

استفاده از تحلیل داده‌های شبکه‌ای در بهبود الگوریتم‌های انطباقی کامپیوتری*
 نادیا سلطانی^۱، شهرام واحدی^۲، منصور بیرامی^۳، محمدحسین زرغامی^۴

Use of network data analysis to improve algorithm of computerized adaptive tests

Nadiyai Soltani¹, Shahram Vahedi², Mansour Bayrami³, Mohammad Hossein Zarghami⁴

چکیده

زمینه: روش‌شناسی شبکه امکان مطالعه همزمان چندین هدف را میسر می‌سازد و قابلیت‌هایی دارد که منجر به ارائه ابزارهای تحلیلی مناسب می‌گردد. پژوهش‌هایی در خصوص استفاده از تحلیل داده‌های شبکه در حوزه سنجش تربیتی وجود دارد اما کمتر پژوهشی است که بصورت عملی به ساخت آزمون انطباقی مبتنی بر نظریه شبکه و یا ارائه الگوریتم آن پرداخته باشد. آیا امکان استفاده از علم شبکه در ساخت آزمون‌های انطباقی وجود دارد؟ **هدف:** این مطالعه با هدف بهبود الگوریتم آزمون‌های انطباقی مبتنی بر تحلیل داده‌های شبکه بر روی زبان آموزان مراکز و مؤسسات آموزش زبان شهر تهران انجام شد. **روش:** این پژوهش از نوع توصیفی و تمرکز اصلی آن بر تحلیل سؤالات آزمون با تکنیک تحلیل داده‌های شبکه بود. جامعه آماری کلیه زبان آموزان ده مؤسسه برتر آموزش زبان شهر تهران در سال ۱۳۹۸ بود که قابلیت خواندن زبان انگلیسی را در حداقل سطح داشتند. از میان آنان ۱۵۵۶ نفر به روش نمونه‌گیری در دسترس انتخاب شدند. ابزار سنجش آزمون ۱۴۰ سؤالی خزانه لغات انگلیسی (VST) (نیشن و وارینگ، ۱۹۹۷) بود. تحلیل سؤالات آزمون از طریق نظریه کلاسیک، نظریه سؤال - پاسخ و تحلیل شبکه مبتنی بر الگوریتم فراچترمن - رینگولد انجام شد. **یافته‌ها:** پارامترهای سؤالات آزمون بر اساس نظریات کلاسیک و سؤال - پاسخ محاسبه گردید. نقشه ارتباط بین سؤالات مبتنی بر الگوریتم فراچترمن - رینگولد، بر اساس آزمون آماری همبستگی تفکیکی با تعدیل آلفا ($P < 0/01$) ترسیم شد، شاخص‌های مرکزیت شبکه و الگوریتم انطباقی آزمون استخراج گردید. **نتیجه‌گیری:** استفاده از مزایای تکنیک تحلیل داده‌های شبکه‌ای مانند قابلیت دیداری‌سازی نتایج، ارائه گزارش‌های ساده و جامع و در نظر گرفتن اهمیت سؤالات در ساختار شبکه ارتباطی بین سؤالات به ارتقاء روش‌های آزمون‌سازی انطباقی منجر شده است. **واژه کلیدها:** آزمون انطباقی کامپیوتری، نظریه شبکه، نظریه سؤال - پاسخ، انتخاب سؤال، الگوریتم.

Background: Network methodology makes it possible to study several goals at the same time and has capabilities leads to providing appropriate analytical tools. There are studies on the use of network data analysis in the field of educational assessment but there is less research that has practically constructed an adaptive test based on network theory or presented its algorithm. Is it possible to use network science to build adaptive tests? **Aims:** This study was performed with the aim of improving the adaptive test algorithm based on network data analysis on language learners of language teaching centers and institutions in Tehran. **Method:** This research was descriptive and its main focus was on the analysis of test questions using network data analysis technique. The statistical population included all students of the top ten language teaching institutes in Tehran in 2019 who had the ability to read English at a minimum level. Among them, 1556 people were selected by convenience sampling method. The tool used was the 140-item English Vocabulary Size Test (VST) (Nation & Waring, 1997). The analysis of test questions was performed through classical test theory, item response theory and network analysis based on Fruchterman-Ringold algorithm. **Results:** The parameters of the test questions were calculated based on classical test theory and item respons theory. The relationship map between the questions was drawn based on Fruchterman-Ringold algorithm and statistical test of partial correlation with alpha modulation ($p < 0/01$), network centrality parameters and adaptive test algorithm were extracted. **Conclusions:** Utilizing the benefits of network data analysis techniques such as results visibility, providing simple and comprehensive reports and considering the importance of the questions in the structure of the communication network between the questions have led to the improvement of adaptive testing methodes. **Key Words:** Computerized adaptive testing, network theory, item response theory, question selection, algorithm.

Corresponding Author: vahedi117@yahoo.com

* این مقاله مستخرج از رساله دکتری نویسنده اول است.

^۱ دانشجوی دکتری روانشناسی تربیتی، گروه علوم تربیتی، دانشگاه تبریز، تبریز، ایران

^۱ Ph.D Student of Educational Psychology, Department of Educational Science, University of Tabriz, Tabriz, Iran

^۲ استاد، گروه علوم تربیتی، دانشگاه تبریز، تبریز، ایران(نویسنده مسئول)

^۲ Professor, Department of educational Sciences, University of Tabriz, Tabriz, Iran(Corresponding Auther)

^۳ استاد، گروه روانشناسی، دانشگاه تبریز، تبریز، ایران

^۳ Professor, Department of Psychology, University of Tabriz, Tabriz, Iran

^۴ استادیار، مرکز تحقیقات علوم رفتاری، دانشگاه علوم پزشکی بقیه‌الله، تهران، ایران

^۴ Assistant Professor, Behavioral Sciences Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran

