴	۲/۶	۰/۸	۱/۴	۱/۲	۵/۳	۲/۴	۱/۷	۱/۶	۲/۲	۲/۸	۰/۶	۳/۴
۲۴	۱	۰/۴	۲	۱/۲	-۰/۴	۱/۳	-۰/۱	۶/۸	۱/۶	۱/۶	۳/۸	۱/۹	۳/۳	۵/۷	۶/۳
۲۵	۲/۳	۰/۹	۲/۴	۱/۲	۰/۷	۳/۷	۰	۳/۲	۳/۸	-۰/۴	۱/۷	۳/۳	۳/۸	۳/۶	۵/۲
۲۶	۳/۱	-۰/۷	-۰/۳	۱/۱	-۰/۴	۱/۱	۱/۱	۳/۹	۳/۹	۰/۷	۳/۲	۴/۷	۲/۵	۱/۴	۴/۳
۲۷	۵/۴	۲/۶	۰/۲	۳/۳	۶/۷	۶/۶	۶/۲	۳/۹	۲/۷	۳/۳	۷/۶	۷/۶	۴/۶	۲/۹	۵/۹
۲۸	۲/۶	۲/۵	۱	۵/۲	۱/۲	۲/۴	۳/۸	۱/۲	۲/۷	۵/۴	۳/۷	۴/۶	۱/۴	-۰/۵	۲/۱
۲۹	۱/۷	۳	-۰/۴	۵	۱/۸	۱/۹	۳/۵	۰/۵	۲/۳	۱/۱	۱/۹	۴/۲	۱/۶	۱/۲	۳/۴
۳۰	۰/۹	۲/۹	۲/۱	۲	۰/۸	۱/۳	۰/۸	۶/۶	۲/۵	۰/۶	۰/۳	۲/۵	۰/۳	۱/۵	۱/۲
سؤال	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹	
۱۷	۲/۴														
۱۸	۵/۴	۳/۳													
۱۹	۲/۲	۱/۶	۵/۳												
۲۰	۱/۹	۳/۶	۲/۷	۲/۴											
۲۱	۸/۶	۱/۹	۱/۷	۱	۲/۹										
۲۲	۲/۱	۵/۱	-۰/۳	۰/۳	-۰/۶	۲/۱									
۲۳	۳/۸	۶	۳	۱/۹	۴/۴	۱۰/۳	۱/۹								
۲۴	۴/۱	۲/۳	۵/۹	۲/۶	۳	۴/۷	۲/۴	۳/۶							
۲۵	۲/۲	۲/۶	۴/۲	۴	۲/۵	۴/۹	-۱/۳	۲/۹	**۳۵/۲**						
۲۶	۳/۴	۵/۲	۴/۴	-۱/۱	۱	۶	۰/۸	۴/۳	**۳٤/۱**	**٤۷/۸**					
۲۷	-۰/۵	۵/۸	۳/۵	۲	۲/۴	۷/۱	۲/۷	۳/۹	۴/۱	۵/۷	۵/۹				
۲۸	-۰/۱	۱/۱	۴/۴	۲/۴	۱/۹	۱	۴/۲	۲/۲	۳/۴	۴/۴	۱/۸	۳/۵			
۲۹	۳/۹	-۰/۷	۲	۰	۱/۲	۰/۱	۱/۷	۰/۸	۳/۴	۱/۱	۰/۳	۲/۲	۱/۶		
۳۰	۰/۶	۲/۵	۵/۸	۰/۵	۲/۷	۴/۴	۴/۶	۴/۶	۱/۹	۱/۳	۰/۲	۱/۹	۱/۵	۳/۷	

همان‌گونه که پیش از این ذکر شد، بر اساس نظر کای و همکاران (۲۰۱۱)، سؤال‌هایی که مقدار شاخص X^2 برای آن‌ها بزرگتر از ۱۰ باشد، احتمال وجود وابستگی موضعی با یکدیگر را نشان می‌دهند. با توجّه به مقادیر آمده در جدول ۱، مشاهده می‌شود که مقادیر شاخص X^2 برای سؤال های ۲۴، ۲۵ و ۲۶ (که به صورت پررنگ و با خطی زیر آن‌ها مشخص شده‌اند)، بزرگتر است ۳۳ می‌باشند. اندازه بزرگی مقادیر شاخص X^2 نشان دهنده وجود وابستگی موضعی بین این سه سؤال یا همان‌گونه که کای و همکاران (۲۰۱۱) بیان می‌کنند، نشان دهنده واریانس باقی مانده‌ای می‌باشد که مدل تک بُعدی در نظر گرفته شده، قادر به تبیین آن نمی‌باشد. به منظور بررسی بیشتر و در نظر نگرفتن تنها یک شاخص

برای قضاوت مبنی بر وجود وابستگی موضعی بین سؤال‌های مذکور، شاخص Q_3 ین نیز با استفاده از بسته نرم افزاری mirt در نرم افزار R و با اعمال مدل پاسخ مدرّج شده سِیم‌جیما بر داده‌ها، با به‌کارگیری الگوریتم باک و ایتکین (EM) (پس از ۶۷ تکرار)، محاسبه شدند (جدول ۲ ملاحظه شود).

جدول ۲. مقادیر شاخص Q_3 سؤال‌های پرسشنامه سنجش توانایی‌های شناختی روی یک بُعد

سؤال	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰
۲	۰/۲۱									
۳	۰/۰۲	۰/۰۶								
۴	۰/۰۱	۰/۰۶	۰/۱۸							
۵	−۰/۰۲	−۰/۰۲	۰/۱۷	۰/۱۴						
۶	۰/۰۱	۰/۰۱	۰/۰۰۱	۰/۰۱	۰/۱۲					
۷	−۰/۰۴	۰/۰۰۲	۰/۰۰۲	۰/۰۷	۰/۱۰	۰/۰۶				
۸	−۰/۱۱	−۰/۰۳	۰/۰۱	۰/۰۰۲	−۰/۱۱	−۰/۰۳	۰/۰۰۲			
۹	−۰/۰۵	۰/۰۰۸	−۰/۰۰۴	۰/۰۰۳	۰/۰۰۷	−۰/۰۷	۰/۰۳	−۰/۰۱		
۱۰	۰/۰۰۷	−۰/۰۲	−۰/۰۵	−۰/۰۰۴	−۰/۰۲	−۰/۰۲	۰/۰۳	۰/۰۲	۰/۰۲	
۱۱	−۰/۱۲	−۰/۱	−۰/۱۰	−۰/۱۱	−۰/۰۳	−۰/۰۵	−۰/۰۶	۰/۰۴	−۰/۰۴	۰/۰۲
۱۲	−۰/۰۳	−۰/۰۷	−۰/۰۲۹	۰/۰۳	−۰/۰۲	۰/۰۰۲	۰/۰۱	۰/۰۲	۰/۰۱	۰/۰۶
۱۳	−۰/۰۳	−۰/۰۴	۰/۰۱	−۰/۰۸	−۰/۰۲	−۰/۰۲	−۰/۰۲	−۰/۰۰۸	−۰/۰۰۸	۰/۰۵
۱۴	−۰/۰۶	−۰/۰۲	−۰/۰۲	−۰/۰۴	−۰/۰۸	−۰/۰۶	−۰/۱۰	۰/۰۱	−۰/۰۱	−۰/۰۱
۱۵	−۰/۰۷	−۰/۰۶	−۰/۰۱	−۰/۰۸	−۰/۰۵	−۰/۰۱	−۰/۰۸	−۰/۰۱	−۰/۰۷	−۰/۰۹
۱۶	−۰/۰۹	−۰/۱۱	−۰/۰۰۷	−۰/۰۷	−۰/۰۳	−۰/۰۹	−۰/۱۰	۰/۰۲	−۰/۰۴	۰/۰۰۸
۱۷	−۰/۰۶	−۰/۰۵	−۰/۱۰	−۰/۰۸	−۰/۰۹	۰/۰۰۸	−۰/۰۴	۰/۰۰۶	−۰/۰۵	−۰/۰۴
۱۸	−۰/۰۷	−۰/۱۲	−۰/۰۶	−۰/۰۸	−۰/۰۹	−۰/۰۹	−۰/۰۵	۰/۰۰۲	−۰/۰۲	−۰/۰۸
۱۹	−۰/۰۱	−۰/۰۸	−۰/۰۲	−۰/۰۶	−۰/۰۴	−۰/۱۱	−۰/۰۴	−۰/۰۶	−۰/۰۱	−۰/۰۹
۲۰	−۰/۰۸	−۰/۰۳	−۰/۰۵	−۰/۰۸	−۰/۰۹	−۰/۰۹	−۰/۰۸	−۰/۰۴	−۰/۰۸	−۰/۱۴
۲۱	−۰/۰۴	−۰/۰۲	−۰/۰۲	−۰/۰۲	−۰/۱۱	−۰/۱۰	−۰/۰۷	۰/۰۱	−۰/۱۱	−۰/۰۴
۲۲	−۰/۰۲	−۰/۰۵	−۰/۰۵	−۰/۰۶	−۰/۰۰۳	−۰/۱۴	−۰/۰۶	−۰/۰۳	−۰/۰۶	−۰/۰۲
۲۳	−۰/۰۶	−۰/۰۷	−۰/۰۳	−۰/۰۲	−۰/۰۹	−۰/۱۰	−۰/۰۳	−۰/۰۵	−۰/۱۰	−۰/۰۵
۲۴	۰/۰۱	۰/۰۰۹	−۰/۰۱	−۰/۰۱	۰/۰۶	۰/۰۱	−۰/۰۱	−۰/۱۱	۰/۰۱	−۰/۰۲
۲۵	۰/۰۲	۰/۰۵	−۰/۰۳	−۰/۰۰۶	۰/۰۱	−۰/۰۰۱	−۰/۰۰۶	−۰/۰۲	−۰/۰۰۶	−۰/۰۳
۲۶	۰/۰۲	−۰/۰۰۸	−۰/۰۴	−۰/۰۰۷	−۰/۰۱	۰/۰۵	۰/۰۱	−۰/۰۱	۰/۰۱	۰/۰۳
۲۷	۰/۰۸	۰/۰۷	−۰/۰۵	−۰/۰۷	−۰/۰۹	−۰/۰۵	−۰/۰۴	−۰/۰۸	−۰/۱۳	−۰/۰۵
۲۸	−۰/۰۹	−۰/۰۷	−۰/۰۴	−۰/۱۱	−۰/۰۷	−۰/۱۰	−۰/۱۵	−۰/۰۴	−۰/۰۸	−۰/۰۲
۲۹	−۰/۰۴	−۰/۰۳	−۰/۱۰	−۰/۰۸	−۰/۱۰	−۰/۰۹	−۰/۱۴	−۰/۰۵	−۰/۰۸	−۰/۰۱
۳۰	−۰/۱۱	−۰/۱۰	−۰/۰۴	−۰/۰۸	−۰/۰۶	−۰/۰۶	−۰/۰۲	−۰/۰۴	−۰/۰۸	−۰/۰۳
سؤال	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰

