

ساخت و رواسازی آزمون تفکر انتقادی: کاربرد نظریه سؤال-پاسخ

The Construction and Validation of a Test for Measuring Critical Thinking of Students: Using Item-Response Theory (IRT)

Hasan Taghian
Ali Moghadamzadeh
Keyvan Salehi

حسن تقیان*
علی مقدم‌زاده**
کیوان صالحی***

Abstract

This study aimed to build and validate the Critical Thinking Scale using item-response theory. For this purpose, a sample size ($n=7478$) was selected from Qom and Sheikh Mofid university students and the questionnaire was implemented with their cooperation. Three-parameter model was detected more suitable among various IRT models. Data was analyzed through three-parameter logistic model. Exploratory factor analysis was applied to evaluate the theory's assumptions (one-dimensional and local independence). Factor analysis's results confirmed the existence of a dominant factor and local independence. Item Information Function was used to evaluate the propriety of questions. Item Information Function examination showed that some of questions had very little information or the information function was not distributed over ability range. The rest of the item parameters were estimated again after removing inappropriate items. New estimations showed that the range of Discriminant Coefficient was .133 to 1.899 and difficulty coefficient -.990 to 2.780. In addition, coefficient conjecture was .0 to .0790. Eventually, the results showed that the Critical Thinking Scale with 7 components and 29 items has desirable properties to measure the structure.

Keywords: critical thinking, test, item-response theory (IRT), three-parameter logistic model

چکیده

مطالعه حاضر با هدف ساخت و رواسازی آزمون تفکر انتقادی با استفاده از نظریه سؤال-پاسخ انجام گرفت. بدین منظور از بین دانشجویان دانشگاه‌های قم و شیخ مفید، یک نمونه به اندازه $(n=7478)$ انتخاب و پرسشنامه موردنظر با همکاری آن‌ها اجرا شد. از بین مدل‌های مختلف در نظریه سؤال-پاسخ، مدل سه پارامتری مناسب تشخیص داده شد. داده‌ها با کاربرد مدل لوجستیک سه پارامتری تحلیل شد. به منظور بررسی مفروضه‌های این نظریه (تک‌بعدی بودن و استقلال موضعی) از تحلیل عاملی اکتشافی استفاده شد. نتایج تحلیل عاملی وجود یک عامل غالب و استقلال موضعی سوالات را تأیید کرد. به منظور بررسی مناسب بودن سوالات از تابع آگاهی سوالات استفاده شد. بررسی توابع آگاهی سوالات نشان داد برخی از سوالات دارای آگاهی بسیار کمی بوده و یا تابع آگاهی در طول بازه توانایی گسترده نشده است. پس از حذف سوالات نامناسب، پارامترهای سوالات باقی‌مانده دوباره برآورد شد. برآوردهای جدید نشان داد دامنه ضریب تشخیص بین $0/113$ تا $1/899$ و ضریب دشواری بین $-0/990$ تا $2/780$ است. همچنین ضریب حدس نیز در دامنه بین $0/000$ تا $0/790$ قرار دارد. به طور کلی، می‌توان گفت پرسشنامه تفکر انتقادی با ۷ مؤلفه و ۲۹ سؤال، دارای ویژگی‌های مطلوب برای سنجش سازه مذکور می‌باشد.

واژه‌های کلیدی: تفکر انتقادی، آزمون، نظریه سؤال-پاسخ، مدل لوجستیک سه پارامتری

* دانشجوی دکتری سنجش آموزش دانشگاه تهران، عضو هیئت علمی جامعه المصطفی العالمیه
** نویسنده مسئول: عضو هیات علمی دانشگاه تهران
*** عضو هیات علمی دانشگاه تهران

email: amoghadamzadeh@ut.ac.ir

Received: 16 Mar 2016 Accepted: 10 Oct 2016

دریافت: ۹۴/۱۲/۲۶ پذیرش: ۹۵/۷/۱۹

مقدمه

در دنیای کنونی هدف تعلیم و تربیت، آموزش مهارت‌های تشخیص، ترکیب، فرضیه‌سازی، تجسم، توجیه و ارزیابی و چگونگی به نتیجه رسیدن است. از این رو، در سال‌های اخیر مهارت‌های تفکر جایگاه ویژه‌ای در سیاست‌های آموزشی پیدا کرده است و پرورش تفکر انتقادی به‌عنوان یکی از مهم‌ترین اهداف تعلیم و تربیت، مطرح شده است (شریعتمداری، ۱۳۸۵). اگرچه تفکر انتقادی قدمتی به اندازه تاریخ اندیشه بشری دارد و در سال‌های اخیر به‌عنوان مهم‌ترین هدف تعلیم و تربیت به شمار رفته است، اما امروزه نه تنها به‌عنوان یکی از مهم‌ترین اهداف تعلیم و تربیت، بلکه از جمله اصلی‌ترین استانداردهای آموزشی در دنیا محسوب می‌شود (هال، ۲۰۰۸، به نقل از علی‌پور و همکاران، ۱۳۹۲). پیشینه تفکر منطقی به‌صورت اولیه در دو حوزه فلسفه و روان‌شناسی ریشه دارد (لوئیس و اسمیت، ۱۹۹۳). استرنبرگ (۱۹۸۶) به این نکته اشاره دارد که ریشه بخشی از تفکر منطقی نیز در آموزش است. این سه حوزه رویکردهای متفاوتی را در ارتباط با تعریف تفکر منطقی در پیش گرفته‌اند که منعکس‌کننده نگاه آن‌ها به این مقوله است. نوشته‌های سقراط، افلاطون، ارسطو و حتی فیلسوفان معاصر نیز نشان‌دهنده نگاه فلسفی به این موضوع است. این نگاه بر متفکران منطقی فرضی، تمرکز دارد که بیشتر شامل کیفیت و مشخصات این چنین افراد است تا رفتارها یا افعالی که متفکر منطقی، می‌تواند انجام دهد (لوئیس و اسمیت، ۱۹۹۳؛ ثایر-باکون، ۲۰۰۰). استرنبرگ (۱۹۸۶) عنوان کرده است که رویکرد این نوع تفکر، به‌صورت ایدئال بوده و بر آنچه افراد می‌توانند تحت بهترین شرایط انجام دهند، تمرکز دارد؛ بنابراین پائول (۱۹۹۲) در مورد تفکر منطقی در بافت «تعالی تفکر» بحث کرده است. این گرایش به متفکر منطقی ایدئال به‌صورت آشکار در انجمن فلسفه آمریکا منجر به تعریف متفکر منطقی ایدئال به‌عنوان فردی که به‌صورت ذاتی کنجکاو، روشنفکر، انعطاف‌پذیر، بی‌طرف، بصیر، دارای قدرت فهم نقطه نظرات متعدد و متمایل به قضاوت بی‌طرفانه^۱ و متوجه به ابعاد دیگر شده است (فاسیون، ۱۹۹۰). این رویکرد در عرف فلسفی همچنین بر کیفیت یا استانداردهای تفکر تأکید دارد. برای مثال بایلین (۲۰۰۲) تفکر منطقی را به‌صورت تفکر با کیفیت خاص و ضرورتاً تفکر خوب که واجد ملاک‌ها یا استانداردهای خاصی برای دقت و شایستگی باشد، تعریف کرده است. علاوه بر آن، رویکرد فلسفی به‌صورت سنتی بر کاربرد قوانین منطقی مرسوم، تأکید دارد (لوئیس و اسمیت، ۱۹۹۳؛ استرنبرگ، ۱۹۸۶). یک محدودیت مهم این رویکرد تعریف تفکر منطقی، این است که معمولاً با واقعیت‌ها منطبق نیست (استرنبرگ، ۱۹۸۶). با تأکید بر متفکر منطقی و آنچه افراد توان انجام آن را دارند، این رویکرد ممکن است در فهم اینکه چگونه افراد واقعاً فکر می‌کنند کمک زیادی نکند. تعریف تفکر منطقی پدید آمده از فلسفه شامل اجزای زیر است:

^۱. Suspend Judgment

گرایش و مهارت در درگیر شدن در فعالیتی با تشکیک فکری (مک‌پک، ۱۹۹۰)؛ تفکر مستدل و تشکیکی که بر تصمیم در مورد آنچه باید انجام شود یا پذیرفته شود، تمرکز دارد (انیس، ۱۹۸۵)؛ تفکر ماهر و پاسخ‌گو که قضاوت را ساده می‌کند، زیرا: ۱ - بر معیار استوار است، ۲ - خودتصحیح است و ۳ - نسبت به زمینه حساس است؛ قضاوت هدفمند و خودتنظیم که به تفسیر، تحلیل، ارزیابی و نتیجه‌گیری و همچنین به تفسیر مبتنی بر مدرک، روشمند و توجه ادراکی بر آنچه مبنای قضاوت است منجر می‌شود (فاسیون، ۱۹۹۰)؛ تفکر هدف‌مبنا و غایت‌نگر، «تفکر در جهت رسیدن به قضاوت» که در آن تفکر به خودی خود شامل استانداردهای شایستگی و دقت است (باینلین، ۱۹۹۹)؛ قضاوت تفکر مبنا در مورد آنچه باید انجام شود یا آنچه باید پذیرفته شود (فاسیون، ۲۰۰۰).

رویکرد روان‌شناسی شناختی از دو جهت با رویکرد فلسفی متفاوت است. نخست اینکه تمرکز روان‌شناسان شناختی، بر این است که چگونه افراد در واقع فکر می‌کنند در برابر اینکه افراد چگونه می‌توانند یا باید در شرایط ایده‌آل فکر کنند (استرنبرگ، ۱۹۸۶). دوم اینکه علاوه بر تعریف تفکر منطقی با اشاره به ویژگی‌های تفکر منطقی ایده‌آل یا در نظر گرفتن ملاک‌ها یا استانداردهای یک تفکر خوب، تلاش‌های آن‌ها در زمینه روان‌شناسی شناختی، منجر شده است به تعریف متفکر منطقی با انواع فعالیت‌ها یا رفتارهایی که متفکران منطقی می‌توانند انجام دهند؛ بنابراین این رویکرد تعریف تفکر منطقی شامل فهرستی از مهارت‌ها یا فرایندهایی است که توسط متفکران منطقی می‌توانند انجام دهند و مبتنی بر الگوی رفتارگرایی و تحقیقات آزمایشی است (لوئیس و اسمیت، ۱۹۹۳).

محققین فلسفه رویکرد روان‌شناسان شناختی را مورد نقد قرار داده و معتقدند که این رویکرد مجموعه‌ای از دانش و مهارت را در مجموعه‌ای از فرایندهای غیر مرتبط محدود می‌کند (استرنبرگ، ۱۹۸۶). برای مثال باینلین (۲۰۰۲) عنوان می‌کند که توجه به تفکر منطقی به‌عنوان یک سری از مهارت‌ها یا مراحل غیر مرتبط به یک تصور غلط منجر می‌شود و این تصور غلط در نیاز رفتارگرایان به تعریف سازه‌ها به‌صورتی مشاهده‌پذیر ریشه دارد؛ بنابراین به این دلیل که فرایند واقعی تفکر غیرقابل مشاهده است، روان‌شناسان شناختی تمایل دارند بر محصول تفکر (رفتارها یا مهارت‌هایی مانند تحلیل، تفسیر، فرمول‌بندی و ...) تأکید کنند. تعریف تفکر انتقادی که از رویکرد روان‌شناسان شناختی استخراج می‌شود شامل اجزای زیر است:

فرایندهای ذهنی، راهبردها و بازنمایی‌هایی که افراد برای حل مسائل، تصمیم‌گیری و یادگیری مفاهیم جدید بکار می‌برند (استرنبرگ، ۱۹۸۶)، به‌کارگیری مهارت‌های شناختی یا راهبردهایی که احتمال نتایج مطلوب را افزایش می‌دهد (هالپرن، ۱۹۹۸) و توجه به ابعاد یک موضوع، آماده بودن برای مواجهه با شواهدی که نظریه فرد را تأیید نکند، استدلال بی‌غرضانه، استنباط و تفسیر کردن با استفاده از حقایق موجود و ... (ویلینگام، ۲۰۰۷). کسانی که در حوزه آموزش فعالیت می‌کنند نیز در ارتباط با تفکر انتقادی

نظراتی دارند. بلوم و همکاران او از این گروه افراد هستند. طبقه‌بندی آن‌ها برای دانش یکی از پرکاربردترین منابع برای محققین حوزه آموزش است که قصد دارند مهارت‌های مراتب بالاتر تفکر را تدریس کنند یا ارزیابی نمایند. طبقه‌بندی بلوم دارای شش عنصر است که عبارتند از دانش، درک، کاربرد، تحلیل، ترکیب و ارزیابی. چهار سطح بالایی (کاربرد، تحلیل، ترکیب و ارزیابی) به سطوح بالاتری از تفکر نیاز دارد که با تفکر منطقی مرتبط است (پائول، ۱۹۹۲).

با وجود اختلافاتی که بین سه حوزه فوق وجود دارد، اما در تعریف تفکر انتقادی اشتراکاتی نیز وجود دارد. نخستین توافق در مورد توانایی‌هایی است که در ارتباط با تعریف تفکر انتقادی مطرح می‌شود که عبارت‌اند از: ۱- تحلیل استدلال‌ها، ادعاها و شواهد (انیس، ۱۹۸۵؛ فاسیون، ۱۹۹۰؛ هالپرن، ۱۹۹۸؛ پائول، ۱۹۹۲). ۲- نتیجه‌گیری با استفاده از دلایل استقرائی و قیاسی (انیس، ۱۹۸۵؛ فاسیون، ۱۹۹۰؛ پائول، ۱۹۹۲؛ ویلینگهام، ۲۰۰۷). ۳- قضاوت یا ارزیابی (کیس، ۲۰۰۵؛ انیس، ۱۹۸۵؛ فاسیون، ۱۹۹۰؛ لیپمن، ۱۹۸۸؛ تیندال و نولت، ۱۹۹۵). ۴- تصمیم‌گیری یا حل مسائل (انیس، ۱۹۸۵؛ هالپرن، ۱۹۹۸؛ ویلینگهام، ۲۰۰۷). سایر توانایی‌های مرتبط با تفکر انتقادی عبارت‌اند از سؤال و جواب به‌منظور روشن شدن موضوع (انیس، ۱۹۸۵)، تشخیص مفروضات (انیس، ۱۹۸۵؛ پائول، ۱۹۹۲)، تفسیر و تشریح (فاسیون، ۱۹۹۰)، استدلال شفاهی (هالپرن، ۱۹۹۸)، پیش‌بینی (تیندال و نولت، ۱۹۹۵) و توجه به ابعاد مختلف یک موضوع (ویلینگهام، ۲۰۰۷).