مقدمه

تمیز سؤال کاملاً به نمونه‌ای از آزمودنی‌ها وابسته هستند که از روی آنها کسب شده‌اند و مقدار این پارامترها از یک گروه نمونه به گروه دیگر متفاوت هستند (ایلهان و گولر، ۲۰۱۸). نظریه سؤال - پاسخ که مدل متغیر مکنون^۴ و یا نظریه نوین هم خوانده می‌شود به کارهای لرد و ناویک (۱۹۶۸) برمی‌گردد. در بیشتر موقعیت‌های عملی تصمیمات بر اساس کمیت‌هایی گرفته می‌شود که نمی‌توان آنها را به طور مستقیم مشاهده نمود. این کمیت‌ها به متغیرهای مکنون مشهورند. هر مدلی که یک ساختار پنهان را به یک ساختار مشاهده شده مربوط کند، مدل متغیر مکنون نام دارد (ضرغامی، قائمی و قائمی، ۱۳۹۲). نظریه سؤال - پاسخ به عنوان یکی از نظریات مبتنی بر متغیر مکنون، یک تئوری ریاضیاتی است که مفروضات آن از طریق داده‌ها و مبتنی بر رویکرد نیکویی برازش قابل آزمون است و ویژگی‌هایی دارد (مانند فقدان وابستگی چرخشی^۵) که جذابیت بالایی برای کاربردهای خاص مانند تهیه بانک سؤال^۶، ساخت آزمون‌های انطباقی و... را فراهم می‌آورد (بورسبوم، ۲۰۰۸). آزمون‌های انطباقی کامپیوتری جایگزینی جدی برای آزمون‌های مداد - کاغذی قلمداد می‌شوند. انعطاف‌پذیری، تسهیل شرایط اجرا، افزایش کارایی، امکان سنجش دقیق و گسترده‌تر، امکان نمره‌دهی ماشینی، فراهم‌سازی امکان بازخورد سریع، تخمین‌های فوری و آنی از سطح توانایی پاسخ‌دهندگان، افزایش امنیت آزمون^۷ و کاهش ریسک تقلب (هر آزمون‌دهنده ترتیب متفاوتی از سؤالات آزمون را دریافت می‌کند)، از جمله دلایل اصلی کاربرد آزمون‌های انطباقی کامپیوتری است. (اپل، رسینجر، ایکمایر و هلم، ۲۰۱۷).

هر فرآیند آزمون‌های انطباقی کامپیوتری، نیازمند یک بانک سؤال مدرج است که معمولاً از نظریه سؤال - پاسخ جهت مدرج کردن سؤالات استفاده می‌شود. مرحله نخست شامل انتخاب یک و یا چند سؤال مناسب به عنوان اولین سؤالات آزمون است، در مرحله دوم سؤالات به صورت پشت سر هم از بانک انتخاب شده و توانایی بعد از ارائه هر سؤال برآورد می‌شود. مرحله سوم، توقف است و پارامترهایی برای قانون توقف آزمون تعیین می‌شود و در نهایت سطح توانایی آزمون‌دهنده برآورد می‌شود (واندرلیندن و گلاس

روانشناسی تربیتی به عنوان شاخه‌ای از علم قلمداد می‌شود که علاوه بر کاربردهای نظریات روانشناسی در حوزه تعلیم و تربیت به اندازه‌گیری روانشناختی^۱ در سطح نظری (آزمون‌سازی نظری) و تکنیک می‌پردازد تا فعالیت‌هایی مانند سنجش، اندازه‌گیری، ارزشیابی و حوزه‌های وابسته را انجام دهد و مسائل مربوط به این حوزه را حل نماید (آزمون‌سازی کاربردی^۲). این وجه از روانشناسی تربیتی در حال حاضر نقش مهم و اصلی در همه ابعاد زندگی ما بازی می‌کند (راست و گلوبوک، ۱۹۸۹). آزمون‌سازی مدرن ریشه در روانشناسی تربیتی دارد و رویکرد های علمی به روز را با در نظر گرفتن موضوعات اخلاقی و سیاسی در مقیاس وسیعی در مطالعات جامعه به شدت شبکه‌ای شده بکار می‌گیرد، خصوصاً که جامعه کنونی دارای تنوع زیاد و در حال بین‌المللی شدن^۳ است (کالبرتسون، ۲۰۱۵). کاربرد فن‌آوری‌های نوین اطلاعاتی در جهان معاصر با سرعت فزاینده‌ای در حال گسترش است و تأثیر چشم‌گیری بر تمام جوانب زندگی از جمله ابعاد اقتصادی، اجتماعی، صنعتی و آموزشی داشته است (شریفی‌یگانه، فلسفی‌نژاد، دلاور، فرخی و جمالی، ۱۳۹۵). در حوزه آموزش، فن‌آوری‌های اطلاعاتی، فرآیند تدریس و یادگیری را متحول کرده‌اند، سنجش نیز بعنوان بخش جدایی‌ناپذیر فرآیند آموزش بی‌تأثیر از ظهور و پیشرفت‌های فن‌آوری‌های نوین از جمله کامپیوتر نبوده است (بنت، ۲۰۰۲). با استفاده از فن‌آوری کامپیوتر می‌توان آزمون‌های مداد - کاغذی مرسوم را به گونه‌ای کارآمدتر به کار برد و اندازه‌ای دقیق‌تر از توانایی آزمودنی بدست آورد (کلندر، ۲۰۱۱). آزمون‌سازی انطباقی کامپیوتری به عنوان جایگزینی برای آزمون‌های ترتیبی (و یا سؤال ثابت) ارائه شده‌اند. در حال حاضر در حوزه سنجش و اندازه‌گیری، نظریه کلاسیک آزمون و نظریه سؤال - پاسخ به عنوان دو نظریه غالب وجود دارند. با اینکه نظریه کلاسیک آزمون بیشتر از نیم قرن بر نهضت آزمون‌سازی حاکم بوده ولی بعد از اینکه برخی مطالعات محدودیت‌های این نظریه را آشکار ساختند، زمینه ارائه نظریه‌های نوین اندازه‌گیری فراهم شد (تی، میادی و کائو، ۲۰۱۵). اولین و مهمترین محدودیت نظریه کلاسیک این است که مقدار دشواری و

4. latent variable model.