۱۲	۰/۱۲									
۱۳	−۰/۰۰۳	۰/۰۴								
۱۴	۰/۰۳	−۰/۰۶	۰/۰۳							
۱۵	۰/۰۲	−۰/۱۱	۰/۰۰۸	۰/۰۸						
۱۶	−۰/۰۳	−۰/۰۶	۰/۰۰۷	۰/۰۲	۰/۰۱					
۱۷	−۰/۰۱	−۰/۰۵	۰/۰۲	۰/۰۴	۰/۰۲	۰/۰۱				
۱۸	−۰/۰۶	−۰/۱۰	−۰/۰۲	۰/۰۴	−۰/۰۴	−۰/۰۷	−۰/۰۰۵			
۱۹	−۰/۱۵	−۰/۰۷	−۰/۰۴	−۰/۰۰۲	−۰/۰۲	−۰/۰۵	−۰/۰۷۹	۰/۰۸		
۲۰	−۰/۰۴	−۰/۰۸	−۰/۰۷	−۰/۰۴	−۰/۰۱	−۰/۰۰۱	−۰/۱۲۳	۰/۰۵	۰/۰۹	
۲۱	−۰/۰۹	−۰/۰۳	۰/۰۰۹	−۰/۰۵	−۰/۰۰۹	۰/۱۴	−۰/۰۶۵	−۰/۰۷	−۰/۰۵	۰/۰۱
۲۲	۰/۰۰۱	−۰/۰۷	−۰/۰۸	−۰/۰۰۸	−۰/۰۲	۰/۰۷	۰/۰۰۶	−۰/۰۷	−۰/۰۳	−۰/۰۲
۲۳	−۰/۰۰۶	−۰/۰۵	−۰/۰۶	−۰/۰۵	−۰/۰۷	۰/۰۵	−۰/۱۳	−۰/۰۵	−۰/۰۱	۰
۲۴	−۰/۰۴	−۰/۰۵	−۰/۰۵	−۰/۰۴	−۰/۰۳	−۰/۰۶	۰/۰۱	−۰/۰۳	۰/۰۰۴	−۰/۰۰۳
۲۵	−۰/۰۴	−۰/۱۱	−۰/۰۴	−۰/۰۱	۰/۰۴	−۰/۰۲	۰/۰۱	۰	−۰/۰۱	−۰/۰۳
۲۶	−۰/۰۵	−۰/۰۶	−۰/۰۳	−۰/۰۰۵	−۰/۰۱	−۰/۰۷	۰/۰۱	−۰/۰۳	۰/۰۱	−۰/۰۵
۲۷	−۰/۰۵	−۰/۰۸	−۰/۰۱	−۰/۰۴	−۰/۱۴	−۰/۰۱۶	۰/۰۶	−۰/۰۶	−۰/۰۱	−۰/۰۲
۲۸	۰/۰۵	۰/۰۵	۰/۰۰۳	−۰/۰۶	−۰/۰۳	−۰/۰۲	−۰/۰۲	−۰/۰۹	−۰/۰۶	−۰/۰۶
۲۹	−۰/۰۵	−۰/۱۴	−۰/۰۵	−۰/۰۳	۰/۰۰۸	۰/۰۰۳	−۰/۰۲	−۰/۰۰۲	−۰/۰۱	−۰/۰۰۴
۳۰	−۰/۰۲	−۰/۰۴	−۰/۰۱	−۰/۰۴	−۰/۰۱	−۰/۰۲	−۰/۰۸	−۰/۰۲	−۰/۰۰۶	۰/۰۵۴
سؤال	۲۱	۲۲	۲۳	۲۴	۲۵	۲۶	۲۷	۲۸	۲۹	۳۰
۲۲	۰/۰۵									
۲۳	۰/۱۷	۰/۰۹								
۲۴	−۰/۰۴	−۰/۰۱	−۰/۰۵							
۲۵	−۰/۰۷	۰/۰۳	−۰/۰۴	**۰/۳۸**						
۲۶	−۰/۱۲	−۰/۰۳	−۰/۰۸	**۰/۳۸**	**۰/٤٤**					
۲۷	−۰/۰۲	۰/۰۳	۰/۰۱	−۰/۰۸	−۰/۱۰	−۰/۰۶				
۲۸	−۰/۰۰۴	−۰/۰۰۵	۰/۰۵	−۰/۰۷	−۰/۱۰	−۰/۰۲	۰/۰۵			
۲۹	۰/۰۱	۰/۰۱	−۰/۰۱	۰/۰۴	۰/۰۹	−۰/۰۰۷	۰/۰۲	−۰/۰۲		
۳۰	۰/۰۴	−۰/۰۲	۰/۰۹	۰/۰۲	−۰/۰۰۲	−۰/۰۰۷	۰/۰۲	۰/۰۲	۰/۰۰۵	

همان‌گونه که پیش از این اشاره شد، نقطه برش برای مقایسه مقادیر شاخص Q_3 برابر ۰/۲ می‌باشد. در صورتی که این مقدار را با مقادیر جدول ۲ مقایسه کنیم، متوجّه خواهیم شد که مقادیر شاخص Q_3 برای سؤال های ۲۴، ۲۵ و ۲۶ (که به صورت پررنگ و با خطی در زیر آن‌ها مشخص شده است)، بزرگ‌تر از ۰/۲ می‌باشند و این یافته‌ها نتایج به‌دست آمده از شاخص X^2 را تایید کردند.

محتوای سؤال‌های ۲۴، ۲۵ و ۲۶ که خرده مقیاس شناخت اجتماعی را تشکیل می‌دهند، عبارت است از:

سؤال ۲۴- در صورتی که یک فرد در یک جلسه اجتماعی معذّب باشد، من کاری می‌کنم که فرد حس راحت‌تری پیدا کند.

سؤال ۲۵- به این که دیگران به حرفهایم گوش کنند، توجّه می‌کنم.

سؤال ۲۶- می‌توانم منظور افراد را با نگاه کردن به آن‌ها متوجّه شوم.

بحث (۱). با توجّه به این‌که در پاسخ‌گویی به سؤال‌های پرسشنامه توانایی‌های شناختی، (الف) وقت کافی در اختیار آزمودنی‌ها قرار داده شده بود؛ (ب) همه آزمودنی‌ها به سؤال‌ها پاسخ داده بودند و (ج) این سؤال‌ها جزء سؤال‌های انتهایی پرسشنامه نبودند، بنابراین فرضیّه وابستگی موضعی ناشی از مکان (یا پارامتر سرعت در پاسخ‌گویی به سؤال‌ها) قابل طرح نبود. همچنین، بررسی محتوای این سه سؤال نشان داد که فرضیّه وابستگی موضعی ناشی از یکسان بودن محتوای سؤال‌ها، که در آن احتمال انتخاب دو مقوله پاسخ برای دو سؤال متفاوت، برابر است با احتمال انتخاب یکی از مقوله‌ها برای هر یک از سؤال‌ها (رابطه ۴)، نیز قابل طرح نبود. زیرا اوّلاً محتوای سؤال‌ها متفاوت است و ثانیاً احتمال انتخاب هر یک از مقوله‌های پاسخ، متفاوت است. جدول ۳ احتمال انتخاب هر یک از مقوله‌های پاسخ برای این سه سؤال در کل نمونه که به کمک شاخص‌های آماری سه سؤال ۲۴، ۲۵ و ۲۶ و با استفاده از نرم افزار SPSS به‌دست آمده است را نشان می‌دهد و این موضوع را تایید می‌کند.

جدول ۳. احتمال انتخاب هر یک از مقوله‌های پاسخ سؤال های ۲٤، ۲۵ و ۲٦

	احتمال				
سوال	مقوله ۱	مقوله ۲	مقوله ۳	مقوله ۴	مقوله ۵
۲۴	۰/۱۳۳	۰/۱۶	۰/۲۲۴	۰/۲۳۶	۰/۲۴۷
۲۵	۰/۱۰۸	۰/۰۸۴	۰/۱۲۶	۰/۲۱	۰/۴۷۲
۲۶	۰/۰۹	۰/۱۱۳	۰/۱۷۶	۰/۲۷۱	۰/۳۴۹

همچنین مقادیر π_{LD} برای زوج سؤال‌های ۲۴، ۲۵ و ۲۶ و بر اساس آن‌چه در رابطه (۶) بیان شد، محاسبه شدند. این مقادیر برای این سؤال‌ها برابر بودند با: $\pi_{LD}\left(۲۴, ۲۵\right) = ۰/۳۳$، $\pi_{LD}\left(۲۴, ۲۶\right) = ۰/۳۴$ و $\pi_{LD}\left(۲۵, ۲۶\right) = ۰/۳۷$. با توجّه به این مقادیر، احتمال این‌که

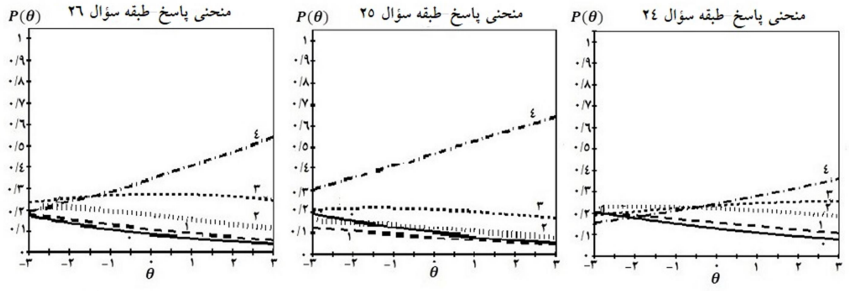
آزمودنی‌ها به این زوج سؤال‌ها به طور مشابه پاسخ داده باشند، کم و این مقادیر اختلاف زیادی با مقدار یک داشتند. بنابراین و با توجّه به این یافته‌ها، مدل وابستگی موضعی سطحی (SLD) نمی‌توانست به عنوان مدل وابستگی موضعی مشاهده شده بین سؤال‌های ۲۴، ۲۵ و ۲۶، در نظر گرفته شود. به همین دلیل در ادامه این فرضیّه مورد بررسی قرار گرفت که وابستگی‌های موضعی مشاهده شده بین سه سؤال ۲۴، ۲۵ و ۲۶، ناشی از وجود بُعدی متفاوت با بُعد سایر سؤال‌های پرسشنامه می‌باشد، و در نتیجه وابستگی‌های مشاهده شده بین این سه سؤال، از نوع وابستگی اساسی (ULD) می‌باشد. در ادامه و در قسمت الف (۲) این فرضیّه مورد بررسی قرار گرفته است.

ب (۱). بررسی پیش‌فرض یکنوایی سؤال‌ها روی یک بُعد

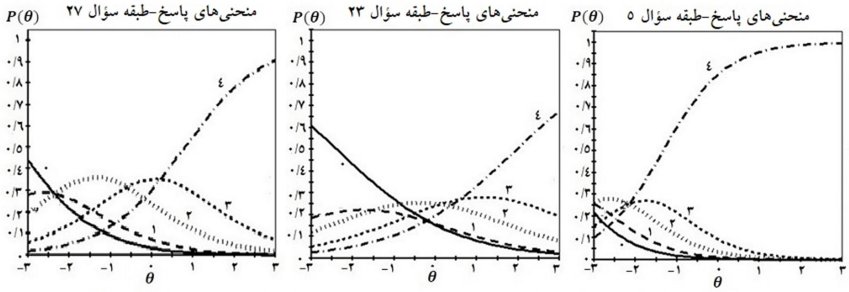
تجزیه و تحلیل بیشتر به منظور مشخص کردن این موضوع که آیا سه سؤال ۲۴، ۲۵ و ۲۶ پرسشنامه سنجش توانایی‌های شناختی روی بُعدی متفاوت از بُعد سایر سؤال‌ها قرار می‌گیرند و خصیصه مکنون و زیربنایی این سه سؤال متفاوت از سایر سؤال‌ها است، از طریق بررسی منحنی‌های پاسخ–طبقه[1] (CRC) سؤال‌ها انجام گرفت. این منحنی‌ها برای سؤال‌های ۲۴، ۲۵ و ۲۶ در شکل ۲ و برای سه سؤال ۵، ۲۳ و ۲۷ (به عنوان نمونه) از بین ۲۷ سؤال دیگر، در شکل ۳ نشان داده شده است. شکل ۲ نشان می‌دهد که منحنی‌های پاسخ–طبقه برای سؤال‌های ۲۴، ۲۵ و ۲۶ تقریباً به یکدیگر شبیه هستند و پیش‌فرض یکنوایی برای آن‌ها برقرار نمی‌باشد. به عبارت دیگر، احتمال انتخاب یکی از طبقات پاسخ ($P(\theta)$)، متناسب با افزایش سطح صفت مکنون (θ)، افزایش نمی‌یابد. (مثلاً اختلاف احتمال انتخاب طبقه صفر که نشان دهنده انتخاب گزینه «تقریباً هرگز» توسط آزمودنی‌ها می‌باشد، برای آزمودنی‌ها با سطح صفت بالا و پایین در هر سه منحنی حدوداً ۰/۱ است). در حالی‌که این وضعیت در مورد منحنی‌های پاسخ–طبقه سه سؤال دیگر (سؤال های ۵، ۲۳ و ۲۷) که از بین سایر سؤال‌ها و به منظور مقایسه انتخاب شده‌اند، متفاوت است. در این سؤال‌ها پیش‌فرض یکنوایی برقرار است و با افزایش سطح صفت آزمودنی‌ها، احتمال انتخاب طبقات پایینی (که در منحنی با اعداد ۰، ۱ و ۲ نشان داده شده و به ترتیب معادل انتخاب گزینه‌های «تقریباً هرگز»، «به‌ندرت» و «گاهی‌اوقات» از سوی آزمودنی‌ها می‌باشد) کاهش و احتمال انتخاب طبقات بالایی (که در منحنی با اعداد ۳ و ۴ نشان داده شده و به ترتیب معادل انتخاب

1. category response curves

گزینه‌های «اغلب» و «تقریباً همیشه» از سوی آزمودنی‌ها می‌باشد) افزایش می‌یابد. این موضوع نشان داد که تک خصیصه مکنون زیربنایی یا تک بُعد در نظر گرفته شده، قادر است به خوبی پاسخ‌های آزمودنی‌ها به این سؤال‌ها را بر خلاف سؤال‌های ۲۴، ۲۵ و ۲۶ مدل سازی کند.