بسیاری از محققین معتقدند که علاوه بر مهارت‌ها و توانایی‌ها، تفکر انتقادی شامل گرایش‌ها نیز می‌شود. برخی محققان معتقدند که توانایی تفکر انتقادی با گرایش به این تفکر متفاوت است (انیس، ۱۹۸۵). معمول‌ترین مواردی که در ارتباط با گرایش به تفکر انتقادی مطرح می‌شود، عبارت‌اند از: ۱- عدم تعصب (بالین و همکاران، ۱۹۹۹؛ انیس، ۱۹۸۵؛ فاسیون، ۱۹۹۰؛ هالپرن، ۱۹۹۸). ۲- انصاف و بی‌طرفی (بالین و همکاران، ۱۹۹۹؛ فاسیون، ۱۹۹۰). ۳- تمایل به جستجوی دلایل (بالین و همکاران، ۱۹۹۹؛ انیس، ۱۹۸۵). ۴- کنجکاوی (بالین و همکاران، ۱۹۹۹؛ فاسیون، ۱۹۹۰). ۵- تمایل به مطلع بودن (انیس، ۱۹۸۵؛ فاسیون، ۱۹۹۰). ۶- انعطاف‌پذیری (فاسیون، ۱۹۹۰؛ هالپرن، ۱۹۹۸) و ۷- پذیرفتن و احترام به نظرات دیگران (بالین و همکاران، ۱۹۹۹؛ فاسیون، ۱۹۹۰). کنترل (۲۰۰۵) با جمع‌بندی نظرات مختلف در مورد تفکر انتقادی و با محور قرار دادن نظرات انیس (۱۹۸۷) عنوان می‌کند که تفکر انتقادی دو محور اصلی دارد که عبارت‌اند از: ۱- توانایی تفکر شکاک؛ ۲- توانایی تفکر مستدل؛ و در نهایت هشت محور را به عنوان مهارت‌ها و نگرش‌های مرتبط با تفکر انتقادی مطرح می‌کند که عبارت‌اند از: تشخیص نقطه نظر دیگران؛ ارزیابی شواهد مرتبط با سایر نظرات؛ سنجش منصفانه استدلال‌های مختلف و شواهد مربوطه؛ تشخیص مفروضات نادرست؛ تشخیص راه‌های مورد استفاده برای درست جلوه دادن نظرات اشتباه؛ تفکر در مورد موضوعات به‌صورت ساختاریافته؛ نتیجه‌گیری در

مورد معتبر و قابل تصدیق بودن استدلال‌ها بر اساس شواهد و مفروضات مناسب؛ ارائه نقطه نظرات خود به شیوه ساختاریافته، واضح و مستدل. کوتزل (۲۰۰۵) عنوان کرد هر چند استفاده از این رویکرد بیشتر برای موفقیت در محیط‌های آموزشی مفید است، اما می‌تواند به عنوان مهارت‌هایی از تفکر انتقادی در زندگی روزمره نیز مورد استفاده مؤثر قرار گیرد.

تاکنون ابزارهای متعددی برای سنجش مهارت‌های تفکر انتقادی ساخته شده و مورد استفاده قرار گرفته است که در ادامه برخی از معروف‌ترین و پرکاربردترین ابزارها معرفی می‌شوند: ۱- آزمون مهارت‌های تفکر انتقادی کالیفرنیا: این آزمون دارای ۳۴ گویه چندگزینه‌ای است که به سنجش مهارت‌های تحلیل، استنباط و ارزشیابی می‌پردازد. مهارت‌های تفسیری مورد سنجش در این ابزار عبارت‌اند از: طبقه‌بندی، رمزگشایی جملات، روشنگری معنا، بررسی ایده‌ها و تحلیل ایده‌ها. مهارت‌های استنباطی عبارت‌اند از: جستجو برای شواهد، گمانه‌زنی در مورد جایگزین‌ها و استخراج نتایج. مهارت‌های ارزشیابی عبارت‌اند از: ارزشیابی ادعا، ارزشیابی بحث‌ها، بیان نتایج، توجیه رویه‌ها و ارائه استدلال و استدلال استقرایی شامل نتیجه‌گیری از بحث به دنبال رویارویی با حقایق مربوط به پیش‌فرض‌ها (معطری و همکاران، ۱۳۸۱). خلیلی و سلیمانی (۱۳۸۲) روایی و پایایی این آزمون را در ایران مورد بررسی قرار دادند که پایایی آن با استفاده از فرمول ۲۰ کودر ریچاردسون ۰/۶۲ بدست آمد. ۲- فهرست عوامل زمینه‌ساز یا گرایش‌های تفکر انتقادی کالیفرنیا: این پرسشنامه دارای ۷۵ گویه با مقیاس لیکرت ۶ گزینه‌ای است و عوامل زمینه‌ساز تفکر انتقادی از جمله جستجوی حقیقت، ذهن باز، تحلیل‌گر بودن، روش‌مندی، اعتماد به نفس، کنجکاوی و کمال تفکر انتقادی تنظیم شده است (زوری و همکاران، ۲۰۱۰). ۳- آزمون تفکر انتقادی نوشتنی انیس و ویر: در این آزمون برهان‌سازی آزمون شوندگان از طریق ارزیابی متن نوشته شده مورد بررسی قرار می‌گیرد. این آزمون برای اندازه‌گیری فکر باز و بازاندیشی دانشجویان مورد استفاده قرار می‌گیرد (ورل و همکاران، ۲۰۰۷). ۴- آزمون تفکر انتقادی واتسون و گلیرز: این آزمون یکی از آزمون‌هایی است که به صورت گسترده استاندارد شده و قادر به اندازه‌گیری اجزاء عقلانی و خلاق تفکر انتقادی است. دو فرم از این آزمون (فرم الف و ب) وجود دارد. در این آزمون ۵ مهارت تفسیر، شناخت فرضیه‌ها، قیاس، تفسیر و ارزیابی استدلال مورد سنجش قرار می‌گیرد (ورل و همکاران، ۲۰۰۷). پرسشنامه‌های فوق به فارسی ترجمه شده است و روایی و پایایی آن‌ها مورد بررسی قرار گرفته است، اما هنوز پرسشنامه‌ای که به صورت مستقل در ایران شکل گرفته باشد ساخته نشده است؛ بنابراین در این پژوهش یک آزمون با استفاده از رویکرد کوتزل (۲۰۰۵) ساخته شده و علاوه بر نظریه کلاسیک، از نظریه سؤال-پاسخ نیز برای بررسی ویژگی‌های سؤال و آزمودنی استفاده شده است. در پرسشنامه‌های کنونی سنجش تفکر انتقادی، از نظریه سؤال-پاسخ استفاده نشده است. هدف این رویکرد این است که علاوه بر برآورد دقیق پارامترهای آزمون و سؤال، نارسایی‌های نظریه

کلاسیک آزمون، از جمله وابستگی پارامتر سؤال به نمونه خاص را از میان بردارد (هومن، ۱۳۷۱)؛ بنابراین مطالعه حاضر به منظور طراحی آزمون تفکر انتقادی برای دانشجویان و بررسی ویژگی‌های روان‌سنجی و هنجاریابی آن انجام شد. بدین منظور سوال‌های زیر طرح شده‌است:

- بر پایه نظریه کلاسیک، ابزار ساخته شده برای اندازه‌گیری تفکر انتقادی دانشجویان، از چه ویژگی‌های روان‌سنجی برخوردار است؟
- بر پایه نظریه سؤال - پاسخ، ابزار ساخته شده برای اندازه‌گیری تفکر انتقادی دانشجویان از چه ویژگی‌های روان‌سنجی برخوردار است؟
- سؤال‌های ابزار ساخته شده با کدامیک از مدل‌های نظریه سؤال - پاسخ، برازندگی مناسب‌تری دارد؟
- پارامترهای حدس، دشواری و تشخیص سوال‌های ابزار ساخته شده از چه وضعیتی برخوردار است؟

ساخت آزمون با استفاده از نظریه سؤال - پاسخ

پیدایش نظریه سؤال - پاسخ به اوایل ۱۹۴۰ مربوط است، اما اقبال عمومی به این نظریه تا سال ۱۹۷۰ به تأخیر افتاد. همان‌گونه که از نام این نظریه پیداست، مدل‌های پاسخ سؤال، رفتار آزمودنی‌ها را در سطح سؤال (و نه در سطح آزمون) مورد نظر قرار می‌دهد. ساخت مدلی در سطح سؤال، برای ساخت آزمون، مطالعه سنجش افتراقی، سنجش انطباقی کامپیوتری و گزارش نمرات، انعطاف بیشتری را به دست می‌دهد. مدل‌های اولیه پاسخ سؤال، برای کار با پاسخ‌های دوقسمتی (مانند صحیح-غلط، موافق-مخالف یا بله-خیر) ساخته شد، اما امروزه مدلهایی برای کارکردن با همه انواع داده‌های آموزشی و روان‌شناختی به وجود آمده است (لیندن و همبلتون، ۱۹۹۷).