5. statistical and mathematical models

6. item bank

7. test security

1. psychological measurement

2. applied psychometrics

3. internationalism

شبکه در حوزه روانشناسی و اندازه‌گیری روانی می‌تواند تلاش بین رشته‌ای قلمداد شود که در آن نظریات روانشناسی با روش‌شناسی^۲ کمی آماری و ریاضی شکل گرفته و گسترش می‌یابد و در کنار آن، برای چالش‌های تحلیلی در این حوزه پاسخ مناسب ارائه می‌دهد. دو دلیل موجب نگاه جدید به رویکرد شبکه‌ای در حوزه اندازه‌گیری روانی شده است. اول این که در سال‌های اخیر مدل‌ها و نظریات مربوط به متغیرهای پنهان سیطره داشته‌اند و دوم این که بر خلاف مدل‌های پنهان، دیدگاه شبکه‌ای به گسترش سیستم‌های پویا کمک کرده است. گستردگی سیستم‌های پویا از لحاظ محتوایی بسیار زیاد است و سیستم‌های مکانیکی تا نشانه‌شناسی بیماری‌ها را در بر می‌گیرد (بورسبوم، ۲۰۰۸).

هدف در آزمون‌سازی تحلیل سؤالات و تعیین شاخص‌های آزمون‌سازی، شناسایی سؤالات نامناسب و اصلاح آنها و در نهایت گسترش یک ابزار اندازه‌گیری است. در سایر روش‌های تحلیل سؤال دستیابی به این هدف با مشخص کردن پارامتر سؤالات بدست می‌آید. در روش شبکه‌ای، تحلیل سؤال بر اساس نقشه و شاخص‌های کاملاً منطبق بر آن، انجام می‌شود که نتیجه ارتباط درونی سؤالات مختلف آزمون است (دورنوی، ۲۰۱۴). با همه مزایایی که نظریه سؤال - پاسخ دارد اما به دلیل اینکه مبتنی بر نظریه خصیصه مکنون و مدل‌های بازتابی^۳ است با مشکلات فلسفی جدی روبروست و انتقادات فلسفی اساسی بر آن وارد است. مدل‌های بازتابی نمی‌توانند تبیین درستی از ارتباط بین سازه‌های روانی و متغیرهای مشاهده شده ارائه نمایند (ادوارد، ۲۰۱۱). همچنین شاخص‌های نظریه سؤال - پاسخ مبتنی بر روش‌های تجربی ریاضیاتی هستند که پیچیدگی بالایی دارند و قدرت تحلیل سریع و گزارش‌دهی را دشوار می‌سازند. تحلیل داده‌های شبکه‌ای دارای توانمندی‌هایی است که امکان بهبود الگوریتم‌های سنجش انطباقی مبتنی بر نظریه سؤال - پاسخ را فراهم می‌آورند (کالرتسون، ۲۰۱۵). مهمترین مزیت این تحلیل در سنجش انطباقی نسبت به شیوه‌های مبتنی بر نظریه سؤال - پاسخ در نظر گرفتن اهمیت هر سؤال در شبکه محتوایی سؤالات است. اگرچه رویکرد شبکه‌ای از ریاضیات قوی برخوردار است اما انطباق ریاضیات آن با نمایش بصری شبکه، قابلیت دیداری‌سازی نتایج (که امکان درک درست ریاضیات را

سبز، ۲۰۰۰). در سنجش انطباقی به طور نظری هر آزمودنی مجموعه منحصر به فردی از سؤالات را دریافت می‌کند و بجای اینکه به هر آزمودنی یک آزمون ثابت داده شود، سؤال‌ها متناسب با توانایی آزمودنی‌ها در جریان آزمون انتخاب و اجرا می‌شوند. آزمون انطباقی دقیق‌ترین اندازه‌گیری از توانایی آزمودنی‌ها را با حداقل سؤالات ارائه می‌دهد. شرایط تستی ایده‌آل آن است که به هر آزمودنی تستی داده شود که با توانایی او برازش یا سازگاری داشته باشد (بورسبوم، ۲۰۰۶). اگر آزمون به گونه‌ای انتخاب شود که بیشترین اطلاعات را در مورد توانایی فرد فراهم آورد، بدون کاهشی در دقت اندازه‌گیری، آزمون‌ها می‌توانند کوتاه‌تر شوند. (لرد و ناویک، ۱۹۶۸). اما این روش‌ها محدودیت‌هایی دارند. انتخاب مطلوب سؤالات تنها زمانی امکان‌پذیر است که بانک سؤال خوب پارامتری شده و در حد لزوم بزرگ و در دسترس باشد. همچنین نظریه سؤال - پاسخ وابسته به نظریه‌ای کلی‌تر به نام نظریه خصیصه مکنون است (فایرز، ۲۰۰۷). در قرن اخیر مدل‌های شبکه گسترش یافت که رابطه احتمالی بین سازه‌های نظری و داده‌های تجربی را فرمول‌بندی می‌کرد. این مدل، ساختاری فرضی را متصور می‌شود و چگونگی جایگاه یک شیء را در این ساختار و ارتباط آن با سایر اشیاء را در مجموعه‌ای از متغیرهای شاخص که منظور همان متغیرهای مشاهده شده است مشخص و معلوم می‌سازد (پلاجنر و ولل، ۲۰۱۵). به یک گراف^۱ بزرگ یا نسبتاً بزرگ، شبکه گفته می‌شود. منظور از گراف ساده (G)، زوج مرتب (V,E) است که V مجموعه‌ای ناتهی و اعضای E زیر مجموعه‌های دو عضوی از V هستند. به هر عضو V یک راس یا گره G و به هر عضو E یک یال G می‌گوییم (فورتونانو، ۲۰۱۰). استفاده از مدل شبکه‌ای در چند سال اخیر در بررسی مجموعه‌هایی از متغیرها که ارتباط درونی با یکدیگر دارند، گسترش پیدا کرده است. دلایل زیادی برای کاربرد نظریه شبکه در حوزه‌های مختلف فنی، ریاضی، پزشکی و انسانی وجود دارد که مهمترین آنها عبارتند از: الف) گرایش رو به گسترشی که به رویکرد سیستمی در حوزه علم است، ب) فرار از تقلیل‌گرایی که ویژگی اصلی علم قرن گذشته است و ج) قابلیت ساده‌سازی، مدیریت، ذخیره و مجموعه‌سازی داده‌های با حجم بالا (اپسکمپ، کرامر، ولدراپ، اشمیتن و بورسبوم، ۲۰۱۲). تحلیل