شکل ۲. منحنی‌های پاسخ-طبقه سؤال‌های ۲٤، ۲۵ و ۲٦ مقیاس سنجش توانایی‌های شناختی در مدل تک بُعدی



شکل ۳. منحنی‌های پاسخ-طبقه سؤال‌های ۵، ۲۳ و ۲۷ مقیاس سنجش توانایی‌های شناختی در مدل تک بُعدی

ج (۱). بررسی پارامترها و بارهای بُعدی مدل سؤال-پاسخ تک بُعدی

به منظور بررسی پارامترهای مدل تک بُعدی، پارامترهای این مدل به همراه خطای استاندارد (s.e) هر یک از سؤال‌ها با اعمال مدل تک بُعدی پاسخ مدرّج شده (GRM) سِیم جیما و با استفاده از نرم افزار IRTPRO محاسبه شدند (جدول ۴ ملاحظه شود).

جدول ٤. مقادیر پارمترهای شیب، آستانه بین طبقات و بارهای بُعدی مدل تک بُعدی

سؤال	a	$s.e$	b_1	$s.e$	b_2	$s.e$	b_3	$s.e$	b_4	$s.e$	λ	$s.e$
۱	۱/۱۴	۰/۰۸	-۳/۷۶	۰/۲۸	-۲/۶۱	۰/۱۸	-۰/۷۷	۰/۰۷	۰/۸۳	۰/۰۸	۰/۵۶	۰/۰۵
۲	۰/۹۴	۰/۰۸	-۴/۲۲	۰/۳۵	-۲/۵۶	۰/۲۰	-۱/۰۵	۰/۱۰	۰/۵۴	۰/۰۹	۰/۴۸	۰/۰۵
۳	۰/۹۹	۰/۱۰	-۴/۱۲	۰/۴۰	-۳/۳۵	۰/۳۱	-۲/۶۰	۰/۲۳	-۱/۵۵	۰/۱۴	۰/۵۰	۰/۰۷
۴	۰/۹۱	۰/۰۸	-۵/۴۵	۰/۵۳	-۳/۳۱	۰/۲۸	-۱/۸۹	۰/۱۶	-۰/۲۰	۰/۰۷	۰/۴۷	۰/۰۵

۵	۱/۳۱	۰/۱۲	−۳/۹۸	۰/۳۵	−۳/۰۸	۰/۲۴	−۲/۱۹	۰/۱۶	−۱/۳۴	۰/۱۰	۰/۶۱	۰/۰۵
۶	۱/۳۷	۰/۱۰	−۳/۵۳	۰/۲۵	−۲/۵۵	۰/۱۶	−۱/۴۴	۰/۰۹	−۰/۲۹	۰/۰۵	۰/۶۳	۰/۰۵
۷	۱/۰۸	۰/۰۸	−۳/۶۸	۰/۲۷	−۲/۲۱	۰/۱۵	−۰/۷۴	۰/۰۸	−۰/۷۶	۰/۰۹	۰/۵۳	۰/۰۵
۸	۰/۷۲	۰/۰۷	−۴/۴۸	۰/۴۲	−۲/۶۶	۰/۲۴	−۰/۶۸	۰/۱۰	۱/۶۵	۰/۱۷	۰/۳۹	۰/۰۵
۹	۱/۲۵	۰/۰۹	−۳/۳۵	۰/۲۳	−۲/۴۰	۰/۱۵	−۱/۱۹	۰/۰۸	۰/۲۱	۰/۰۶	۰/۵۹	۰/۰۵
۱۰	۰/۸۷	۰/۰۷	−۳/۲۱	۰/۲۶	−۲/۱۰	۰/۱۷	−۰/۶۷	۰/۰۹	۱/۱۶	۰/۱۲	۰/۴۵	۰/۰۵
۱۱	۱/۰۸	۰/۰۸	−۳/۶۵	۰/۲۷	−۲/۲۵	۰/۱۵	−۰/۷۵	۰/۰۸	۰/۸۳	۰/۰۹	۰/۵۳	۰/۰۵
۱۲	۰/۹۷	۰/۰۷	−۳/۰۳	۰/۲۲	−۱/۴۹	۰/۱۲	−۰/۰۸	۰/۰۷	۲/۰۶	۰/۱۶	۰/۴۹	۰/۰۵
۱۳	۰/۷۱	۰/۰۷	−۴	۰/۳۸	−۲/۵۰	۰/۲۴	−۰/۹۶	۰/۱۲	۰/۶۶	۰/۱۱	۰/۳۸	۰/۰۵
۱۴	۰/۸۶	۰/۰۷	−۳/۵۳	۰/۲۹	−۲/۳۴	۰/۱۹	−۰/۷۹	۰/۰۹	۰/۷۴	۰/۱۰	۰/۴۵	۰/۰۵
۱۵	۱/۱۲	۰/۰۸	−۳/۲۱	۰/۲۳	−۲/۲۹	۰/۱۶	−۰/۹۸	۰/۰۸	۰/۲۱	۰/۰۷	۰/۵۵	۰/۰۵
۱۶	۰/۸۱	۰/۰۷	−۳/۲۲	۰/۲۸	−۲/۲۲	۰/۱۹	−۰/۸۲	۰/۱۰	۰/۶۶	۰/۱۰	۰/۴۳	۰/۰۵
۱۷	۱/۳۰	۰/۰۹	−۳/۰۲	۰/۲۰	−۲/۰۴	۰/۱۲	−۰/۹۲	۰/۰۷	۰/۳۵	۰/۰۶	۰/۶۱	۰/۰۴
۱۸	۱/۳۳	۰/۰۹	−۲/۹۹	۰/۱۹	−۱/۹۶	۰/۱۲	−۰/۷۷	۰/۰۶	۰/۴۵	۰/۰۷	۰/۶۲	۰/۰۴
۱۹	۱/۲۰	۰/۰۹	−۲/۷۲	۰/۱۸	−۱/۷۳	۰/۱۲	−۰/۷۶	۰/۰۷	۰/۰۹	۰/۰۶	۰/۵۸	۰/۰۵
۲۰	۱/۲۱	۰/۰۹	−۲/۸۲	۰/۱۹	−۱/۸۷	۰/۱۲	−۰/۸۱	۰/۰۷	۰/۲۰	۰/۰۶	۰/۵۸	۰/۰۵
۲۱	۰/۸۳	۰/۰۷	−۲/۷۶	۰/۲۳	−۱/۵۸	۰/۱۴	−۰/۱۳	۰/۰۸	۱/۳۹	۰/۱۴	۰/۴۴	۰/۰۵
۲۲	۰/۸۹	۰/۰۸	−۳/۴۸	۰/۲۹	−۲/۵۵	۰/۲۱	−۱/۲۱	۰/۱۱	۰/۲۶	۰/۰۸	۰/۴۶	۰/۰۵
۲۳	۰/۷۲	۰/۰۷	−۲/۳۸	۰/۲۲	−۱/۱۱	۰/۱۳	۰/۳۵	۰/۱۰	۱/۹۷	۰/۱۹	۰/۳۹	۰/۰۵
۲۴	۰/۱۸	۰/۰۶	−۱۰/۲۸	۳/۲۴	−۴/۸۷	۱/۵۵	۰/۳۱	۰/۳۴	۶/۰۷	۱/۹۳	۰/۱۱	۰/۰۶
۲۵	۰/۲۳	۰/۰۶	−۹/۱۳	۲/۳۶	−۶/۲۴	۱/۶۲	−۳/۳۴	۰/۸۹	۰/۴۵	۰/۲۸	۰/۱۴	۰/۰۶
۲۶	۰/۲۷	۰/۰۶	−۸/۸۱	۱/۹۷	−۵/۲۲	۱/۱۷	−۱/۸۸	۰/۴۷	۲/۳۹	۰/۵۷	۰/۱۵	۰/۰۶
۲۷	۱/۰۴	۰/۰۸	−۳/۲۳	۰/۲۳	−۲/۰۸	۰/۱۵	−۰/۶۲	۰/۰۷	۰/۸۰	۰/۰۹	۰/۵۲	۰/۰۴
۲۸	۱/۳۶	۰/۰۹	−۲/۵۲	۰/۱۵	−۱/۶۹	۰/۱۰	−۰/۳۸	۰/۰۶	۰/۹۹	۰/۰۸	۰/۶۲	۰/۰۵
۲۹	۱/۳۲	۰/۰۹	−۲/۶۲	۰/۱۷	−۱/۹۴	۰/۱۲	−۱/۰۵	۰/۰۸	−۰/۰۴	۰/۰۶	۰/۶۱	۰/۰۴
۳۰	۰/۸۵	۰/۰۷	−۲/۸۷	۰/۲۳	−۱/۸۹	۰/۱۶	−۰/۶۳	۰/۰۹	۰/۷۱	۰/۱۰	۰/۴۵	۰/۰۴

همان‌گونه که از جدول ۴ مشاهده می‌شود، اندازه پارامتر شیب (a) برای سؤال‌های ۲۴، ۲۵ و ۲۶، نسبت به اندازه‌های این پارامتر برای سایر سؤال‌ها کوچک است (به ترتیب، ۰/۱۸، ۰/۲۳ و ۰/۲۷). این یافته‌ها نشان داد که این سه سؤال در تک بُعد در نظر گرفته

شده، قابلیت تمیز بین آزمودنی‌ها را ندارند. همچنین، مقایسه مقدار پارامترهای آستانه بین طبقات (b) برای سؤال‌های ۲۴، ۲۵ و ۲۶ با مقادیر این پارمتر برای سایر سؤال‌ها در هر چهار طبقه نشان داد که مقادیر این پارامتر برای این سه سؤال کاملاً غیر عادی می‌باشند. مثلاً مقدار b_1 برای سؤال ۲۴ برابر با ۱۰/۲۸- که خیلی کوچک و مقدار b_4 برای این سؤال برابر با ۶/۰۷ که خیلی بزرگ می‌باشد. این موضوع از شکل ۲ و منحنی‌های پاسخ-طبقه سؤال ۲۴ نیز قابل استنباط است. برای دو سؤال ۲۵ و ۲۶ نیز وضعیتی مشابه برقرار است. این غیر عادی بودن اندازه‌های پارامتر آستانه بین طبقات برای سه سؤال ۲۴، ۲۵ و ۲۶، صرفاً نمی‌تواند ناشی از نامناسب بودن این سؤال‌ها باشد، بلکه همان‌گونه که پیش از این و در بحث ۱ نیز اشاره شد، می‌تواند ناشی از انتخاب غیر صحیح تعداد ابعاد در نظر گرفته شده برای مدل‌سازی داده‌ها باشد. این فرضیّه در قسمت الف (۲) مورد بررسی قرار گرفته است. در پایان این مرحله از بررسی، تحلیل بُعدیّت اکتشافی[1] (EDA)، روی داده‌ها با اعمال مدل تک بُعدی پاسخ مدرّج شده (GRM) سِیم‌جیما و با به‌کارگیری الگوریتم بیشینه‌سازی انتظار (EM) باک و ایتکین، انجام گرفت. نتایج این تحلیل در جدول ۴ و در ستون λ آمده است. مقادیر λ بارهای بُعدی[2] هر یک از سؤال‌ها روی تک بُعد در نظر گرفته شده برای کل سؤال‌ها را نشان می‌دهد. بررسی اندازه مقادیر این بارها نشان داد که اندازه بارهای بُعدی سؤال های ۲۴، ۲۵ و ۲۶ روی بُعد اوّل، ناچیز (به ترتیب برابر ۰/۱۱، ۰/۱۴ و ۰/۱۵) می‌باشد و لازم است که این سؤال‌ها از روی این بُعد حذف شوند.