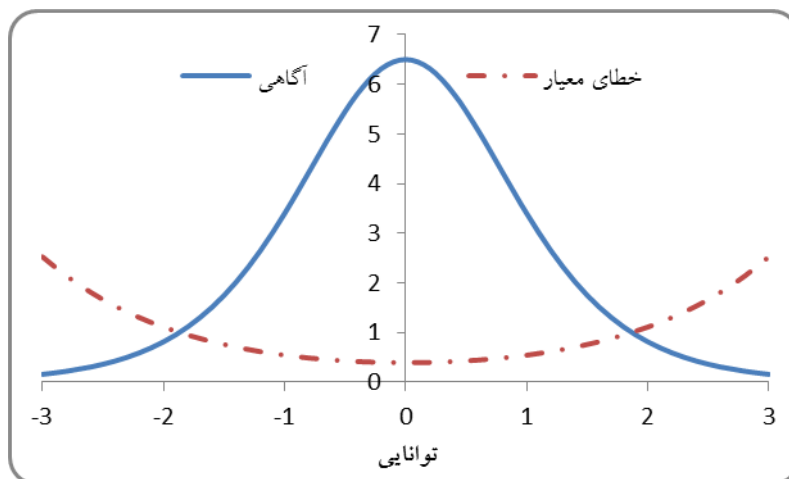
دو مفروضه اساسی مدل‌های سؤال - پاسخ، تک‌بعدی بودن و استقلال موضعی است (وندربلین، ۲۰۱۶). مفروضه تک‌بعدی بودن به این معنی است که مجموعه‌ای از سؤال‌ها تنها یک صفت مکنون (θ) را اندازه‌گیری می‌کند. استقلال موضعی مربوط به این مفروضه است که بین پاسخ‌های آزمودنی‌ها به زوج‌های سؤال‌ها یک آزمون بعد از حذف صفت مورد اندازه‌گیری، همبستگی آماری وجود نداشته باشد. در واقع دو مفروضه مذکور به طرق مختلف یک مطلب را در مورد داده‌ها بیان می‌کنند. سومین مفروضه در مورد مدل‌بندی رابطه بین صفت مورد اندازه‌گیری و پاسخ‌های سؤال‌ها است؛ بنابراین مدل‌های مختلف، مفروضات متفاوتی را در مورد این ارتباط در نظر می‌گیرند.

مدل اجایو نرمال اولین مدل پاسخ سؤال است که برای اندازه‌گیری صفت‌های مکنون آموزشی و یا روان‌شناختی مورد استفاده قرار گرفته است. مدل اجایو نرمال توسط لورد و نویک (۱۹۶۸) مورد تصحیح قرار گرفت. در این مدل، منحنی مشخصه یک سؤال از تابع احتمال تجمعی یک توزیع نرمال استخراج می‌شود. جورج راش با رویکرد متفاوتی به این نظریه، یک تابع لوجستیک را برای استخراج تابع مشخصه

به جای تابع اجایو نرمال ارائه داد (راش، ۱۹۶۰). مدل او مدل اجایو نرمال را ساده کرده و پیچیدگی‌های محاسباتی را تقلیل داد. در این مدل برای همه سؤالات، ضریب تشخیص به صورت ثابت (برابر با یک) در نظر گرفته شده و احتمال پاسخ صحیح به سؤال توسط یک آزمودنی به توانایی آزمودنی و دشواری سؤال وابسته است. در مدل لوجستیک دو پارامتری به جای اینکه برای همه سؤالات ضریب تشخیص برابر فرض شود، هر سؤال دارای ضریب تشخیص خاص خود است. مدل سه پارامتری اجازه می‌دهد که منحنی مشخصه دارای یک مجانب غیر صفر باشد. این مدل زمانی مفید است که سؤالات به گونه‌ای باشد که آزمودنی‌های با توانایی پایین بتوانند به سؤال با استفاده از شانس، پاسخ صحیح بدهند (مثلاً در سؤالات چندگزینه‌ای). همان‌گونه که مشاهده می‌شود در نظریه سؤال - پاسخ، سؤالات با ضریب تشخیص بیشتر و پایایی^۱ بزرگ‌تر، دارای وزن بیشتری می‌باشد و بنابراین، نمرات می‌توانند پایاتر از نمرات روش کلاسیک اندازه‌گیری باشند (دیمارس، ۲۰۱۰). دانستن دشواری سؤالات می‌تواند در ساخت آزمون‌هایی با هدف دستیابی به سطوح صفت جامعه هدف مفید باشد. ضریب تشخیص برای انتخاب سؤالاتی که به خوبی بین آزمودنی‌ها با سطوح متفاوت مهارت یا صفت تمایز قائل شود، مفید است. همچنین هر دو شاخص دشواری و تشخیص برای محاسبه خطای استاندارد اندازه‌گیری یا پایایی نمرات کاربرد دارد. در نظریه کلاسیک اندازه‌گیری، پایایی^۲ به صورت نسبت واریانس واقعی به واریانس مشاهده شده و یا مربع همبستگی بین نمرات واقعی و نمرات مشاهده شده تعریف می‌شود. برای خطای استاندارد اندازه‌گیری نیز تنها یک برآورد وجود خواهد داشت. در نظریه سؤال - پاسخ تابع آگاهی برای محاسبه خطای استاندارد اندازه‌گیری و پایایی مورد استفاده قرار می‌گیرد؛ بنابراین آگاهی آزمون با توجه به سطح مهارت تغییر می‌کند. خطای استاندارد اندازه‌گیری برابر است با معکوس ریشه دوم آگاهی؛ بنابراین مقادیر بزرگ آگاهی، به مقادیر کوچک‌تر خطای استاندارد و مقادیر بزرگ‌تر پایایی منجر می‌شود (نمودار ۱).

^۱. Reliable

^۲. Reliability



(هر جا توانایی بیشتر است، خطای استاندارد کمتر است)

نمودار ۱- رابطه بین آگاهی آزمون و خطای استاندارد برآورد توانایی

در این نظریه، آگاهی آزمون حاصل جمع آگاهی سؤالات است. منحنی مشخصه آزمون^۱ نیز حاصل جمع منحنی‌های سؤالات است. از آنجا که، برای هر سؤال مقدار منحنی بین کمترین مجانب (برای مدل‌های یک و دو پارامتری برابر با صفر است) تا یک می‌باشد، این مقدار برای آزمون بین جمع کمترین مجانب‌های سؤالات تا تعداد سؤالات می‌باشد. خطای استاندارد برآورد توانایی، انحراف معیار توزیع مجانباً نرمال برآورد بیشینه درست‌نمایی برای یک مقدار واقعی توانایی مفروض است. وقتی آزمون طولانی است توزیع نرمال است. با این همه برای بیشتر مقاصد سنجش، حتی با آزمون‌هایی به کوتاهی ۱۰ تا ۲۰ سؤال، نیز تقریب‌های نرمال رضایت‌بخش به دست می‌آید (سامیجیما، ۱۹۷۷؛ به نقل از فلسفی‌نژاد، ۱۳۸۹). به‌طور کلی، مقدار خطای استاندارد اندازه‌گیری بستگی به (الف) تعداد سؤالات آزمون (خطای استاندارد کوچک‌تر متعلق به آزمون‌های طولانی‌تر است)؛ (ب) کیفیت سؤالات آزمون (به‌طور کلی، خطاهای استاندارد کوچک‌تر متعلق به سؤالاتی است که قدرت تمیز بالایی دارند و پاسخ صحیح آن‌ها را نمی‌توان حدس زد)؛ (ج) جور بودن دشواری سؤال با توانایی آزمودنی (برخلاف سؤالات خیلی سخت یا خیلی آسان، خطاهای استاندارد اندازه‌گیری کوچک‌تر مربوط به آزمون‌هایی است که دشواری سؤالات آن‌ها تقریباً معادل توانایی آزمودنی‌ها است). اندازه خطای استاندارد به‌سرعت پایدار می‌شود،

^۱. Test Characteristic Curve

ساخت و روانسازی آزمون تفکر انتقادی: کاربرد نظریه سؤال - پاسخ

به طوری که افزایش مقدار آگاهی به بیش از یک مقدار (مثلاً) ۲۵ تنها اثر کمی در مقدار خطاهای برآورد توانایی دارد (فلسفی نژاد، ۱۳۸۹).