2. methodological

3. reflective interpret

1. graph

روش

روش این پژوهش از نوع توصیفی است. داده‌های حاصل از اجرای پرسشنامه در دو مرحله مجزا تحلیل شدند. ابتدا پارامتر سؤالات بر اساس نظریات کلاسیک و سؤال - پاسخ (مدل پاسخ مدرج) بدست آمد سپس به منظور رسیدن به الگوریتم انطباقی، شبکه ارتباط بین سؤالات مبتنی بر الگوریتم فراچترمن - رینگولد ترسیم گردید و بر اساس داده‌های شبکه‌ای و شاخص‌های مربوط به آن، سؤالات پرسشنامه مورد تحلیل قرار گرفت.

به منظور بررسی سؤالات، آزمون روی گروهی از شرکت کنندگان اجرا شد. لازم به ذکر است تمرکز اصلی این پژوهش بر روش‌شناسی و تحلیل سؤالات آزمون است و تعمیم نتایج از اهمیت کمتری برخوردار است. جامعه آماری پژوهش کلیه زبان آموزان ده مؤسسه برتر آموزش زبان انگلیسی شهر تهران بود که قابلیت خواندن زبان انگلیسی را در حداقل سطح داشتند و حجم نمونه ۱۵۵۶ نفر داده‌های معتبر است که به صورت در دسترس و داوطلبانه از این مؤسسات انتخاب شدند. تحلیل سؤالات آزمون و ارائه الگوی ساخت آزمون انطباقی با استفاده از نرم‌افزار R و بسته‌های نرم افزاری mirt و qgraph انجام شد.

ابزار

آزمون ۱۴۰ سؤالی خزانه لغات انگلیسی است که توسط نیشن و وارینگ (۱۹۹۷) تهیه شده است. این آزمون از سطوح مختلف دشواری لغات تشکیل شده و سعی کرده است که کلیه لغات زبان انگلیسی را پوشش دهد. انتخاب لغات در این آزمون بر اساس خانواده لغاتی هستند که از پیکره ملی بریتانیایی^۱ گرفته شده است که در برگیرنده ۱۰۰ میلیون لغت است و در بخش‌های مختلف ملل انگلیسی زبان وجود دارد. مجموعه لغات پیکره ملی بریتانیایی شش سطح مختلف دارد که در این آزمون همگی سطوح وجود دارند. در این آزمون هر لغت نماینده ۱۰۰ تا ۲۰۰ خانواده لغتی است. اجرای این آزمون می‌تواند برای زبان‌آموزان در سطوح مختلف تسلط استفاده شود. اجرای مختلف این آزمون توسط نیشن و بگلار (۲۰۰۷) روی زنان و مردان با سطوح مختلف توانایی زبانی به مقادیر پایایی بالایی منجر شده است (مقدار پایایی راش در حدود ۰/۹۶ برآورد شده است). نمره‌دهی و تفسیر آزمون آسان است. سؤالات

میسر می‌سازد)، ارائه گزارش‌های ساده و جامع و میزان پیچیدگی کمتر تحلیل‌های شبکه نسبت به نظریه سؤال - پاسخ از جمله قابلیت‌های این رویکرد محسوب می‌شوند (سرین، کاتر و لیتمن، ۲۰۲۰). از جمله مطالعات پیشین می‌توان به پژوهش کیونگ و یانگ (۲۰۱۲) اشاره کرد، آنها در پژوهش خود به بررسی مبانی نظری علم شبکه و مقایسه آن با نظریات کلاسیک و سؤال - پاسخ پرداخته و بیان کرده‌اند که با استفاده از تحلیل داده‌های شبکه می‌توان توانایی آزمودنی‌ها را بطور مؤثر و کارآمد ارزیابی کرد و دقت و سرعت در برآورد توانایی را بهبود بخشید. دورنوی (۲۰۱۴) در پژوهش خود به بررسی شبکه‌های پویا و معرفی شاخص‌های شبکه پرداخته است. پلاجز و ولل (۲۰۱۵) با مطالعه بر روی دانش‌آموزان دبیرستانی چندین مدل شبکه بیز را برای طراحی آزمون انطباقی ریاضی پیشنهاد کردند و به مقایسه نتایج با روش‌های سنتی آزمون‌سازی پرداختند و در نهایت به ارجحیت استفاده از تحلیل داده‌های شبکه در آزمون‌سازی بعنوان روشی کم‌هزینه و آسان در مقایسه با روش‌های سنتی تأکید کردند. کالبرتسون (۲۰۱۵) در بررسی خود به مبانی نظری علم شبکه بعنوان روشی کارآمد در ارزیابی‌های تشخیصی و آموزشی پرداخته و بر اهمیت و قابلیت دیداری‌سازی نتایج با استفاده از شبکه‌های بیز تأکید کرده به نحوی که شبکه‌های بیز یک چارچوب ساده و بصری برای توزیع احتمالات مشترک فراهم می‌کنند. از جمله دیگر مطالعات پیشین می‌توان به مطالعه عراقی، طاهری، جزء رمضانی، عباسپور و ضرغامی (۱۳۹۶) اشاره نمود که از تحلیل داده‌های شبکه‌ای در فرآیند تحلیل سؤالات پرسشنامه استفاده نموده و نتایج حاصل را با شاخص‌های روان‌سنجی مرسوم مقایسه کرده است.