بحث (۲). با توجّه به شاخص‌های ارائه شده در قسمت‌های الف (۱)، ب (۱)، ج (۱) و نکات مطرح شده در بحث (۱)، و همچنین در نظر گرفتن این موضوع که سؤال‌های ۲۴، ۲۵ و ۲۶ پرسشنامه سنجش توانایی‌های شناختی، یک خرده مقیاس کامل یعنی خرده مقیاس شناخت اجتماعی را تشکیل می‌دادند، در ادامه این فرضیّه مورد بررسی قرار گرفت که عدم برقراری پیش‌فرض‌های استقلال موضعی، یکنوایی، کوچک بودن اندازه‌های مقادیر پارامتر تمیز، غیرعادی بودن (خیلی کوچک یا بزرگ) مقادیر پارامترهای آستانه بین طبقات و کوچک بودن اندازه بارهای بُعدی سه سؤال ۲۴، ۲۵ و ۲۶ روی تک بُعد در نظر گرفته شده، ناشی از در نظر نگرفتن تعداد صحیح ابعاد برای پرسشنامه سنجش توانایی‌های

1. exploratory dimensionality analysis
2. dimension loading

شناختی می‌باشد. تایید این فرضیّه یعنی این‌که وابستگی‌های موضعی مشاهده شده بین این سؤال‌ها از نوع وابستگی موضعی اساسی (ULD) می‌باشد. همان‌گونه که پیش از این و در بحث از مدل‌های وابستگی موضعی مطرح شد، در مدل وابستگی موضعی اساسی (ULD)، فرض می‌شود که یک خصیصه مکنون مجزّا وجود دارد، به گونه‌ای که این خصیصه بین سؤال‌هایی که با یکدیگر وابستگی موضعی دارند، مشترک است، ولی با سایر سؤال‌ها اشتراکی ندارد (چن و تیسن، ۱۹۹۷). به همین دلیل برای سه سؤال ۲۴، ۲۵ و ۲۶، بُعد مجزّایی در نظر گرفته شد. در صورتی که این فرضیّه تایید نشود لازم است این سه سؤال از مقیاس حذف شوند. به همین دلیل در ادامه مدل چند بُعدی پاسخ مدرّج شده (MGRM) موراکی و کارلسون (۱۹۹۵)، با در نظر گرفتن دو بُعد، یک بُعد برای سؤال‌های ۲۴، ۲۵ و ۲۶ (خرده مقیاس شناخت اجتماعی) و یک بُعد برای ۲۷ سؤال دیگر، بر داده‌ها اعمال شد.

الف (۲). بررسی پیش‌فرض استقلال موضعی سؤال‌ها روی دو بُعد

به منظور بررسی مقادیر شاخص وابستگی موضعی X^2 سؤال‌ها روی دو بُعد، مدل دو بُعدی پاسخ مدرّج شده موراکی و کارلسون (۱۹۹۵) با در نظر گرفتن یک بُعد برای سؤال‌های ۲۴، ۲۵ و ۲۶ و یک بُعد برای سایر سؤال‌ها بر داده‌ها اعمال شد. با استفاده از نرم افزار IRTPRO و با به‌کارگیری الگوریتم بیشینه‌سازی انتظار باک و ایتکین (EM) (با ۶۱ تکرار)، فراوانی‌های مورد انتظار برای هر یک از مقوله‌های پاسخ سؤال‌ها برآورد شدند. در ادامه مقادیر شاخص X^2 با به‌کارگیری روش برآورد بیشینه درستنمایی (MLE)، برای سؤال‌ها محاسبه شدند. همچنین، مقادیر شاخص Q_3 ین نیز با استفاده از بسته نرم‌افزاری mirt در نرم افزار R و با اعمال مدل دو بُعدی پاسخ مدرّج شده موراکی و کارلسون (۱۹۹۵) محاسبه شدند. برای این منظور با به‌کارگیری الگوریتم بیشینه‌سازی انتظار (EM) باک و ایتکین (با ۶۱ تکرار)، فراوانی‌های مورد انتظار برای هر یک از مقوله‌های پاسخ سؤال‌ها برآورد شدند و سپس مقادیر شاخص Q_3 با به‌کارگیری روش برآورد بیشینه درستنمایی (MLE) برای سؤال‌ها محاسبه گردید. اطلاعات مربوط به مقادیر شاخص‌های X^2 و Q_3 برای سؤال‌های ۲۴، ۲۵ و ۲۶ و برخی سؤال‌های دیگر، در جدول ۵ که قسمتی از جدول‌های اصلی می‌باشد (جدول های ۱ و ۲ مجدداً و به طور کامل ارائه نشدند)، آورده شده است.

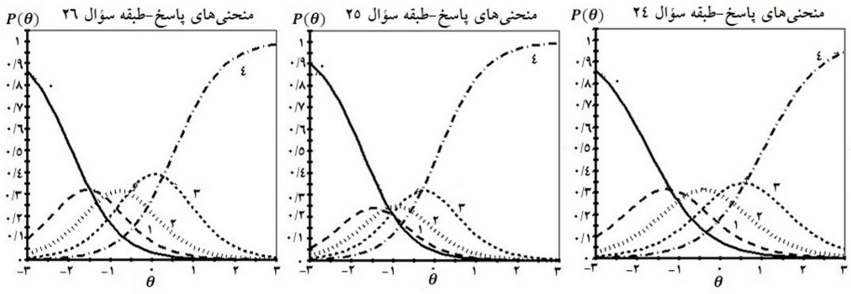
جدول ۵. مقادیر شاخص‌های X^2 و Q_3 سؤال‌های ۲٤، ۲۵ و ۲٦ روی دو بُعد

سؤال	۲۲		۲۳		۲۴		۲۵		۲۶	
	X^2	Q_3	X^2	Q_3	X^2	Q_3	X^2	Q_3	X^2	Q_3
۲۳	۱/۸	۰/۱۳								
۲۴	۲/۴	۰/۰۴	۳/۷	۰/۰۷						
۲۵	−۱/۱	۰/۱۲	۳	۰/۱۱	۲/۱	−۰/۱۲				
۲۶	۰/۷	۰/۰۲	۴/۲	۰/۰۴	۰/۵	−۰/۰۸	۶	−۰/۱۶		
۲۷	۲/۷	۰/۰۸	۳/۹	۰/۰۷	۴/۲	۰/۰۶	۵/۵	۰/۰۷	۵/۷	۰/۱

نتایج به‌دست آمده نشان داد (جدول ۵) مقادیر شاخص‌های وابستگی موضعی (X^2 و Q_3) مدل دو بُعدی، برای سؤال های **۲۴**، **۲۵** و **۲۶** (خرده مقیاس شناخت اجتماعی) و سایر سؤال‌ها در محدوده قابل قبول قرار دارند (کوچک‌تر از ۱۰ برای شاخص X^2 و کوچک‌تر از ۰/۲ برای شاخص Q_3). این مقادیر استقلال موضعی سؤال‌ها روی دو بُعد در نظر گرفته شده را نشان دادند.

ب (۲). بررسی پیش‌فرض یکنوایی سؤال‌ها روی دو بُعد

در ادامه، فرضیّه یکنوایی سؤال‌ها با بررسی منحنی‌های پاسخ-طبقه سؤال‌های پرسشنامه روی دو بُعد مورد بررسی قرار گرفت. با توجّه به این‌که منحنی‌های پاسخ-طبقه برای سایر سؤال‌ها روی بُعد اوّل بدون تغییر بودند، به همین دلیل در شکل ۴ صرفاً منحنی‌های پاسخ-طبقه سؤال‌های **۲۴**، **۲۵** و **۲۶** روی بُعد دوّم، نشان داده شده است.



شکل ٤. منحنی‌های پاسخ-طبقه سؤال‌های ۲٤، ۲۵ و ۲٦ مقیاس سنجش توانایی‌های شناختی در مدل دو بُعدی

مقایسه منحنی‌های پاسخ طبقه این سؤال‌ها در مدل تک بُعدی (شکل ۲) و در مدل دوبُعدی (شکل ۴)، تغییرات موجود در این منحنی‌ها را به خوبی نشان می‌دهد. همان‌گونه که از شکل ۴ مشاهده می‌شود، با افزایش سطح صفت آزمودنی‌ها (در بُعد یا خصیصه

مکنون دوّم) احتمال انتخاب طبقات پایین پاسخ کاهش و احتمال انتخاب طبقات بالا افزایش یافته است. این موضوع نشان داد که پیش فرض یکنوایی برای سه سؤال ۲۴، ۲۵ و ۲۶ روی بُعد دوّم برقرار است.

ج (۲). بررسی پارامترها و بارهای بُعدی مدل سؤال-پاسخ دو بُعدی

به منظور بررسی پارامترهای مدل دو بُعدی و همچنین بارهای بُعدی سؤال‌ها روی هر یک از دو بُعد، مقادیر این پارامترها با اعمال مدل دو بُعدی پاسخ مدرّج شده موراکی و کارلسون (۱۹۹۵) و با به‌کارگیری الگوریتم بیشینه‌سازی انتظار (EM) باک و ایتکین (با استفاده از نرم افزار IRTPRO)، محاسبه شدند. با توجّه به این‌که مقادیر این پارامترها برای سؤال‌هایی که روی بُعد اوّل قرار داشتند، همان مقادیر ذکر شده در جدول۴ بود، مقادیر این پارامترها صرفاً برای سؤال های ۲۴، ۲۵ و ۲۶ در جدول ۶ آورده شده است.

جدول ٦. مقادیر پارمترهای شیب، آستانه بین طبقات و بارهای بُعدی سؤال‌های ٢٤، ٢٥ و ٢٦ در مدل دو بُعدی

سؤال	a	$s.e$	b_1	$s.e$	b_2	$s.e$	b_3	$s.e$	b_4	$s.e$	λ_2	$s.e$
۲۴	۱/۴۵	۰/۱۲	−۱/۷۳	۰/۱۲	−۰/۸۴	۰/۰۷	۰/۰۶	۰/۰۶	۱/۰۵	۰/۰۸	۰/۶۵	۰/۰۵
۲۵	۱/۷۹	۰/۱۶	−۱/۷۴	۰/۱۱	−۱/۲۰	۰/۰۸	−۰/۶۳	۰/۰۶	۰/۱۱	۰/۰۵	۰/۷۲	۰/۰۵
۲۶	۱/۷۳	۰/۱۵	−۱/۹۳	۰/۱۲	−۱/۱۶	۰/۰۸	−۰/۴۰	۰/۰۶	۰/۵۵	۰/۰۶	۰/۷۲	۰/۰۵

مقایسه مقادیر پارامترهای برآورد شده برای این سؤال‌ها در جداول ۴ و ۶ ، تغییرات به‌وجود آمده را به خوبی نشان می‌دهد. همان‌گونه که از جدول ۶ مشاهده می‌شود، اندازه مقادیر پارامتر شیب (a) برای سؤال‌های ۲۴، ۲۵ و ۲۶ در مدل دو بُعدی (به ترتیب، ۱/۴۵، ۱/۷۹ و ۱/۷۳) با اندازه این مقادیر در مدل تک بُعدی (به ترتیب، ۰/۱۸، ۰/۲۳ و ۰/۲۷) تفاوت کاملاً محسوسی دارند. بر خلاف آن‌چه از مقادیر این پارامترها در مدل تک بُعدی استنباط شد، این سؤال‌ها از توانایی بالایی در تمیز بین آزمودنی‌ها روی بُعد دوّم برخوردار می‌باشند. همچنین، مقایسه پارامترهای آستانه بین طبقات (b ها)، برای سؤال‌های ۲۴، ۲۵ و ۲۶ در مدل تک بُعدی (جدول ۴) و مدل دوبُعدی (جدول ۶) نیز انتخاب بُعدی مجزّا برای این سؤال‌ها را تایید کرد. به عنوان مثال، مقدار b_1 برای سؤال ۲۴ در مدل تک بُعدی برابر با ۱۰/۲۸- (غیر طبیعی) به مقدار ۱/۷۳- (طبیعی) در مدل دو بُعدی تغییر یافته و این روند برای سایر سؤال‌ها و مقادیر سایر پارامترهای آستانه بین طبقات (b ها) این سه سؤال نیز برقرار بود.