روش

پژوهش حاضر به لحاظ نحوه گردآوری داده‌ها، از نوع مطالعات توصیفی به شمار می‌رود. جامعه آماری این مطالعه را تمامی دانشجویان مقطع کارشناسی و کارشناسی ارشد دختر و پسر دانشگاه‌های شهر قم تشکیل دادند. نمونه‌ای به حجم ۴۷۸ نفر از دانشجویان دختر و پسر مقاطع کارشناسی و کارشناسی ارشد دانشگاه‌های قم و شیخ مفید به روش نمونه‌گیری تصادفی مورد بررسی قرار گرفت. ۲۹۶ نفر (۶۱/۹ درصد) آزمودنی‌ها زن و ۲۰۹ نفر (۴۳/۷ درصد) مرد بودند. ۳۱۳ نفر (۶۵/۵ درصد) در گروه علوم انسانی، ۵۷ نفر (۱۱/۹ درصد) در گروه علوم پایه و ۱۰۸ نفر (۲۲/۶ درصد) در گروه فنی و مهندسی مشغول به تحصیل بوده‌اند. ۳۶۲ نفر (۷۵/۷ درصد) در مقطع کارشناسی و ۱۱۶ نفر (۲۴/۳ درصد) در مقطع کارشناسی ارشد مشغول تحصیل بوده‌اند.

برای ساخت ابراز اندازه‌گیری مهارت‌های تفکر انتقادی دانشجویان، ابعاد تفکر انتقادی کوترل (۲۰۰۷) مورد استفاده قرار گرفت. از نظر کوترل (۲۰۰۷)، دانشجویان برای موفقیت تحصیلی به هشت بُعد تفکر انتقادی نیاز دارند. به این ترتیب، برای تشخیص نقطه‌نظر دیگران، ۴ سؤال؛ برای ارزیابی شواهد مرتبط با سایر نظرات، ۹ سؤال؛ برای سنجش منصفانه استدلال‌های مختلف و شواهد مربوطه، ۴ سؤال؛ برای تشخیص مفروضات نادرست، ۵ سؤال؛ برای تفکر در مورد موضوعات به صورت ساختاریافته، ۷ سؤال؛ برای نتیجه‌گیری در مورد معتبر و قابل تصدیق بودن استدلال‌ها بر اساس شواهد و مفروضات مناسب ۳ سؤال و برای ارائه نقطه‌نظرات خود به شیوه ساختاریافته، واضح و مستدل، ۳ سؤال در نظر گرفته شد. پس از تدوین سؤال‌های پرسشنامه توسط ۴ نفر از استادان رشته فلسفه و علوم تربیتی بررسی و اصلاح شد و فرم نهایی بر روی نمونه مورد بررسی اجرا گردید.

یافته‌ها

شاخص‌های کلاسیک

نتایج بررسی همگنی درونی کل پرسشنامه و مؤلفه‌های آن نشان داد آلفای کرونباخ پرسشنامه اولیه ۰/۵۶ و مؤلفه‌های آن به ترتیب برای تشخیص نقطه‌نظر دیگران، ۰/۸۱؛ برای ارزیابی شواهد مرتبط با سایر نظرات، ۰/۸۵؛ برای سنجش منصفانه استدلال‌های مختلف و شواهد مربوطه ۰/۴۷؛ برای تشخیص مفروضات نادرست، ۰/۳۱؛ تفکر در مورد موضوعات به صورت ساختاریافته ۰/۹؛ برای نتیجه‌گیری در مورد معتبر و قابل تصدیق بودن استدلال‌ها بر اساس شواهد و مفروضات مناسب، ۰/۷۴ و برای ارائه نقطه نظرات خود به شیوه ساختاریافته، واضح و مستدل ۰/۷۸ بود که پس از حذف سؤالات نامناسب و فرم نهایی به ترتیب ۰/۷۱، ۰/۸۵، ۰/۸۵، ۰/۶۴، ۰/۴۶، ۰/۹۰، ۰/۷۹ و ۰/۸۶ به دست آمد؛ بنابراین می‌توان گفت برای دو مؤلفه سنجش منصفانه استدلال‌های مختلف و شواهد مربوطه و تشخیص مفروضات نادرست کمتر از ۰/۷ است، اما وضعیت سایر مؤلفه‌ها تقریباً مطلوب است.

شاخص دشواری^۱ و شاخص تشخیص^۲

جدول ۱ شاخص‌های کلاسیک اندازه‌گیری شامل دشواری سؤالات و ضریب تشخیص را نشان می‌دهد. همان‌گونه که مشاهده می‌شود، دشواری سؤالات در فاصله بین ۰/۱۸ تا ۰/۹۰ قرار دارد. از آنجا که مقدار مناسب برای دشواری سؤالات بین ۰/۳ تا ۰/۷ است؛ می‌توان گفت دشواری اکثر سؤالات در فاصله مناسب قرار دارد. همچنین ضریب تشخیص سؤالات در بازه بین ۰/۱۱ تا ۰/۵۳ قرار دارد که بازه‌ای بین ضریب تشخیص کم تا متوسط را شامل می‌شود.

شاخص‌های نظریه سؤال - پاسخ

بررسی مفروضه‌های نظریه به کارگیری نظریه سؤال - پاسخ مستلزم برقراری برخی مفروضات است. تک بُعدی بودن^۳، از مهم‌ترین مفروضه‌های نظریه سؤال پاسخ است منظور این است که سازه مورد اندازه‌گیری به واقع تک‌بعدی باشد. به عبارت دیگر، کوواریانس بین سؤالات باید توسط یک بُعد، تبیین شود. به منظور بررسی این مفروضه می‌توان با استفاده از تحلیل عاملی اکتشافی، اولین مقدار ویژه و بار عاملی سؤالات بر اولین عامل را مورد بررسی قرار داد (کتل، ۱۹۶۶ و لوهلین، ۱۹۸۷).