اگرچه در بررسی ادبیات پیشین پژوهش‌هایی در خصوص مبانی نظری و مزایای استفاده از تحلیل داده‌های شبکه در حوزه سنجش تربیتی وجود دارد اما کمتر پژوهشی است که بصورت عملی به ساخت آزمون انطباقی مبتنی بر نظریه شبکه و یا ارائه الگوریتم آن پرداخته باشد. لذا این پژوهش با هدف ارائه الگوی آزمون انطباقی بر اساس علم شبکه اجرا شد. بطور کلی این پژوهش در صدد پاسخگویی به این سؤال است که آیا می‌توان از مزایای تحلیل داده‌های شبکه‌ای در جهت انتخاب سؤالات در سنجش انطباقی استفاده کرد؟

^۱. British National Corpus (BNC)

برای دستیابی به مقادیر پارامتر سؤالات با هدف تشکیل بانک سؤال، تعداد ۱۴۰ سؤال آزمون خزانه لغات از طریق نظریه سؤال - پاسخ تحلیل شد. مدل استفاده شده برای برازش با داده‌ها، مدل تک بعد سه پارامتری است که پارامترهای دشواری، شیب و حدس در آن متغیر است. مقدار پارامتر سؤالات با استفاده از بسته ltm تحت نرم‌افزار R استخراج شده است که نتایج حاصل در جدول ۲ مشاهده می‌شود.

آزمون روشن و بدون ابهام‌اند. اجرای این آزمون برای تمام زبان‌آموزان از کسانی که تنها ۵ لغت تا ۱۰۰۰ لغت هر سطح را می‌دانند، امکان‌پذیر است (نیشن و بگلار، ۲۰۰۷).

یافته‌ها

دشواری و ضریب تمیز ۱۴۰ سؤال مختلف مبتنی بر نظریه کلاسیک اندازه‌گیری در جدول ۱ ارائه شده است. درصد پاسخ های صحیح به عنوان شاخص دشواری و همبستگی دو رشته‌ای نقطه ای بین سؤال و نمره کل به عنوان شاخص تمیز محسوب می‌شوند.

جدول ۱. پارامتر سؤالات آزمون مبتنی بر نظریه کلاسیک

شاخص دشواری		شاخص تمیز بر اساس همبستگی دو رشته ای نقطه‌ای	
درصد پاسخ غلط	درصد پاسخ صحیح	Logit	سؤال در نمره کل موجود است
۰/۰۷۶	۰/۹۲۴	۲/۴۹۷۸	۰/۲۳۶۹
۰/۰۵۳	۰/۹۴۷	۲/۸۸۳	۰/۲۵۲۸
۰/۰۸۹۳	۰/۹۱۰۷	۲/۳۲۲۷	۰/۳۱۷۱
۰/۷۵۵۲	۰/۲۴۴۸	-۱/۱۲۶۷	۰/۳۲۹۱
۰/۰۴۱۸	۰/۹۵۸۲	۳/۱۳۱۱	۰/۲۶۰۹
۰/۰۳۵۶	۰/۹۶۴۴	۳/۳۰۰۲	۰/۲۶۴۵
۰/۲۴۲	۰/۷۵۸	۱/۱۴۱۹	۰/۴۲۶۳
۰/۰۶۱۴	۰/۹۳۸۶	۲/۷۲۷۶	۰/۳۷۵۲
۰/۰۶۹	۰/۹۳۱	۲/۶۰۱۶	۰/۳۶۹۱
۰/۱۰۶	۰/۸۹۴	۲/۱۳۲۳	۰/۳۹۷۶
۰/۳۳۴	۰/۶۶۶	۰/۶۹	۰/۴۱۰۳
۰/۱۱۸۵	۰/۸۸۱۵	۲/۰۰۶۲	۰/۳۵۹۹
۰/۱۰۱۱	۰/۸۹۸۹	۲/۱۸۴۹	۰/۳۶۲۴
۰/۲۰۸۵	۰/۷۹۱۵	۱/۳۳۳۹	۰/۳۰۵۴
۰/۳۵۵۶	۰/۶۴۴۴	۰/۵۹۴۳	۰/۴۳۷
۰/۷۳۲۹	۰/۲۶۷۱	-۱/۰۰۹۵	۰/۲۲۰۵
۰/۳۸۰۱	۰/۶۱۹۹	۰/۴۸۹۳	۰/۳۵۸۵
۰/۱۳۷۴	۰/۸۶۲۶	۰/۸۳۷۲	۰/۳۹۹
۰/۰۹۸۳	۰/۹۰۱۷	۲/۲۱۶	۰/۳۲۶۱
۰/۶۰۰۴	۰/۳۹۹۶	-۰/۴۰۷۲	۰/۳۳۱۹
۰/۰۹۲۷	۰/۹۰۷۳	۲/۲۸۰۵	۰/۴۸۲۲
۰/۲۰۱۵	۰/۷۹۸۵	۱/۳۷۶۷	۰/۴۵۹۷
۰/۴۴۳۵	۰/۵۵۶۵	۰/۲۲۶۹	۰/۴۱۶۹
۰/۴۲۷۵	۰/۵۷۲۵	۰/۲۹۲۲	۰/۴۰۷۲
۰/۲۸۵۹	۰/۷۱۷۴۱	۰/۹۱۵۳	۰/۳۶۴۳
۰/۶۳۸۱	۰/۳۶۱۹	۰/۵۶۷	۰/۳۱۴۶
۰/۴۸۰۵	۰/۵۱۹۵	۰/۰۷۸۱	۰/۳۷۵۶
۰/۴۹۳	۰/۵۰۷	۰/۰۲۷۹	۰/۳۰۱۶
۰/۷۸۴۵	۰/۲۱۵۵	-۱/۲۹۲۲	۰/۱۰۷۴
۰/۲۶۴۳	۰/۷۳۵۷	۱/۰۲۳۸	۰/۴۵۳۲
۰/۳۶۳۳	۰/۶۳۶۷	۰/۵۶۱	۰/۳۴۲۲
۰/۵۰۰۷	۰/۴۹۹۳	-۰/۰۰۲۸	۰/۴۰۰۴