به علاوه، اجرای تحلیل بُعدیّت تاییدی[1] (CDA)، با در نظر گرفتن دو بُعد؛ یک بُعد برای سؤال‌های ۲۴، ۲۵ و ۲۶ و یک بُعد برای سایر سؤال‌ها، با به‌کارگیری مدل دو بُعدی پاسخ مدرّج شده موراکی و کارلسون (۱۹۹۵)، و استفاده از الگوریتم بیشینه‌سازی انتظار (EM) باک و ایتکین (با استفاده از نرم افزار IRTPRO و پس از ۶۱ تکرار)، نشان داد مقادیر بارهای بُعدی سؤال‌های ۲۴، ۲۵ و ۲۶ که در مدل تک بُعدی به ترتیب برابر با ۰/۱۱، ۰/۱۴ و ۰/۱۵ بودند، در مدل دو بُعدی به ترتیب به ۰/۶۵، ۰/۷۲ و ۰/۷۲ تغییر یافته‌اند. با توجّه به این‌که مقادیر این بارها نسبت به بارهای آمده در جدول ۴ فقط برای سؤال‌های ۲۴، ۲۵ و ۲۶ دچار تغییر شده بودند و برای سایر سؤال‌هایی که روی بُعد اوّل قرار داشتند، بدون تغییر باقی مانده بود، مقادیر این بارها (λ_2) فقط برای این سه سؤال در جدول ۶ ذکر شده است. تغییرات محسوس اندازه بارهای بُعدی نشان داد که در نظر گرفتن بُعد مناسب برای سؤال‌های یک آزمون تا چه میزان می‌تواند برآورد بارهای بُعدی سؤال‌ها را دچار تغییر کند. افزایش بارهای بُعدی سؤال‌ها روی بُعد دوّم، شواهد تایید کننده بیشتری برای انتخاب بُعدی مجزّا برای این سؤال‌ها و دو بُعدی بودن ساختار توانایی‌های شناختی ارائه کرد.

بحث (۳). بررسی پیش‌فرض استقلال موضعی سؤال‌ها روی یک بُعد و دو بُعد در قسمت‌های الف (۱) و الف (۲)؛ بررسی پیش‌فرض یکنوایی سؤال‌ها در قسمت‌های ب (۱) و ب (۲) و بررسی پارامترهای مدل‌های تک بُعدی و دو بُعدی در قسمت‌های ج (۱) و ج (۲)، انجام گرفت. نتایج این بررسی‌ها و بحث‌های انجام شده (بحث ۱ و ۲)، نشان داد که شرط برقراری پیش فرض‌های استقلال موضعی و یکنوایی برای سؤال‌ها و داده‌های حاصل از آن‌ها، در نظر گرفتن دو بُعد برای ساختار داده‌ها می باشد. در ادامه و به منظور مقایسه و بررسی نیکویی برازش مدل‌های تک بُعدی و دو بُعدی با داده‌ها، شاخص‌های برازش این دو مدل نیز مورد بررسی قرار گرفته است.

د. بررسی شاخص‌های برازش مدل‌های تک بُعدی و دو بُعدی با داده‌ها

به منظور ارزیابی نیکویی برازش داده‌ها با مدل‌های تک بُعدی و دو بُعدی، شاخص ریشه میانگین مجذورات خطای تقریب ($RMSEA_2$) مبتنی بر آماره M_2 مایدئو-الیوارس و جو[2]

1. confirmatory dimensionality analysis
2. Maydeu-Olivares & Joe

(۲۰۰۵ و ۲۰۰۶) (نماد ۲ به تفاوت آن با $RMSEA$ اشاره دارد)، برای هر دو مدل تک بُعدی و دو بُعدی توانایی‌های شناختی با استفاده از نرم افزار IRTPRO محاسبه شدند (جدول ۷).

جدول ۷. شاخص‌های برازش مدل تک بُعدی و دو بُعدی پرسشنامه توانایی‌های شناختی

مدل / شاخص	M_2	df	P	$RMSEA_2$
تک بُعدی	۱۰۲۵/۸۸	۳۱۵	۰/۰۰۰۱	۰/۰۵
دو بُعدی	۶۲۴/۲۹	۲۸۶	۰/۰۰۰۱	۰/۰۳

با توجّه به مقادیر جدول ۷، مقدار شاخص $RMSEA_2$ برای مدل دو بُعدی برابر با ۰/۰۳ و کمتر از مقدار این شاخص برای مدل تک بُعدی (۰/۰۵) بود. این یافته نشان داد مدل دو بُعدی برازش نزدیک‌تری با داده‌ها دارد و در نظر گرفتن دو بُعد برای ساختار توانایی‌های شناختی را تایید کرد.

همچنین جهت مقایسه بیشتر دو مدل تک بُعدی و دوبُعدی، از شاخص‌های معیار اطلاعات آکائیک (AIC)، شاخص اصلاح شده معیار اطلاعات آکائیک[1] ($AICc$)، ملاک اطلاعات بیزی (BIC)، ملاک اطلاعات بیزی تنظیم شده با نمونه[2] ($SABIC$) وشاخص منفی دو برابر لگاریتم درست نمایی[3] (-2loglik) استفاده شد. سن و برادشاو[4] (۲۰۱۷)، از طریق چندین مطالعه شبیه‌سازی شده[5] نشان داده‌اند هنگامی که داده‌ها با مدل‌های سؤال-پاسخ (IRT) برازش داده می‌شوند، هر یک از شاخص AIC، BIC و $SABIC$ می‌توانند مدل صحیح را با هر تعداد تکرار برای مدل‌های نظریه سؤال پاسخ تک بُعدی (UIRT) و در ۷۰٪ تکرارها برای مدل‌های نظریه سؤال-پاسخ چند بُعدی (MIRT)، به طور صحیح مشخص کنند. اطلاعات مربوط به این شاخص‌ها در جدول ۸ ارائه شده است.

جدول ۸. شاخص‌های مقایسه مدل تک بُعدی و دو بُعدی مقیاس سنجش توانایی‌های شناختی

مدل / شاخص	AIC	$AICc$	$SABIC$	BIC	$-2logLik$
تک بُعدی	۸۴۹۸۳/۱۹	۸۵۰۳۰/۶۸	۸۵۲۵۷/۹۰	۸۵۷۳۴/۳۳	۸۴۶۸۳/۱۹
دو بُعدی	۸۴۴۸۹/۵۴	۸۴۵۷۹/۳۴	۸۴۸۰۶/۵۶	۸۵۳۸۵/۹۰	۸۴۱۳۱/۵۴
Δ تک بُعدی و دو بُعدی	۴۹۳/۶۵	۴۵۱/۳۴	۴۵۱/۳۴	۳۴۸/۴۳	۵۵۱/۶۵

1. Akaike information criterion corrected
2. sample-size-adjusted BIC
3. log likelihood
4. Sen & Bradshaw
5. simulation study

نتیجه بررسی شاخص‌های برازش دو مدل تک بُعدی و دو بُعدی (جدول ۸)، در مجموع نشان دهنده برازندگی بهتر مدل دو بُعدی نسبت به مدل تک بُعدی بود. زیرا اندازه مقادیر شاخص‌های AIC، $AICc$، $SABIC$ و BIC مدل دو بُعدی کوچک‌تر از اندازه مقادیر این شاخص‌ها در مدل تک بُعدی بودند. همچنین اختلاف بین مقادیر شاخص منفی دو برابر لگاریتم درستنمایی برای مدل تک بُعدی (۸۴۶۸۳/۱۹) و مدل دو بُعدی (۸۴۱۳۱/۵۴) برابر با ۵۵۱/۶۵ به‌دست آمد. از نظر کای و همکاران (۲۰۱۱)، به دلیل این که مدل تک بُعدی درون مدل دوبُعدی قرار دارد (مدل آشیانه‌ای)، اختلاف بین مقادیر شاخص منفی دو برابر لگاریتم درستنمایی می‌تواند به عنوان آماره توزیع χ^2 با یک درجه آزادی تفسیر شود. بنابراین اختلاف ۵۵۱/۶۵ بین دو مدل کاملاً معنی‌دار است ($df = ۱$ و $p < ۰/۰۰۰۱$). این شواهد به طور قوی مؤیّد آن بود که داده‌های به‌دست آمده از مقیاس سنجش توانایی‌های شناختی، برازش بهتری با مدل دو بُعدی دارند. دو بُعدی بودن ساختار داده‌های حاصل از اجرای پرسشنامه سنجش توانایی‌های شناختی، نشان داد که وابستگی‌های موضعی مشاهده شده بین سؤال‌های ۲۴، ۲۵ و ۲۶، از نوع وابستگی اساسی (ULD) می‌باشد. به عبارت دیگر این وابستگی‌ها ناشی از وجود خصیصه مکنون مجزّایی است که بین این سه سؤال مشترک، ولی با سایر سؤال‌ها اشتراکی ندارد (چن و تیسن، ۱۹۹۷).

با توجّه به بحث مطرح شده در قسمت مدل‌های وابستگی موضعی و رابطه ۵، در ادامه مدل دو بُعدی به‌دست آمده برای پرسشنامه توانایی‌های شناختی، به لحاظ همبسته یا غیر همبسته بودن ابعاد آن مورد بررسی قرار گرفت.

ه. بررسی نوع مدل وابستگی موضعی اساسی مشاهده شده بین سؤال‌ها

به منظور بررسی این موضوع که مدل دو بُعدی به‌دست آمده به منظور تبیین وابستگی‌های موضعی اساسی مشاهده شده بین سؤال‌ها، یک مدل دو بُعدی با ابعاد همبسته یا غیرهمبسته می‌باشد، وزن‌های هر یک از سؤال‌ها روی هر یک از دو بُعد، بر اساس رابطه ۵ و با استفاده از نرم افزار IRTPRO محاسبه گردید. وزن‌های هر یک از سؤال‌ها روی بُعد اوّل ($wt_{۱i_{سؤال}}$) و روی بُعد دوّم ($wt_{۲i_{سؤال}}$) در جدول ۹ آمده است.

جدول ۹. وزن سؤال‌های پرسشنامه سنجش توانایی‌های شناختی روی دو بُعد

سؤال	$wt_۱$	$wt_۲$	سؤال	$wt_۱$	$wt_۲$	سؤال	$wt_۱$	$wt_۲$
۱	۰/۵۵	۰/۱۱	۱۱	۰/۵۴	−۰/۰۱	۲۱	۰/۴۶	−۰/۰۹
۲	۰/۴۸	۰/۱۱	۱۲	۰/۵۱	−۰/۰۷	۲۲	۰/۴۷	۰/۰۲

۳	۰/۵۰	۰/۰۳	۱۳	۰/۳۹	۰	۲۳	۰/۴۰	−۰/۰۷
۴	۰/۴۷	۰/۰۷	۱۴	۰/۴۵	۰/۰۴	۲۴	۰/۰۴	۰/۶۴
۵	۰/۶۰	۰/۱۶	۱۵	۰/۵۴	۰/۱۰	۲۵	۰/۰۵	۰/۷۳
۶	۰/۶۲	۰/۱۲	۱۶	۰/۴۴	−۰/۰۷	۲۶	۰/۰۹	۰/۶۹
۷	۰/۵۳	۰/۰۹	۱۷	۰/۶۰	۰/۱۰	۲۷	۰/۵۳	−۰/۰۷
۸	۰/۴۰	−۰/۰۷	۱۸	۰/۶۱	۰/۰۹	۲۸	۰/۶۴	−۰/۰۴
۹	۰/۵۸	۰/۰۹	۱۹	۰/۵۷	۰/۱۰	۲۹	۰/۶۱	۰/۱۱
۱۰	۰/۴۵	۰/۰۳	۲۰	۰/۵۸	۰/۰۵	۳۰	۰/۴۵	۰/۰۵

همان‌گونه که از جدول ۹ مشخص است، وزن سؤال‌های ۲۴، ۲۵ و ۲۶ روی بُعد دوّم قابل توجّه (به ترتیب ۰/۶۴، ۰/۷۳ و ۰/۶۹) و بزرگ‌تر از وزن این سؤال‌ها روی بُعد اوّل (به‌ترتیب ۰/۰۴، ۰/۰۵ و ۰/۰۹) می‌باشد. برعکس وزن سایر سؤال‌ها روی بُعد اوّل بزرگ‌تر از وزن آن‌ها روی بُعد دوّم است (جدول ۹ ملاحظه شود). با توجّه به این‌که وزن‌های به‌دست آمده (wt_i ها)، همگی غیر صفر بودند، بر اساس آن‌چه در مورد مدل‌های وابستگی موضعی بیان شد (رابطه ۵)، مدل وابستگی موضعی مشاهده شده بین سؤال‌ها، از نوع مدل وابستگی موضعی اساسی (ULD) با ابعاد همبسته می‌باشد. به عنوان مثال، در صورتی که سؤال ۲۴ با یک بُعد تحلیل شود (θ)، این بُعد ترکیبی خطی از دو بُعد (θ_1 وθ_2) و به صورت $\theta = (۰/۰۴)\theta_1 + (۰/۶۴)\theta_2$ خواهد بود و نشان می‌دهد هنگامی که مدل به صورت تک بُعدی در نظر گرفته شود (فقط θ_1)، برآورد پارامترهای این سؤال توسط مدل تک بُعدی تا چه میزان می‌تواند غیردقیق باشد. این موضوع را می‌توان با مشاهده مقادیر پارامترهای برآورد شده برای سؤال ۲۴ در جدول‌های ۴ و ۶ نیز مشاهده کرد. با توجّه به این‌که مدل وابستگی مشاهده شده بین سؤال‌ها از نوع مدل وابستگی اساسی با ابعاد همبسته می‌باشد، در ادامه میزان همبستگی بین دو بُعد، محاسبه و اندازه آن برابر با ۰/۱۷ به‌دست آمد. این مقدار کم از همبستگی، نشان دهنده زاویه ۸۰/۲ درجه بین دو بُعد بود و نشان داد که این دو بُعد تقریباً متعامد و مجزّا می‌باشند.