^۱. difficulty Index

^۲. Discriminate Index

^۳. Unidimensional

ساخت و رواسازی آزمون تفکر انتقادی: کاربرد نظریه سؤال - پاسخ

جدول ۱- شاخص‌های کلاسیک سؤالات آزمون

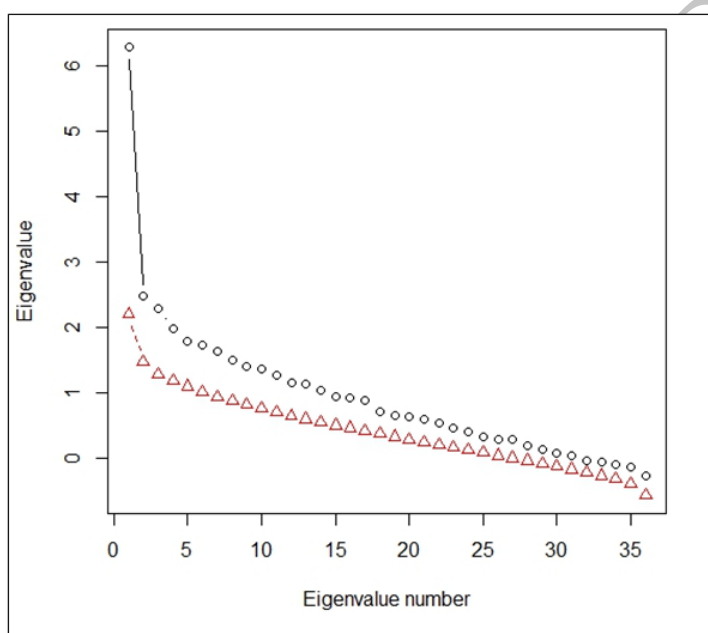
شماره سؤال	ضریب تشخیص	دشواری سؤال	شماره سؤال	ضریب تشخیص	دشواری سؤال
۱	۰/۲۸	۰/۵۰	۱۹	۰/۴۹	۰/۸۳
۲	۰/۳۱	۰/۵۳	۲۰	۰/۴۶	۰/۸۶
۳	۰/۳۵	۰/۶۱	۲۱	۰/۴۸	۰/۸۲
۴	۰/۲۱	۰/۳۱	۲۲	۰/۵۲	۰/۶۹
۵	۰/۳۲	۰/۳۳	۲۳	۰/۱۹	۰/۳۵
۶	۰/۲۷	۰/۵۱	۲۴	۰/۲۴	۰/۲۹
۷	۰/۴۰	۰/۳۳	۲۵	۰/۳۵	۰/۳۶
۸	۰/۱۲	۰/۴۸	۲۶	۰/۲۵	۰/۵۳
۹	۰/۳۴	۰/۶۰	۲۷	۰/۲۰	۰/۲۶
۱۰	۰/۲۵	۰/۴۹	۲۸	۰/۲۱	۰/۴۱
۱۱	۰/۳۹	۰/۳۹	۲۹	۰/۳۴	۰/۵۲
۱۲	۰/۳۳	۰/۷۱	۳۰	۰/۳۵	۰/۵۰
۱۳	۰/۲۱	۰/۴۶	۳۱	۰/۲۰	۰/۵۱
۱۴	۰/۱۳	۰/۳۰	۳۲	۰/۲۱	۰/۳۷
۱۵	۰/۳۲	۰/۵۶	۳۳	۰/۱۹	۰/۲۹
۱۶	۰/۵۱	۰/۷۵	۳۴	۰/۲۶	۰/۴۸
۱۷	۰/۵۳	۰/۸۱	۳۵	۰/۱۱	۰/۱۸
۱۸	۰/۴۷	۰/۹۰	۳۶	۰/۲۶	۰/۴۶

دومین مفروضه در مدل‌های پاسخ سؤال این است که سؤالات دارای استقلال موضعی^۱ باشند. این مفروضه با مفروضه تک‌بعدی بودن در ارتباط است و به این معنی است که با حذف اثر سازه زیربنایی سؤالات، کوواریانس سیستماتیک دیگری بین سؤالات وجود ندارد. وقتی که مفروضه تک‌بعدی بودن برقرار است، استقلال موضعی نیز حاصل می‌شود؛ از این حیث دو مفهوم با هم معادل اند (لرد و ناویک، ۱۹۶۸؛ به نقل از فلسفی‌نژاد، ۱۳۸۹). برای بررسی مفروضه تک‌بعدی بودن از تحلیل عاملی اکتشافی برای داده‌های دوارزشی با استفاده از ماتریس همبستگی چندحالتی^۲ استفاده شد. نتایج نشان داد یک

^۱. Local Independence

^۲. Polychoric

ارزش ویژه غالب (۶/۳) نسبت به سایر ارزش‌های ویژه ۲/۵، ۲/۳ و ... وجود دارد. همچنین نمودار اسکری (نمودار ۲) ارزش‌های ویژه ماتریس همبستگی بین سؤالات - که از داده‌های آزمون بدست آمده است - را با نمودار ارزش ویژه ماتریس همبستگی بین سؤالات داده‌های تصادفی مقایسه می‌کند. دو نمودار با هم مشابه هستند؛ به جز اولین ارزش ویژه مربوط به داده‌های آزمون؛ این ارزش ویژه به طور قابل ملاحظه‌ای از هم‌تای تصادفی‌اش بزرگ‌تر است؛ بنابراین می‌توان فرض تک‌بعدی بودن را پذیرفت.



نمودار ۲- مقایسه نمودار اسکری مربوط به ماتریس همبستگی سؤالات آزمون و ماتریس همبستگی داده‌های تصادفی برای بررسی مفروضه تک‌بعدی بودن

انتخاب مدل

به منظور برآورد پارامترها، دو مدل در نظر گرفته شد. مدل اول یک مدل دوپارامتری است که در آن فرض می‌شود پارامترهای شیب و تشخیص به صورت آزادانه برآورد می‌شوند، اما پارامتر حدس مقداری ثابت (برابر با صفر) است. مدل دوم، مدل سه پارامتری است که در آن پارامتر حدس نیز برآورد می‌شود. نتایج بررسی ۲- برابر لگاریتم درست‌نمایی مدل‌ها نشان داد، تفاوت بین دو مدل ۸۵/۳ با درجه آزادی ۳۶ معنی‌دار است. بررسی برازش سؤالات با مدل‌ها و مقایسه سطح زیر نمودار تابع آگاهی در بازه توانایی ۴- تا ۴ در دو مدل استفاده از مدل سه‌پارامتری را توصیه می‌کند. با توجه به اینکه سؤالات آزمون از نوع

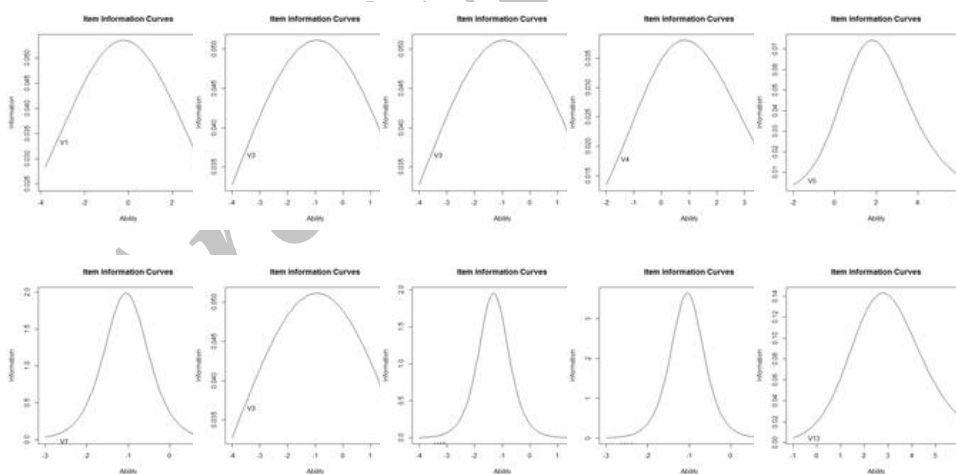
ساخت و روانسازي آزمون تفكر انتقادي: کاربرد نظريه سؤال - پاسخ