جدول ۱. پارامتر سؤالات آزمون مبتنی بر نظریه کلاسیک						
شاخص دشواری			شاخص تمیز بر اساس همبستگی دو رشته ای نقطه‌ای			
شماره سؤال	درصد پاسخ غلط	درصد پاسخ صحیح	Logit	سؤال در نمره کل موجود است	سؤال در نمره کل موجود نیست	
سؤال ۳۳	۰/۵۵۵۱	۰/۴۴۴۹	-۰/۲۲۱۳	۰/۲۴۴۹	۰/۲۱۸۸	
سؤال ۳۴	۰/۵۳۴۹	۰/۴۶۵۱	-۰/۱۳۹۷	۰/۴۲۳۷	۰/۴۰۰۶	
سؤال ۳۵	۰/۰۹۲۱	۰/۹۰۷۹	۲/۲۸۸۹	۰/۴۳۱۸	۰/۴۱۸۶	
سؤال ۳۶	۰/۲۰۸۵	۰/۷۹۱۵	۱/۳۳۳۹	۰/۴۰۳۹	۰/۳۸۴۳	
سؤال ۳۷	۰/۳۵۷۷	۰/۶۴۲۳	۰/۵۸۵۲	۰/۴۵۸۹	۰/۴۳۷۵	
سؤال ۳۸	۰/۰۸۱۶	۰/۹۱۸۴	۲/۴۲۰۹	۰/۴۹۰۲	۰/۴۷۸۵	
سؤال ۳۹	۰/۳۳۳۳	۰/۶۶۶۷	۰/۶۹۳۱	۰/۴۶۲۳	۰/۴۴۱۴	
سؤال ۴۰	۰/۶۵۶۲	۰/۳۴۳۸	-۰/۶۴۶۴	۰/۲۰۸	۰/۱۸۲۷	
سؤال ۴۱	۰/۶۷۲۹	۰/۳۲۷۱	-۰/۷۲۱۵	۰/۱۴۲۱	۰/۱۱۶۵	
سؤال ۴۲	۰/۵۸۲۳	۰/۴۱۷۷	-۰/۳۳۲۲	۰/۴۶۵۸	۰/۴۴۴	
سؤال ۴۳	۰/۴۹۵۱	۰/۵۰۴۹	۰/۰۱۹۵	۰/۴۲۴۳	۰/۴۰۱۲	
سؤال ۴۴	۰/۴۸۱۲	۰/۵۱۸۸	۰/۰۷۵۳	۰/۴۷۶۱	۰/۴۵۴۳	
سؤال ۴۵	۰/۶۶۵۳	۰/۳۳۴۷	-۰/۶۸۶۹	۰/۲۶۸۱	۰/۲۴۳۶	
سؤال ۴۶	۰/۳۳۴۷	۰/۶۶۵۳	۰/۶۸۶۹	۰/۵۰۹۹	۰/۴۹۰۲	
سؤال ۴۷	۰/۲۹۹۲	۰/۷۰۰۸	۰/۸۵۱۳	۰/۴۶۳	۰/۴۴۲۸	
سؤال ۴۸	۰/۴۸۷۴	۰/۵۱۲۶	۰/۰۵۰۲	۰/۴۴۱۴	۰/۴۱۸۷	
سؤال ۴۹	۰/۵۲۳	۰/۴۷۷	-۰/۰۹۲۱	۰/۳۳۵	۰/۳۱۰۱	
سؤال ۵۰	۰/۱۸۲	۰/۸۱۸	۱/۵۰۲۸	۰/۵۲۶۱	۰/۵۱۰۴	
سؤال ۵۱	۰/۵۹۱۴	۰/۴۰۸۶	-۰/۳۶۹۶	۰/۳۰۸۳	۰/۲۸۳۴	
سؤال ۵۲	۰/۵۲۰۲	۰/۴۷۹۸	-۰/۰۸۰۹	۰/۳۸۳۵	۰/۳۵۹۵	
سؤال ۵۳	۰/۵۹۹	۰/۴۰۱	-۰/۴۰۱۴	۰/۳۷۷۳	۰/۳۵۳۶	
سؤال ۵۴	۰/۳۵۸۴	۰/۶۴۱۶	۰/۵۸۲۲	۰/۵۰۵۶	۰/۴۸۵۵	
سؤال ۵۵	۰/۵۲۲۳	۰/۴۷۷۷	-۰/۰۸۹۳	۰/۲۸۳۷	۰/۲۵۸	
سؤال ۵۶	۰/۲۳۹۹	۰/۷۶۰۱	۱/۱۵۳۳	۰/۵۴۵۱	۰/۵۲۸۲	
سؤال ۵۷	۰/۶۴۹۹	۰/۳۵۰۱	-۰/۶۱۸۷	۰/۲۸۴۳	۰/۲۵۹۸	
سؤال ۵۸	۰/۷۴۲۷	۰/۲۵۷۳	-۱/۰۵۹۹	-۰/۰۷۱	-۰/۰۹۵	
سؤال ۵۹	۰/۵۶۲۱	۰/۴۳۷۹	-۰/۲۴۹۵	۰/۳۷۹۴	۰/۳۵۵۵	
سؤال ۶۰	۰/۶۹۱۸	۰/۳۰۸۲	-۰/۸۰۸۴	۰/۱۴	۰/۱۱۴۸	
سؤال ۶۱	۰/۲۱۲	۰/۷۸۸	۰/۳۱۲۹	۰/۵۳۵۳	۰/۵۱۸۹	
سؤال ۶۲	۰/۷۳۷۱	۰/۲۶۲۹	-۱/۰۳۰۹	۰/۰۸۴	۰/۰۵۹۷	
سؤال ۶۳	۰/۵۹۰۷	۰/۴۰۹۳	-۰/۳۶۶۷	۰/۲۸۱۳	۰/۲۵۵۹	
سؤال ۶۴	۰/۷۱۲	۰/۲۸۸	۰/۹۰۵۱	۰/۱۳۲۲	۰/۱۴۷۷	
سؤال ۶۵	۰/۷۹۳۶	۰/۲۰۶۴	-۱/۳۴۶۷	۰/۱۴۱۷	۰/۱۱۹۶	
سؤال ۶۶	۰/۶۵۴۱	۰/۳۴۵۹	-۰/۶۳۷۲	۰/۲۳۷۶	۰/۲۱۲۶	
سؤال ۶۷	۰/۶۴۵۷	۰/۳۵۴۳	-۰/۶۰۰۴	۰/۲۵۵۲	۰/۲۳۰۲	
سؤال ۶۸	۰/۷۸۰۳	۰/۲۱۹۷	-۱/۲۶۷۶	-۰/۰۳۹۴	-۰/۰۶۲۳	
سؤال ۶۹	۰/۴۷۹۸	۰/۵۲۰۲	۰/۰۸۰۹	۰/۳۰۴۹	۰/۲۷۹۵	
سؤال ۷۰	۰/۱۸۵۵	۰/۸۱۴۵	۱/۴۷۹۶	۰/۴۷۲۳	۰/۴۵۶۳	
سؤال ۷۱	۰/۶۰۱۱	۰/۳۹۸۹	-۰/۴۱۰۱	۰/۳۴۲۳	۰/۳۱۸	
سؤال ۷۲	۰/۵۰۵۶	۰/۴۹۴۴	-۰/۰۲۲۳	۰/۳۵۹۲	۰/۳۳۴۷	
سؤال ۷۳	۰/۳۸۲۱	۰/۶۱۷۹	۰/۴۸۰۴	۰/۴۱۹۴	۰/۳۹۶۹	
سؤال ۷۴	۰/۱۸۳۴	۰/۸۱۶۶	۱/۴۹۳۵	۰/۴۸۴۶	۰/۴۶۷۹	
سؤال ۷۵	۰/۴۲۱۲	۰/۵۷۸۸	۰/۳۱۷۹	۰/۵۴۲	۰/۵۲۲۳	