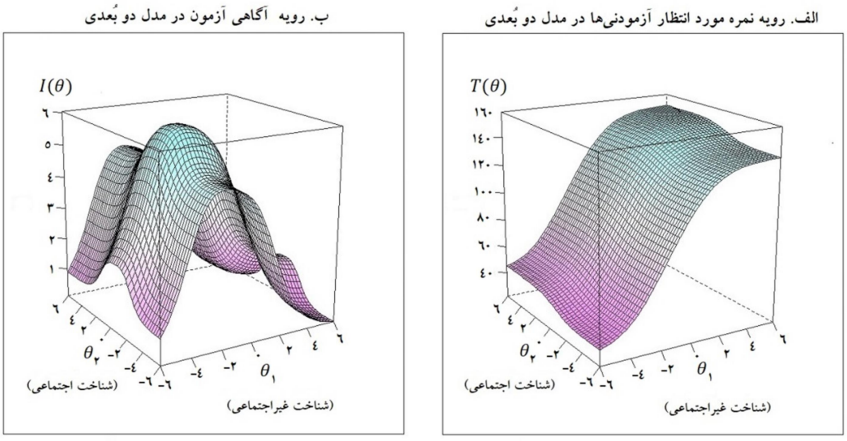
با توجّه به این‌که سؤال‌های ۲۴، ۲۵ و ۲۶ پرسشنامه توانایی‌های شناختی، خرده مقیاس شناخت اجتماعی را تشکیل می‌دادند، بُعد مربوط به این سؤال‌ها به عنوان بُعد شناخت اجتماعی و بُعد مربوط به سایر سؤال‌ها به عنوان بُعد شناخت غیر اجتماعی[1] نام گذاری شد.

1. non-social cognition

در ادامه و به منظور بررسی رابطه بین نمره آزمودنی‌ها با سطح صفت آن‌ها و همچنین بررسی توان آزمون در تمیز بین آزمودنی‌ها بر اساس سطح صفت آن‌ها در این دو بُعد، رویه‌های (سطوح)[1] مربوط به نمره مورد انتظار آزمودنی‌ها و آگاهی آزمون بر اساس سطح صفت آزمودنی‌ها در این دو بُعد مورد بررسی قرار گرفت.

و. بررسی رویه‌های آگاهی آزمون و نمره مورد انتظار آزمودنی‌ها در مدل دو بُعدی

به منظور مشاهده رابطه بین نمره آزمودنی‌ها با سطح صفت آن‌ها و همچنین بررسی توان آزمون در تمیز بین آزمودنی‌ها بر اساس سطح صفت آن‌ها در دو بُعد شناخت اجتماعی و شناخت غیراجتماعی، رویه‌های مربوط به نمره مورد انتظار آزمودنی‌ها و آگاهی آزمون بر اساس سطح صفت آزمودنی‌ها با استفاده از نرم افزار R رسم شدند. این رویه‌ها در شکل ۵ نشان داده شده‌اند.



شکل ۵. رویه‌های نمره مورد انتظار آزمودنی‌ها و آگاهی آزمون بر حسب سطح صفت در مدل دو بُعدی

بررسی رویه نمره نهایی مورد انتظار آزمون (قسمت الف شکل ۵) نشان داد که با افزایش سطح صفت آزمودنی‌ها در هر دو بُعد شناخت غیراجتماعی و شناخت اجتماعی، نمرات آزمودنی‌ها افزایش می‌یابد. نمره‌های کوچک‌تر از ۴۰ برای نقطه (۶- و۶-) معادل پایین‌ترین سطح صفت و نمره‌های بزرگ‌تر از ۱۴۰ برای نقطه (۶ و۶) معادل بالاترین سطح صفت در دو بُعد توانایی‌های شناختی به‌دست آمده است. (آزمون دارای ۳۰ سؤال و نمره آزمودنی‌ها در فاصله ۳۰ تا ۱۵۰ قرار دارد).

1. surfaces

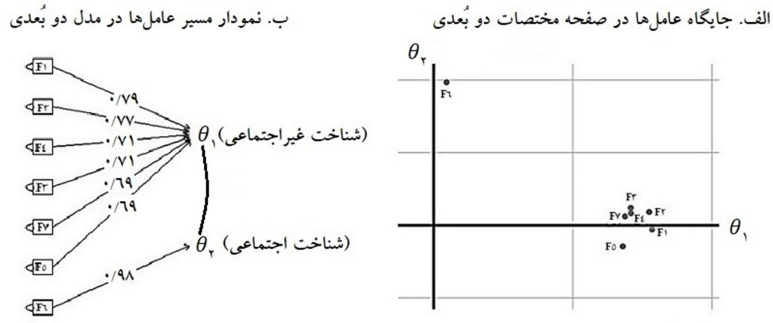
بررسی رویه آگاهی آزمون (قسمت ب شکل ۵) نیز نشان داد که آزمون دارای یک نقطه برافراشتگی (ماکزیمم) در همسایگی نقطه (۰ و۰) می‌باشد و به خوبی قادر است آزمودنی‌ها را به دو گروه تفکیک کند (آزمودنی‌هایی که سطح صفت آن‌ها در دو بُعد، کوچک‌تر یا بزرگ‌تر از (۰ و۰) می‌باشند). همچنین بررسی رویه آگاهی آزمون نشان داد که آزمون قادر است به خوبی آزمودنی‌ها را در سرتاسر دامنه صفت مکنون، به جز برای سطوح بالای صفت (آزمودنی‌هایی که سطح صفت آن‌ها در فاصله (۴ و۴) تا (۶ و۶) قرار دارد) از یکدیگر تفکیک کند.

ز. بررسی بُعدیّت فضای عاملی و نمودار مسیر عامل‌ها

همان‌گونه که پیش از این اشاره شد، پرسشنامه سنجش توانایی‌های شناختی که در این پژوهش مورد استفاده قرار گرفته است، بر اساس تحلیل عاملی کلاسیک صورت گرفته توسط سازنده آن (نجاتی، ۱۳۹۲)، از هفت عامل تشکیل شده است. بر اساس آن‌چه که پیش از این در مورد بُعدیّت فضای عاملی بیان شد، به منظور مشخص کردن تعداد محورهای مختصات یا همان ابعاد فضای عاملی، تحلیل عاملی مرتبه اوّل روی ۳۰ سؤال پرسشنامه سنجش توانایی‌های شناختی با استفاده از روش تحلیل مؤلّفه‌های اصلی و چرخش واریماکس انجام گرفت. (مقدار شاخص کفایت نمونه برداری کایزر- مه‌یر- اولکین[1] (KMO) برابر با ۰/۹۱ به‌دست آمد؛ مقدار شاخص مجذور خی (χ^2) در آزمون کرویّت بارتلت[2] برابر با ۶۳۸۵/۰۲۵ با ۴۳۵ درجه آزادی به‌دست آمد و سطح معنی‌داری کوچک‌تر از ۰/۰۰۰۱ بود؛ واریانس تجمعی تبیین شده توسط ۷ عامل مرتبه اوّل برابر با ۴۷/۵۶٪ و نسبت حاصل از ارزش ویژه عامل اوّل (۶/۴۱) به ارزش ویژه عامل دوّم (۱/۸۹) برابر با ۳/۴۳ بود). در ادامه، تحلیل عاملی مرتبه دوّم روی هفت عامل به‌دست آمده از تحلیل عاملی مرتبه اوّل در این پژوهش، با روش تحلیل مؤلّفه‌های اصلی و با استفاده از منوی R-Factor در نرم افزار SPSS انجام گرفت و دو عامل مرتبه دوّم به‌دست آمد. (مقدار شاخص کفایت نمونه برداری کایزر- مه‌یر- اولکین (KMO) برابر با ۰/۸۶ به‌دست آمد؛

1. Kaiser-Meyer-Olkin
2. Bartlett's test of sphericity

مقدار شاخص مجذور خی (χ^2) در آزمون کرویّت بارتلت برابر با ۱۹۴۷/۰۹۶ با ۲۱ درجه آزادی به‌دست آمد و سطح معنی‌داری آن کوچک‌تر از ۰/۰۰۰۱ بود؛ واریانس تجمعی تبیین شده توسط دو عامل مرتبه دوّم برابر با ۵۹/۶۹٪ بود و نسبت حاصل از ارزش ویژه عامل اوّل (۳/۱۷) به ارزش ویژه عامل دوّم (۱/۰۰۹) برابر با ۳/۱۴ بود). شکل ۶ نتایج تحلیل عاملی مرتبه دوّم بر اساس جایگاه عامل‌ها در صفحه مختصات؛ نمودار مسیر عامل‌ها در مدل دو بُعدی و همچنین وزن استاندارد هر یک از عامل ها در بُعد مربوطه را نشان می‌دهد.



شکل ۶. مختصات و نمودار مسیر عامل‌های پرسشنامه توانایی‌های شناختی بر اساس تحلیل بُعدیّت فضای عاملی

همان‌گونه که از شکل ۶ (قسمت الف) مشخص است، شش عامل از هفت عامل تشکیل دهنده توانایی‌های شناختی روی بُعد اوّل (محور افقی که بُعد شناخت غیراجتماعی را نشان می‌دهد) و تک عامل شناخت اجتماعی که با F_6 مشخص شده است، روی بُعد دوّم (محور قائم که بُعد شناخت اجتماعی را نشان می‌دهد) قرار گرفته است. این یافته‌ها، دو بُعدی بودن فضای عاملی و نتایج به‌دست آمده در قسمت‌های قبل را تایید کرد. قسمت ب شکل ۶ نیز نمودار مسیر عامل‌ها و وزن استاندارد هر یک از عامل‌ها روی بُعد مربوطه را نشان می‌دهد.

در پایان و پس از مشخص شدن مدل داده‌ها، مجدداً شاخص‌های برازش مدل مبتنی بر رویکرد کلاسیک، به کمک نرم‌افزار AMOS، محاسبه شدند. مقایسه شاخص‌های برازش دو مدل (مدل عاملی مرتبه اوّل ارائه شده در شکل ۱ و مدل دو بُعدی به‌دست آمده بر اساس نظریه سؤال-پاسخ چند بُعدی (MIRT) و نمودار مسیر شکل ۶ قسمت ب) نشان داد که شاخص‌های برازش برای مدل دو بُعدی نسبت به مدل عاملی مرتبه اوّل، بهبود یافته‌اند. مقدار شاخص مجذور خی (χ^2) مدل دو بُعدی برابر با ۸۰۴/۹۲ با ۳۹۸ درجه آزادی

($P < ۰/۰۰۰۱$) به‌دست آمد که این مقدار نشان دهنده ۰/۲۰ کاهش در مقدار شاخص مجذور خی هنجار شده (χ^2/df) نسبت به مدل عاملی مرتبه اوّل بود (مقدار شاخص مجذور خی هنجار شده برای مدل عاملی مرتبه اوّل برابر با ۲/۲۲ و برای مدل دو بُعدی برابر با ۲/۰۲ بود). مقدار شاخص *RMSEA* برای مدل دو بُعدی با فاصله اطمینان ۹۰ درصدی (۰/۰۳۳ و ۰/۰۲۷) برابر با ۰/۰۳ به‌دست آمد که کاهشی به مقدار ۰/۰۰۳ را نسبت به مقدار به‌دست آمده برای مدل عاملی مرتبه اوّل نشان داد. همچنین مقدار شاخص *TLI* برای مدل دو بُعدی برابر با ۰/۹۲۶، ۰/۱۶ افزایش و مقدار شاخص *CFI* برابر با ۰/۹۳۲، ۰/۱۲ افزایش را نسبت به مقادیر به‌دست آمده برای مدل عاملی مرتبه اوّل، نشان دادند. به‌علاوه، اختلاف مقدار شاخص *AIC* برای مدل عاملی مرتبه اوّل (۱۰۱۵/۹۸) با مدل دو بُعدی (۹۳۸/۰۹۲)، برابر با ۷۷/۸۸۸ و اختلاف مقدار شاخص *BIC* برای مدل عاملی مرتبه اوّل (۱۴۲۰/۸۱۳) با مدل دو بُعدی (۱۲۷۳/۶۰۲)، برابر با ۱۴۷/۲۱۱ بود. تغییرات مقادیر به‌دست آمده برای شاخص‌های برازش، همگی دلالت بر برازش بهتر مدل دو بُعدی نسبت به مدل عاملی مرتبه اوّل داشتند و صحت نتایج به‌دست آمده از تحلیل مبتنی بر نظریه سؤال-پاسخ چند بُعدی (MIRT) و تحلیل بُعدیّت فضای عاملی را تایید کردند.