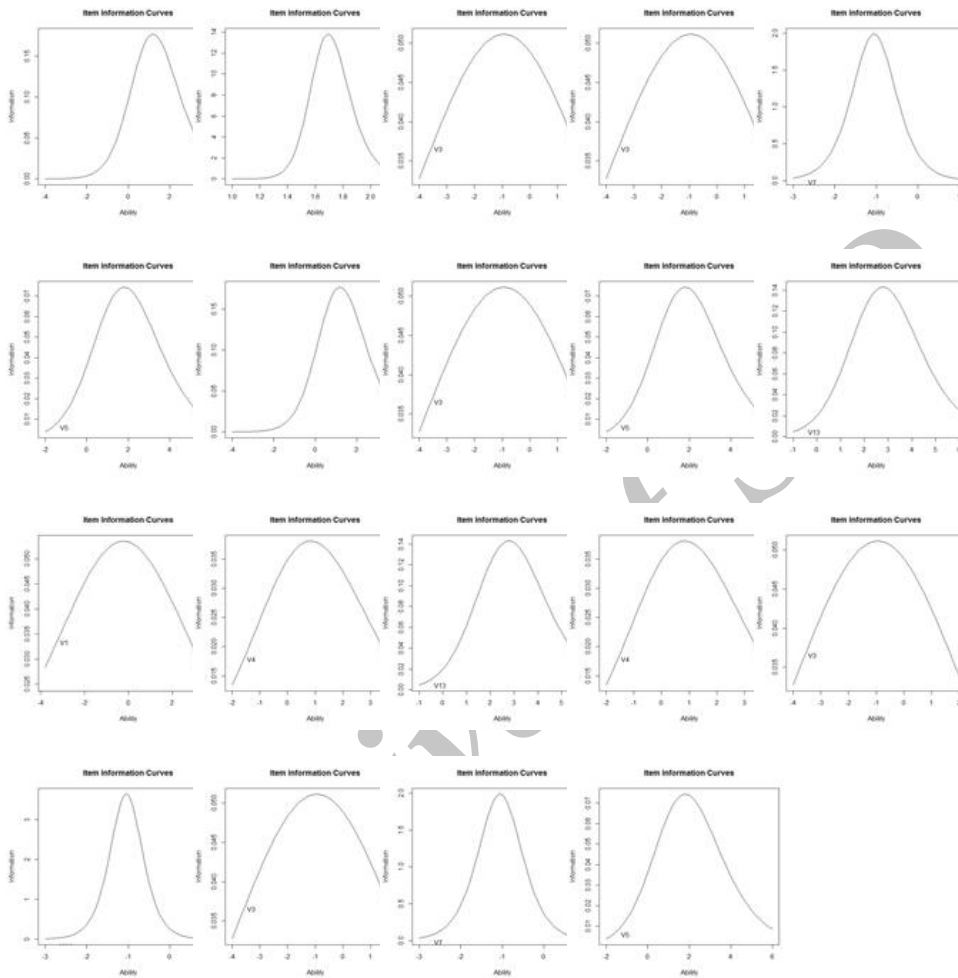
سؤالات چندگزينه‌اي است، احتمال ارائه پاسخ صحيح با استفاده از حدس دور از ذهن نيست؛ بنا بر اين استفاده از مدل سه پارامتری برای محاسبه پارامترهای سؤالات قابل توجه است.

برآوردهای اولیه پارامترهای سؤالات

برآورد پارامترها و خطای استاندارد آنها در مدل لوجستیک سه پارامتری نشان داد ضریب تشخیص سؤالات در دامنه بین ۰/۳ تا ۱/۴۸ است. پارامترهای مربوط به دشواری نیز در بازه بین ۱/۳- تا ۱/۵۳ است. همچنین پارامتر حدس بین ۰ تا ۰/۵ می‌باشد. بررسی نمودار ویژه مربوط به سؤالات نشان می‌دهد متحنی‌های ویژه سؤالات علاوه بر پارامتر مکان، به پارامتر شیب نیز وابسته بوده و برخی از سؤالات دارای مجانب غیر صفر در توانایی‌های پایین است. به عنوان مثال سؤال ۱۰ دارای شیب بسیار زیاد و سؤال ۳۶ دارای مجانب زیاد در توانایی پایین است.

نمودار تابع آگاهی مربوط به ۳۶ سؤال نشان داد برخی از سؤالات از تابع آگاهی مناسبی برخوردار نیستند. تابع آگاهی برخی از سؤالات در کل دامنه توانایی گسترده نیست و تابع آگاهی برخی از سؤالات در دامنه توانایی ۴- تا ۴+ به صورت افزایشی است. پس از بررسی توابع آگاهی سؤالات و حذف سؤالات نامناسب، سطح زیر نمودار تابع آگاهی کل آزمون در بازه توانایی ۴- تا ۴+ از ۲۵۲/۹۵ به ۱۰۸/۱۱ تغییر کرد؛ اما نتایج مربوط به تابع آگاهی سؤالات نشان‌دهنده افزایش در بیشتر سؤالات است. در ادامه نمودار ۳، تابع آگاهی سؤالات بعد از حذف سؤالات نامناسب را نشان می‌دهد.





نمودار ۳- تابع آگاهی سؤالات بعد از حذف سؤالات نامناسب

ساخت و روانسازی آزمون تفکر انتقادی: کاربرد نظریه سؤال - پاسخ

جدول ۲ نتایج تحلیل پارامترهای سؤالات را نشان می‌دهد. نتایج نشان داد دامنه ضریب تشخیص بین ۰/۱۱۳ تا ۱/۸۹۹ و ضریب دشواری بین ۰/۹۹۰- تا ۲/۷۸۰ است. همچنین ضریب حدس نیز در دامنه بین ۰/۰۰۰ تا ۰/۷۹۰ قرار دارد.

جدول ۲- پارامترهای حدس، دشواری و تشخیص سوالات پرسشنامه نهایی

پارامتر تشخیص	پارامتر دشواری	پارامتر حدس	شماره سؤال	مؤلفه
۱/۱۴۰	۱/۱۳	-۰/۱۲	۱	تشخیص نقطه نظر دیگران
۰/۲۸۳	۰/۶۳	-۰/۰۰	۲	
۱/۲۱۶	-۰/۱	-۰/۰۰	۳	
۱/۳۷۰	۱/۷۹	-۰/۰۰	۴	
۱/۱۷۱	۲/۳۲	-۰/۰۰	۵	ارزیابی شواهد مرتبط با سایر نظرات
۰/۱۹۸	۱/۱۴	-۰/۰۲	۶	
۰/۳۳۶	۲/۵۴	-۰/۵۹۲	۷	
۰/۶۸۷	۲/۷۸	-۰/۳۸۷	۸	
۱/۶۲۲	-۰/۵۵	-۰/۱۸۲	۹	
۰/۵۷۱	۱/۳۹	-۰/۰۰	۱۰	
۰/۲۵۷	۲/۳۶	-۰/۰۰	۱۱	سنجش متصفانه استدلال‌های مختلف و شواهد مربوطه
۱/۷۲۹	-۰/۵۴	-۰/۰۰	۱۲	
۰/۱۱۳	۰/۵۹	-۰/۳۸۳	۱۳	
۱/۷۳۹	۱/۹۶	-۰/۳۵۹	۱۴	
۰/۶۵۴	۲/۱۴	-۰/۰۰	۱۵	تشخیص مفروضات نادرست
۰/۳۹۳	-۰/۷	-۰/۱۰۴	۱۶	
۱/۸۳۵	۲/۷۳	-۰/۱۹۶	۱۷	
۱/۳۵۶	۲/۲۸	-۰/۳۰۴	۱۸	
۱/۴۸۳	۲/۶۷	-۰/۰۰	۱۹	
۱/۸۹۹	۰/۹۹	-۰/۲۰۴	۲۰	تفکر در مورد موضوعات به صورت ساختاریافته
۱/۰۱۶	۱/۳۷	-۰/۴۲۹	۲۱	
۰/۸۵۰	۰/۳۷	-۰/۱۶۳	۲۲	
۰/۵۶۳	-۰/۵	-۰/۲۹۸	۲۳	نتیجه‌گیری در مورد معتبر و قابل تصدیق بودن؛ استدلال‌ها بر اساس شواهد و مفروضات مناسب
۰/۸۲۰	-۰/۵۲	-۰/۷۹۰	۲۴	
۱/۰۸۷	۰/۰۵	-۰/۰۳۳	۲۵	
۱/۱۰۱	-۰/۰۶	-۰/۶۴۷	۲۶	ارائه نقطه نظرات خود به شیوه ساختاریافته، واضح و مستدل
۰/۷۶۱	-۰/۹۹	-۰/۳۱۵	۲۷	
۰/۶۵۹	۰/۸۲	-۰/۱۲۸	۲۸	
۱/۵۰۳	-۰/۴۵	-۰/۱۰۳	۲۹	