جدول ۱. پارامتر سؤالات آزمون مبتنی بر نظریه کلاسیک						
شاخص دشواری			شاخص تمیز بر اساس همبستگی دو رشته ای نقطه‌ای			
شماره سؤال	درصد پاسخ غلط	درصد پاسخ صحیح	Logit	سؤال در نمره کل موجود است	سؤال در نمره کل موجود نیست	
سؤال ۷۶	۰/۶۳۸۸	۰/۳۶۱۲	-۰/۵۷	۰/۳۲۱۲	۰/۲۹۷	
سؤال ۷۷	۰/۵۶۲۸	۰/۴۳۷۲	-۰/۲۵۲۴	۰/۲۳۷۷	۰/۲۱۱۶	
سؤال ۷۸	۰/۵۱۳۲	۰/۴۸۶۸	-۰/۰۵۳	۰/۴۳۲۱	۰/۴۰۹۲	
سؤال ۷۹	۰/۴۹۰۹	۰/۵۰۹۱	۰/۰۳۶۳	۰/۳۵۴۴	۰/۳۲۹۸	
سؤال ۸۰	۰/۴۹۱۶	۰/۵۰۸۴	۰/۰۳۳۵	۰/۴۸۱۱	۰/۴۵۹۴	
سؤال ۸۱	۰/۷۱۲۷	۰/۲۸۷۳	-۰/۰۸۵	۰/۱۵۳۴	۰/۱۲۸۸	
سؤال ۸۲	۰/۶۵۴۸	۰/۳۴۵۲	-۰/۰۶۴۰۳	۰/۲۷۶۲	۰/۲۵۱۶	
سؤال ۸۳	۰/۴۶۱۶	۰/۵۳۸۴	۰/۱۵۳۷	۰/۴۳۴۳	۰/۴۱۱۵	
سؤال ۸۴	۰/۷۵۷۳	۰/۲۴۲۷	-۱/۱۳۸۱	۰/۲۰۶۲	۰/۱۸۳۳	
سؤال ۸۵	۰/۵۶۹۷	۰/۴۳۰۳	-۰/۲۸۰۸	۰/۲۴۲۲	۰/۲۱۶۱	
سؤال ۸۶	۰/۷۳۷۱	۰/۲۶۲۹	-۱/۰۳۰۹	۰/۰۸۴۱	۰/۰۵۹۸	
سؤال ۸۷	۰/۷۲۰۴	۰/۲۷۹۶	-۰/۹۴۶۳	۰/۰۱۵۳	-۰/۰۰۹۵	
سؤال ۸۸	۰/۲۵۲۴	۰/۷۴۷۶	۱/۰۸۵۶	۰/۵۷۴۲	۰/۵۵۷۷	
سؤال ۸۹	۰/۷۹۴۳	۰/۲۰۵۷	-۱/۳۵۰۹	-۰/۰۲۴۹	-۰/۰۴۷۲	
سؤال ۹۰	۰/۷۱۴۱	۰/۲۸۵۹	-۰/۹۱۵۳	۰/۱۰۷۴	۰/۰۸۲۶	
سؤال ۹۱	۰/۴۹۵۱	۰/۵۰۴۹	۰/۰۱۹۵	۰/۲۹۸۶	۰/۲۷۳۱	
سؤال ۹۲	۰/۶۷۵	۰/۳۲۵	-۰/۷۳۱	۰/۳۴۶۸	۰/۳۲۳۷	
سؤال ۹۳	۰/۷۱۷۶	۰/۲۸۲۴	-۰/۹۳۲۵	۰/۲۷۱۵	۰/۲۴۸۲	
سؤال ۹۴	۰/۶۰۰۴	۰/۳۹۹۶	-۰/۴۰۷۲	۰/۳۲۳۸	۰/۲۹۹۲	
سؤال ۹۵	۰/۷۳۷۸	۰/۲۶۲۲	-۱/۰۳۴۵	۰/۱۹۱۱	۰/۱۶۷۵	
سؤال ۹۶	۰/۸۱۲۴	۰/۱۸۷۶	-۱/۴۶۵۸	-۰/۰۵۴۸	-۰/۰۷۶۳	
سؤال ۹۷	۰/۵۹۲۱	۰/۴۰۷۹	-۰/۳۷۲۴	۰/۳۱۷۲	۰/۲۹۲۵	
سؤال ۹۸	۰/۶۶۱۱	۰/۳۳۸۹	-۰/۶۶۸۱	۰/۱۶۹۴	۰/۱۴۳۸	
سؤال ۹۹	۰/۷۲۳۸	۰/۲۷۶۲	-۰/۹۶۳۶	-۰/۰۴۳۳	-۰/۰۰۶۸	
سؤال ۱۰۰	۰/۶۹۸	۰/۳۰۲	-۰/۸۳۸	۰/۰۲۴۴	-۰/۰۰۱	
سؤال ۱۰۱	۰/۷۰۸۵	۰/۲۹۱۵	-۰/۸۸۸۱	۰/۱۳۵۹	۰/۱۱۱	
سؤال ۱۰۲	۰/۷۳۲۹	۰/۲۶۷۱	-۱/۰۰۹۵	۰/۲۲۱۷	۰/۱۹۸۲	
سؤال ۱۰۳	۰/۱۹۶۷	۰/۸۰۳۳	۰/۴۰۷۳	۰/۴۹۷۷	۰/۴۸۰۹	
سؤال ۱۰۴	۰/۵۶۰۷	۰/۴۳۹۳	-۰/۲۴۳۹	۰/۳۱۰۱	۰/۲۸۴۹	
سؤال ۱۰۵	۰/۳۳۳۳	۰/۶۶۶۷	۰/۶۹۳۱	۰/۵۲۳۶	۰/۵۰۴۳	
سؤال ۱۰۶	۰/۷۵۰۳	۰/۲۴۹۷	-۱/۱۰۰۵	۰/۲۹۵۵	۰/۲۷۳۴	
سؤال ۱۰۷	۰/۷۷۸۹	۰/۲۲۱۱	-۱/۲۵۹۵	۰/۱۵۲۶	۰/۱۳	
سؤال ۱۰۸	۰/۷۴۱۳	۰/۲۵۸۷	-۱/۰۵۲۶	۰/۱۷۷۱	۰/۱۵۳۵	
سؤال ۱۰۹	۰/۶۲۲۷	۰/۳۷۷۳	-۰/۵۰۱۲	۰/۲۶۲۲	۰/۲۳۶۹	
سؤال ۱۱۰	۰/۷۹۷۱	۰/۲۰۲۹	-۱/۳۶۸۱	-۰/۰۱۶۳	-۰/۰۳۸۶	
سؤال ۱۱۱	۰/۴۸۶۵	۰/۵۰۳۵	۰/۰۱۳۹	۰/۴۴۳۷	۰/۴۲۱	
سؤال ۱۱۲	۰/۶۳۴۶	۰/۳۶۵۴	-۰/۵۵۲	۰/۱۸۱۸	۰/۱۵۵۹	
سؤال ۱۱۳	۰/۷۲۵۲	۰/۲۷۴۸	-۰/۹۷۰۶	۰/۳۱۰۳	۰/۲۸۷۷	
سؤال ۱۱۴	۰/۴۶۱۶	۰/۵۳۸۴	۰/۱۵۳۷	۰/۴۴۰۷	۰/۴۱۸	
سؤال ۱۱۵	۰/۶۳۱۱	۰/۳۶۸۹	-۰/۵۳۶۹	۰/۳۴۱۳	۰/۳۱۷۴	
سؤال ۱۱۶	۰/۷۶۹۲	۰/۲۳۰۸	-۱/۲۰۳۷	-۰/۰۰۰۱	-۰/۰۲۳۴	
سؤال ۱۱۷	۰/۳۶۶۱	۰/۶۳۳۹	۰/۵۴۸۹	۰/۵۰۴۹	۰/۴۸۴۷	