بحث و نتیجه‌گیری

پژوهش حاضر با هدف تحلیل بُعدیّت و مدل‌سازی ساختار توانایی‌های شناختی، بر اساس داده‌های به‌دست آمده از اجرای پرسشنامه سنجش توانایی‌های شناختی (نجاتی، ۱۳۹۲)، بر روی دانش‌آموزان پایه نهم تحصیلی شهر اهواز، انجام گرفت. برای نیل به این هدف از روش‌های نوین تحلیل آزمون، مبتنی بر نظریه سؤال-پاسخ چند بُعدی (MIRT)، با تاکید بر بررسی شاخص‌های وابستگی موضعی سؤال‌ها و همچنین روش‌های کلاسیک تحلیل آزمون، مبتنی بر تحلیل عامل اکتشافی، تحلیل عامل تاییدی مرتبه اوّل و مرتبه دوّم و تحلیل بُعدیّت فضای عاملی، استفاده شد. نتایج تحلیل‌های انجام شده کارایی شاخص‌های وابستگی موضعی در تعیین بُعدیّت ساختار آزمون توانایی‌های شناختی را در عمل، نشان دادند. نتایج به‌دست آمده همچنین نشان داد که در نظر گرفتن شاخص‌هایی از قبیل شاخص برازش مدل و یا نسبت حاصل از ارزش ویژه عامل اوّل به ارزش ویژه عامل دوّم،

نمی‌توانند شاخص‌های مناسبی برای تعیین بُعدیّت ساختار یک آزمون باشند. یونسی و همکاران (۱۳۸۹)، از این اصل که «اگر ارزش ویژه عامل اوّل، دو تا سه برابر ارزش ویژه عامل دوّم باشد، می‌توان نتیجه گرفت که آزمون تک بُعدی است و یک عامل غالب، عملکرد افراد را تبیین می‌کند»، استفاده کردند. همچنین فرّخی و همکاران (۱۳۹۳) در پژوهش خود، با بیان این‌که نسبت ارزش ویژه مؤلّفه اوّل به مؤلّفه دوّم برابر با ۴/۴۸ بوده است، همین اصل را برای بررسی تک بُعدی بودن مورد استفاده قرار داده‌اند. نتایج به‌دست آمده از این پژوهش نشان داد که این اصل می‌تواند غیردقیق و نامعتبر باشد. همان‌گونه که پیش از این ذکر شد، بر اساس نتایج حاصل از تحلیل عامل اکتشافی مرتبه اوّل در این پژوهش، نسبت ارزش ویژه عامل اوّل به ارزش ویژه عامل دوّم برابر با ۳/۴۳ به‌دست آمد. اما بررسی شاخص‌های وابستگی موضعی نشان داد که ساختار داده‌ها دو بُعدی می‌باشد و این یافته، اصل مذکور را نقض می‌کند. نتایج این پژوهش همچنین نشان داد در نظر گرفتن شاخص‌های برازش مدل به تنهایی، نمی‌تواند تعیین کننده بُعدیّت ساختار یک آزمون باشد. بر اساس نتایج ارائه شده در مورد مدل شکل ۱ و همچنین جدول ۷، اگر چه شاخص برازش مدل تک بُعدی در محدوده قابل قبول قرار داشتند (مقادیر مساوی یا کوچک‌تر از ۰/۰۵ برای $RMSEA$ و $RMSEA_2$)، اما بررسی شاخص‌های وابستگی موضعی، ناکافی بودن مدل تک بُعدی به منظور مدل‌سازی داده‌ها را نشان دادند. این یافته‌ها نتایج هتی (۱۹۸۴ و ۱۹۸۵) مبنی بر این‌که «برخی از روش‌هایی که بُعدیّت یک آزمون از طریق آن‌ها بررسی می‌شود، از قبیل در نظر گرفتن نسبت ارزش ویژه عامل اوّل به ارزش ویژه عامل دوّم یا در نظر گرفتن شاخص‌های برازش مدل، نسبت به سنجش بُعدیّت آزمون، یا بی‌ارتباط هستند یا نارسایی شدید دارند» (به نقل از امبرتسون و رایس،۲۰۰۰، ترجمه شریفی و همکاران، ۱۳۸۸) را تایید کرد. این موضوع همچنین یافته‌های مینایی (۱۳۹۳) را تایید کرد. از نظر مینایی (۱۳۹۳)، به دلیل ضعف آماره‌های سنتی برازش، عاقلانه‌تر این است که جهت بررسی حالت‌های گوناگون نقض استقلال موضعی، از آزمون‌های آماری که مشخصاً برای این منظور طراحی شده‌اند، استفاده نمود. مینایی (۱۳۹۳) با استناد به مقاله اسمیت (۲۰۰۴)،

بیان کرده است که آماره‌های برازش از حساسیّت لازم برای تشخیص چند بُعدی بودن داده‌ها برخوردار نمی‌باشند و منطقی است که در کنار آماره‌های برازش، از روش تحلیل مؤلّفه‌های اصلی روی پس مانده‌ها نیز استفاده گردد.

بر اساس نتایج حاصل از این پژوهش، پیشنهاد می‌شود در صورتی که برای تعیین بُعدیّت یک آزمون، قصد استفاده از روش تحلیل عامل اکتشافی کلاسیک را داریم، به جای در نظر گرفتن نسبت حاصل از ارزش ویژه عامل اوّل به ارزش ویژه عامل دوّم، از تحلیل بُعدیّت فضای عاملی، یعنی اجرای تحلیل عامل اکتشافی مرتبه دوّم، روی عامل‌های مرتبه اوّل، استفاده شود. همگرایی نتایج حاصل از تحلیل بُعدیّت فضای عاملی، با نتایج حاصل از به‌کارگیری شاخص‌های وابستگی موضعی بر اساس نظریه سؤال پاسخ در این پژوهش، مناسب بودن این روش را نشان داد. نتایج این پژوهش در مجموع نشان داد که استفاده از نظریه سؤال-پاسخ به منظور تعیین بُعدیّت، به دلیل این‌که در این نظریّه اطلاعات بیشتری مورد استفاده قرار می‌گیرد (از قبیل بارهای بُعدی، پارامتر شیب یا تمیز، پارامتر آستانه بین طبقات، یکنوایی و شاخص‌های وابستگی موضعی)، نسبت به روش‌های کلاسیک که عمدتاً همبستگی‌ها را مورد استفاده قرار می‌دهند، برتری دارد.

بر اساس یافته‌های این پژوهش، در صورتی که بُعد شناخت غیراجتماعی را معادل عامل g فرض کنیم (اشنایدر و مک‌گریو[1]، ۲۰۱۸)، این عامل قادر به تبیین واریانس شناخت اجتماعی نمی‌باشد. همان‌گونه که مشاهده کردیم، مقدار همبستگی بین دو بُعد شناخت غیراجتماعی و شناخت اجتماعی در این پژوهش برابر با ۰/۱۷ به‌دست آمد. این مقدار نشان می‌دهد که خصیصه مکنون مبنای عملکرد شناخت غیراجتماعی و خصیصه مکنون مبنای عملکرد شناخت اجتماعی، اشتراک کمی با یکدیگر دارند و این موضوع تک بُعدی بودن ساختار توانایی‌های شناختی، که بر اساس آن، عامل g به عنوان یک توانایی کلی یا عمومی قادر است واریانس سایر عامل‌ها را تبیین کند، نقض می‌کند. این یافته آماری همسو با برخی نظریّه‌ها و یافته‌های پژوهشی می‌باشد. به عنوان مثال، از لحاظ عملکرد افراد در

1. Schneider & McGrew

برخورد با محرک‌های غیراجتماعی و اجتماعی، یافته‌ها حاکی از آن بوده است که برای محرک‌های غیراجتماعی، مفهوم عملکرد، محدود و مشخص است و به سادگی به عمل در می‌آید، اما در تعاملات اجتماعی، مفهوم عملکرد بسیار پیچیده‌تر از موقعیت‌های غیراجتماعی است و در این حالت عملکرد مغز به مراتب دشوارتر بوده و مستلزم اجرای تعداد بی‌شماری از فرآیندهای شناختی می‌باشد (جهانی تابش، ۱۳۹۶). همچنین به لحاظ فرآیندهای شناختی درگیر در پردازش اطلاعات نیز تفاوت‌هایی بین شناخت غیراجتماعی و شناخت اجتماعی گزارش شده است. از نظر کیلستروم[1] (۲۰۱۰)، بین پردازش اطلاعات در شناخت اجتماعی و شناخت غیراجتماعی، به لحاظ عصب‌شناختی، تفاوت‌هایی وجود دارد و این احتمال می‌رود که مناطق خاصی از مغز صرفاً در برخی از جنبه‌های شناخت اجتماعی تخصص داشته باشند. از نظر وی، بررسی این موضوع از دیدگاه علوم اعصاب نشان داده است که جایگاه برخی از جنبه‌های شناخت اجتماعی در مغز، کاملاً منحصربه‌فرد است، به طوری که تاکنون هیچ فعالیت غیراجتماعی برای آن نواحی در نظر گرفته نشده است.

نتایج این پژوهش همسو با تقسیم‌بندی مادریگال[2] (۲۰۰۸) از توانایی‌های شناختی به دو بخش شناخت سرد[3] و شناخت گرم[4] می‌باشد. از نظر مادریگال، توانایی‌های شناختی رابط بین رفتار و ساختار مغز بوده و گستره وسیعی از توانایی‌ها از قبیل برنامه‌ریزی، توجّه، بازداری پاسخ، حل مسئله، انجام همزمان تکالیف و انعطاف‌پذیری شناختی را در بر می‌گیرد. از نظر وی، این موارد صرفاً شامل بخشی از توانایی‌های شناختی می‌باشند که پایه آن‌ها منطق است و این بخش از توانایی‌ها شناخت سرد نامیده می‌شوند. مادریگال، پایه بخش دیگری از توانایی‌های شناختی که به شناخت گرم معروف هستند و در تجربه پاداش و گزند، تنظیم رفتارهای اجتماعی و تصمیم‌گیری در حالت‌های هیجانی نقش دارند را، خواسته‌ها، باورها و هیجان‌ها می‌داند. بر این اساس می‌توانیم بُعد اوّل به دست آمده در این

1. Kihlstrom
2. Madrigal
3. cold cognition
4. hot cognition

پژوهش (بُعد شناخت غیراجتماعی) را بُعد سرد توانایی‌های شناختی و بُعد دوّم (بُعد شناخت اجتماعی) را بُعد گرم توانایی‌های شناختی نام‌گذاری کنیم.

مشاهده و بررسی رویه نمره مورد انتظار آزمودنی‌ها و همچنین رویه آگاهی آزمون، نشان داد که ابعاد به‌دست آمده برای توانایی‌های شناختی، به خوبی قادر هستند آزمودنی‌ها را بر اساس سطح صفت آن‌ها در دو بُعد شناخت غیراجتماعی و شناخت اجتماعی از یکدیگر تفکیک کنند. بیشترین مقدار آگاهی در دو بُعد، برای سطوح میانی صفت به‌دست آمد و این یافته بیان کننده این است که آزمون قادر است به خوبی آزمودنی‌ها را به دو گروه ضعیف و قوی تقسیم کند. با توجّه به این‌که آلووی[1] (۲۰۰۹) شواهد متعددی از نقص توانایی‌های شناختی از قبیل حافظه را در دانش‌آموزان دارای مشکلات یادگیری ارائه کرده است و همچنین با توجّه به این‌که همبستگی‌های بین نمره کل توانایی‌های شناختی با میانگین معدّل تحصیلی و نمرات ریاضی دانش‌آموزان در پایه‌های هفتم و هشتم تحصیلی مثبت و معنی‌دار به‌دست آمد (۰/۴۶ با میانگین نمره ریاضی پایه هفتم و هشتم و ۰/۴۹ با میانگین معدل پایه هفتم و هشتم)، سطح صفت دانش‌آموزان در دو بُعد شناخت غیراجتماعی و شناخت اجتماعی را می‌توان برای بررسی نقص‌های شناختی آن‌ها مورد استفاده قرار داد.