بحث و نتیجه‌گیری

این مطالعه با هدف ساخت و ارزشیابی ویژگی‌های آزمون تفکر انتقادی به کمک روش‌شناسی نظریه سؤال-پاسخ صورت گرفت. به‌منظور بررسی ویژگی‌های روان‌سنجی آزمون، از نظریه کلاسیک و نظریه سؤال-پاسخ به‌صورت مکمل استفاده شد. در بررسی ویژگی‌های سؤالات بر اساس نظریه کلاسیک، از آلفای کرونباخ، دشواری سؤالات و ضریب تشخیص سؤالات استفاده شد. بر اساس آلفای کرونباخ، پایایی کلی ۰/۷ بوده و برای مؤلفه‌ها نیز پایایی بین ۰/۴۶ تا ۰/۹۰ به دست آمد. برای دو مؤلفه سنجش منصفانه استدلال‌های مختلف و شواهد مربوطه و تشخیص مفروضات نادرست کمتر از ۰/۷ است. دشواری سؤالات در فاصله بین ۰/۱۸ تا ۰/۹۰ قرار دارد؛ بنابراین دشواری اکثر سؤال‌ها، در فاصله مناسب قرار دارد. همچنین ضریب تشخیص سؤالات در بازه بین ۰/۱۱ تا ۰/۵۳ قرار دارد که بازه‌ای بین ضریب تشخیص «کم تا متوسط» را شامل می‌شود. در استفاده از نظریه سؤال - پاسخ از مدل لوجستیک سه پارامتری برای ارزشیابی سؤالات استفاده شد. به‌منظور بررسی مفروضه‌های این نظریه (تک‌بعدی بودن و استقلال موضعی) از تحلیل عاملی اکتشافی استفاده شد. نتایج تحلیل عاملی وجود یک عامل غالب و استقلال موضعی سؤالات را تأیید کرد. به‌منظور بررسی مناسب بودن یا مناسب نبودن سؤالات از تابع آگاهی سؤالات استفاده شد. بررسی توابع آگاهی سؤالات نشان داد برخی از سؤالات دارای آگاهی بسیار کمی بوده و یا تابع آگاهی در طول بازه توانایی گسترده نشده است. پس از حذف سؤالات نامناسب، پارامترهای سؤالات باقی‌مانده دوباره برآورد شد. برآوردهای جدید نشان داد دامنه ضریب تشخیص بین ۰/۱۱۳ تا ۱/۸۹۹ و ضریب دشواری بین ۰/۹۹۰- تا ۲/۷۸۰ است. همچنین ضریب حدس نیز در دامنه بین ۰/۰۰۰ تا ۰/۷۹۰ قرار دارد. آزمون تفکر انتقادی با ۷ مؤلفه و ۲۹ دارای ویژگی‌های مطلوب برای سنجش سازه مذکور است. این آزمون تنها در جامعه دانشجویان کارشناسی و کارشناسی ارشد روسازی شده است؛ بنابراین لازم است قبل از استفاده از این آزمون در سایر جوامع، این آزمون در آن جوامع نیز روسازی شود. هر چند برای ساخت این آزمون مدل لوجستیک سه پارامتری نظریه سؤال-پاسخ مناسب تشخیص داده شده و مورد استفاده قرار گرفت، پیشنهاد می‌شود سایر مدل‌های نظریه مذکور از جمله مدل‌های چندبعدی نیز برای بررسی مناسب آزمون مورد بررسی قرار گیرد.

Reference

- Ali pour, V., Seifnaraghi, M., Naderi, E., & Shariatmadari, A. (2013). Reflection on the barriers to critical thinking in the curriculum of secondary education. *Curriculum Planning*, 9 (2), 1-15. {Persian}.
- Bailin, S. (2002). Critical thinking and science education. *Science & Education*, 11(4), 361–375.
- Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Conceptualizing critical thinking. *Journal of Curriculum Studies*, 31(3), 285–302.
- Case, R. (2005). Moving critical thinking to the main stage. *Education Canada*, 45(2), 45–49.
- Cattell, R. B. (1966). *Handbook of Multivariate Experimental Psychology*. Chicago: Rand McNally.
- Cottrell, S. (2005). *Critical Thinking Skills: Developing Effective Analysis and Argument*. Basingstoke. Palgrave Macmillan.
- DeMars, C. (2010). *Item Response Theory*. New York: Oxford University Press.
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership*, 43(2), 44–48.
- Facione, P. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae, CA: The California Academic Press.
- Facione, P. (2000). The disposition toward critical thinking: Its character, measurement, and relation to critical thinking skill. *Informal Logic*, 20(1), 61–84.
- Falsafinejad, M.R. (2010). *Fundation of Item-Response Theory*. Tehran: Allameh tabatabaei press. {Persian}.
- Hale, E.S. (2008). *Project demonstrating excellence. A Critical Analysis of Richard Paul's Substantive Trans-disciplinary Conception of Critical Thinking*. Submitted in partial fulfillment of the Requirements for the Degree of Doctor of Philosophy in Interdisciplinary Studies. Union Institute & University, Cincinnati, OHIO.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449–455.
- Hooman, HA (1992). *Psychological and educational measurement (test preparation technique)*. Tehran: Entesharateh-Dorsa. {Persian}.
- Khalili, H., & Soleimani, M. (2003). Determining the trust, credibility and norm scores of the California Critical Thinking Skills Test-Form B (TSTCC-B). *Journal of Babol University of Medical Sciences*, 2, 84-90, {Persian}
- Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into Practice*, 32(3), 131–137.
- Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. Springer, New York.
- Lipman, M. (1988). Critical thinking-What can it be? *Educational Leadership*, 46(1), 38–43.
- Loehlin, J. C. (1987). *Latent variable models: An introduction to factor, path, and structural analysis*. Hillsdale, NJ: Erlbaum.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Addison Wesley, Reading, MA.
- McPeck, J. E. (1990). Critical thinking and subject specificity: A reply to Ennis. *Educational Researcher*, 19(4), 10–12.
- Moatari, M., Abedi, HA., Amini., & Fathi-Azar, A. (2001). the impact of Rethinking on critical thinking skills of nursing students in Tabriz city. *Iranian Journal of Medical Education*, 4, 55-60. {Persian}.
- Paul, R. W. (1992). Critical thinking: What, why, and how? *New Directions for Community Colleges*, 1992(77), 3–24.
- Rasch, G.(1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research, Copenhagen, Denmark.

- Samejima, F. (1977). A use of the information function in tailored testing. *Applied Psychological Measurement*, 1, 233-347.
- Shariatmadari, A. (2006). The principles of education, Isfahan: Entezarateh Mashaal. {Persian}.
- Sternberg, R. J. (1986). *Critical thinking: Its nature, measurement, and improvement National Institute of Education*
- Thayer-Bacon, B. J. (2000). *Transforming critical thinking: Thinking constructively*. New York, NY: Teachers College Press.
- Tindal, G., & Nolet, V. (1995). Curriculum-based measurement in middle and high schools: Critical thinking skills in content areas. *Focus on Exceptional Children*, 27(7), 1-22.
- Vander Linden, W. J. (2016). *Handbook of Item Response Theory, Volume Two: Statistical Tools*. Chapman & Hall/CRC Statistics in the Social and Behavioral Sciences: CRC Press
- Willingham, D. T. (2007). Critical thinking: Why is it so hard to teach?. *American Educator*, 8-19.
- Worrell JA., & Profetto-McGrath, J. (2007). Critical thinking as an outcome of context-based learning among post RN students: A literature review. *Nurse Education Today*; 27:420-426.
- Zori, S., Nosek, J. L., & Musil, C. M. (2010). Critical Thinking of Nurse Managers Related to Staff RNs' Perceptions of the Practice Environment. *Journal of Nursing Scholarship*, 42(3), 305-313.