یکی از محدودیّت‌های پژوهش حاضر، تعداد کم سؤال‌هایی بود که شناخت اجتماعی را مورد سنجش قرار می‌دادند. با توجّه به این‌که خرده‌مقیاس شناخت اجتماعی در این پژوهش، تنها شامل سه سؤال بود (سؤال‌های ۲۴، ۲۵ و ۲۶) و با توجّه به یافته‌های این پژوهش مبنی بر این‌که شناخت اجتماعی بُعدی مجزّا از توانایی‌های شناختی را تشکیل می‌دهد، پیشنهاد می‌شود پژوهشگران این حوزه، پرسشنامه‌ای را با تعداد سؤال‌های بیشتر و به منظور بررسی کامل‌تر ساختار این نوع از توانایی‌های شناختی، تهیه کنند و به لحاظ ویژگی‌های روان‌سنجی مورد بررسی قرار دهند.

1. Alloway

منابع

احمدی ده‌قطب‌الدینی، محمد؛ خدایی، ابراهیم؛ فرزاد، ولی الله؛ مقدم زاد، علی و کبیری، مسعود. (۱۳۹۶). کاربرد مدل دوعاملی چند بُعدی نظریّه سؤال-پاسخ جهت تحلیل بُعدیّت و کارکرد افتراقی سؤال، سؤال‌های مبتنی بر متن، *فصلنامه اندازه گیری تربیتی*، سال هفتم، شماره ۲۸، ۱۵۳-۱۲۱.

امبرتسون، سوزان. ای. و رایس استیون، پی. (۲۰۰۰). *نظریّه‌های جدید روان‌سنجی برای روان‌شناسان (IRT)*. ترجمهٔ حسن‌پاشا شریفی، ولی‌الله فرزاد، مجتبی حبیبی عسگرآباد و بلال ایزانلو. (۱۳۸۸). چاپ اوّل، تهران، انتشارات رشد.

ایزانلو، بلال؛ بازرگان، عباس؛ فرزاد، ولی‌الله؛ صادقی، ناهید و کاووسی، امیر. (۱۳۹۳). تفکیک خوشه‌های سؤال از ابعاد متعامد بر اساس هشت روش تعیین بُعد در داده‌های دو ارزشی: مورد آزمون ریاضی رشته ریاضی فیزیک کنکور ۹۲-۹۱. *فصلنامه اندازه گیری تربیتی*، سال پنجم، شماره ۱۸،۲۴۰-۲۰۷.

کرلینجر، فردریک نیکلز. (۱۹۸۶). *مبانی پژوهش در علوم رفتاری*. ترجمهٔ حسن پاشا شریفی و جعفر نجفی زند. (۱۳۸۸). چاپ سوم، تهران: انتشارات آوای نور.

جهانی تابش، عذرا. (۱۳۹۶). شناخت اجتماعی. در خرازی، کمال با همکاری جمعی از استادان، *مقدمه ای بر علوم و فناوری های شناختی و کاربرد های آن*. (۱۳۹۶). چاپ اوّل. تهران: انتشارات سمت. ۱۹۲-۱۶۲.

فرخی، نور علی؛ حرفه دوست، منصور؛ محمدجانی، هیوا؛ جباری، ثریا و رعایی، فرزانه. (۱۳۹۳). ویژگی های روان‌سنجی باورهای کارآمدی در تدریس ریاضی با استفاده از مدل اعتبار پاره ای تعمیم یافته. *فصلنامه اندازه گیری تربیتی*، سال پنجم، شماره ۱۸، ۱۷۸-۱۵۷.

مینایی، اصغر. (۱۳۹۳). کاربرد مدل اندازه‌گیری راش در ارزشیابی ویژگی‌های اندازه‌گیری آزمون مهارت‌های دیداری-حرکتی (TVSM-R). *فصلنامه اندازه گیری تربیتی*، سال پنجم، شماره ۱۸، ۱۱۳-۷۷.

نجاتی، وحید. (۱۳۹۲). پرسش نامهٔ توانایی‌های شناختی: طراحی و بررسی خصوصیات روانسنجی. *فصلنامه تازه‌های علوم شناختی*، سال ۱۵، شماره ۲، ۱۹-۱۱.

یونسی، جلیل؛ دلاور، علی و فلسفی نژاد، محمد رضا. (۱۳۸۹). بررسی ویژگی‌های روان-سنجی سؤال‌های تخصصی آزمون‌های فراگیر رشته روانشناسی دانشگاه پیام نور در سال ۱۳۸۵. *فصلنامه اندازه گیری تربیتی*، دوره ۱، شماره ۲، ۱۶۱-۱۳۱.

Alloway, T. P. (2009). Working memory, but not IQ, predicts subsequent learning in children with learning difficulties. *European Journal of Psychological Assessment*, *25*, 8-92.

Bishop, Y. M. ML., Fienberg, S. E., & Holland, P. W. (1975). *Discrete multivariate analysis*. Cambridge, MA: MIT Press.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: Application of an EM algorithm. *Psychometrika*, *46*(4), 443-459.

Cai, L., Du Toit, S. H. C., & Thissen, D. (2011). *IRTPRO: User guide*. Lincolnwood, IL: Scientific Software International.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies.* New York: Cambridge University Press.

Chalmers, R. P. (2012). Mirt: A multidimensional item response theory package for the R environment, *Journal of Statistical Software*, *48*(6), 1-29.

Chen, W. H., & Thissen, D. (1997). Local dependence indices for item pairs using item response theory. *Journal of Educational and Behavioral Statistics, 22,* 265-289.

Cheng, Y. L. (2016). *The dimensionality of cognitive structure: A MIRT approach and the use of subscores*. **(**Doctoral dissertation), Michigan State University. Retrieved from https://etd.lib.msu.edu/.

Edwards, M. C., Houts, C. R., & Cai, L. (2018). A diagnostic procedure to detect departures from local independence in item response theory models. *Psychological Methods*, *23*(1), 138-149.

Gibbons, R. D., & Cai, L. (2018). Dimensionality analysis. In W. J. Van der Linden (Eds.), *Handbook of item response theory models.*Vol. 3, Published in the Taylor & Francis e-Library.

Hambleton, R. K., Van der Linden, W. J., & Wells, C. S. (2010). IRT models for the analysis of polytomously scored data: Brief and selected history of model building advances. In M. L. Nering & R. Ostini (Eds.), *Handbook of polytomous item response theory models*. Published in the Taylor & Francis e-Library.

Harris, J., Hirsh-Pasek, K., & Newcombe, N. S. (2013). Understanding spatial transformations: Similarities and differences between mental rotation and mental folding. *Cognitive Processing*, *14*(2), 105-115.

Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and space in parietal cortex. *Nature Reviews/Neuroscience, 6,* 435-448.

Houts, C. R., & Edwards, M. C. (2013). The performance of local dependence measures with psychological data. *Applied Psychological Measurement,37,* 541-562.

Jiang, S., Wang, C., & Weiss, D. (2016). The sample size requirements for estimation of item parameters in the multidimensional graded response model. *Frontiers in Psychology , 7*, 109, doi.org/10.3389/fpsyg.2016.00109 .

Kihlstrom, J. F. (2010). Social neuroscience: The footprints of phineas gage. *Social Cognition*, *28*(6), 757-783.

Madrigal, R. (2008). Hot vs. cold cognitions and consumers reactions to sporting event outcomes. *Consumer Psychology*, *18*(4), 304-319.

Maydeu-Olivares, A., & Joe, H. (2005). Limited- and full-information estimation and goodness-of-fit testing in 2n contingency tables. *Journal of the American Statistical Association*, *100*(471), 1009-1020.

Maydeu-Olivares, A., & Joe, H. (2006). Limited information goodness-of-fit testing in multidimensional contingency tables. *Psychometrika*, *71*(4), 713-732.

Maydeu-Olivares, A. (2013). Goodness-of-fit assessment of item response theory models. *Measurement*, *11*, 71-101.

Mumford, M. D., Schultz, R. A., & Osburn, H. K. (2002). Planning in organizations: Performance as a multi-level phenomenon. In J. Francis & F. Yammarino (Eds.), *The many faces of multi-level issues* (pp. 3-65). Emerald Group Publishing Limited.

Muraki, E., & Carlson, J. E. (1995). Full-information factor analysis for polytomous item responses. *Applied Psychological Measurement*, *19*(1),73-90.

Naglieri, J. A., & Otero, T. (2011). Cognitive assessment system: Redefining intelligence from a neuropsychological perspective. In A. S. Davis (Eds.), *Handbook of Pediatric Neuropsychology* (pp. 320-333). Springer Publishing Company.

Otoole, A. J., Jiang, F., Abdi, H., Pénard, N., Dunlop, J. P., & Parent, M. A. (2007). Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. *Journal of Cognitive Neuroscience*, *19*(11), 1735-1752.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory, *Psychological Bulletin*, *116*(2), 220.

Reese, L. M. (1995). The impact of local dependencies on some LSAT outcomes (Statistical Report 95-02), Newtown, PA: Law School Admission Council.

Reckase, M. D. (2009). *Multidimensional item response theory*. New York: Springer.

Reckase, M. D. (2018). *Logistic multidimensional models*. In W. J. Van der Linden (Eds.), *Handbook of Item Response Theory Models,* vol.1. Published in the Taylor & Francis e-Library.

Reckase, M. D., & McKinley, R. L. (1991). The discriminating power of items that measure more than one dimension. *Applied Psychological Measurement*, *15*, 361–373.

Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*(3), 287-297.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika*, Monograph 17, 34.

Samejima, F. (1972). A general model for free-response data.
Psychometrika, Monograph 18, 37.

Sen, S., & Bradshaw, L. (2017). Comparison of relative fit indices for diagnostic model selection. *Applied Psychological Measurement*, *41*(6), 422-438.

Sen, S., Cohen, A. S., & Kim, S. H. (2013). A comparison of algorithms for dimensionality analysis. *New Developments in Quantitative Psychology*, Springer Proceedings in Mathematics & Statistics, 66.

Schneider, W. J., & McGrew, K. S. (2018). The Cattell–Horn–Carroll theory of cognitive abilities. In D. P. Flanagan & E. M., McDonough (Eds.), *Contemporary intellectual assessment theories, tests, and issues* (4th ed.). New York: The Guilford Press.

Shaw, P., Greenstein, D., Lerch, J., Clasen, L., Lenroot, R., Gogtay, N., Evans, A., Rapoport, J., & Giedd, J. (2006). Intellectual ability and cortical development in children and adolescents. *Nature, 440* (7084), 676-679.

Svetina, D., & Levy, R. (2014). A framework for dimensionality assessment for multidimensional item response models. *Educational Assessment*, *19*, 35-57.

Stout, W. F. (1990). A new item response theory modeling approach with applications to unidimensional assessment and ability estimation. *Psychometrika, 55,* 293-326.

Tate, R. (2003). A comparison of selected empirical methods for assessing the structure of responses to test items. *Applied Psychological Measurement*, *27*, 159-203.

Tatsuoka, K. K. (1990). Toward an integration of item response theory and cognitive error diagnosis. In N. Frederiksen, R. Glaser, A. Lesgold, & M. G. Shafto (Eds.). *Diagnostic monitoring of skill and knowledge acquisition* (pp. 453-488). Hillsdale, NJ: Lawrence Erlbaum.

Van der Linden, W. J. (2018). Unidimensional logistic response models. In W. J. Van der Linden (Eds.), *Handbook of Item Response Theory Models*,Vol.3. Published in the Taylor & Francis e-Library.

Yen, W. M. (1993). Scaling performance assessments: Strategies for managing local item dependence. *Journal of Educational Measurement, 30,* 187-213.

Yen, W. M. (1984). Effect of local item dependence on the fit and equating performance of the three-parameter logistic model. *Applied Psychological Measurement,8,* 125-145.