جدول ۱. پارامتر سؤالات آزمون مبتنی بر نظریه کلاسیک						
شاخص تمیز بر اساس همبستگی دو رشته ای نقطه‌ای			شاخص دشواری			
سؤال در نمره کل موجود نیست	سؤال در نمره کل موجود است	Logit	درصد پاسخ صحیح	درصد پاسخ غلط	شماره سؤال	
-۰/۰۵۷۱	-۰/۰۳۲	-۰/۸۸۴۸	۰/۲۹۲۲	۰/۷۰۷۸	سؤال ۱۱۸	
۰/۱۴۲۳	۰/۱۶۵۹	-۱/۰۸۱۹	۰/۲۵۳۱	۰/۷۴۶۹	سؤال ۱۱۹	
۰/۰۵۷۹	۰/۰۸۲۷	-۰/۹۴۲۸	۰/۲۸۰۳	۰/۷۱۹۷	سؤال ۱۲۰	
۰/۱۱۴۷	۰/۱۳۹۳	-۰/۹۰۸۵	۰/۲۸۷۳	۰/۷۱۲۷	سؤال ۱۲۱	
۰/۰۵۳۱	۰/۰۷۸۳	-۰/۸۶۴۶	۰/۲۹۶۴	۰/۷۰۳۶	سؤال ۱۲۲	
۰/۲۷۵۲	۰/۳۰۰۳	-۰/۳۷۵۳	۰/۴۰۷۳	۰/۵۹۲۷	سؤال ۱۲۳	
۰/۱۴۷۴	۰/۱۷۲۷	-۰/۷۵۰۲	۰/۳۲۰۸	۰/۶۷۹۲	سؤال ۱۲۴	
-۰/۲۷۰۴	۰/۲۹۳۴	-۰/۹۲۹	۰/۲۸۳۱	۰/۷۱۶۹	سؤال ۱۲۵	
۰/۴۵۱۱	۰/۴۷۰۷	۰/۹۶۳۶	۰/۷۲۳۸	۰/۲۷۶۲	سؤال ۱۲۶	
۰/۲۰۹۳	۰/۲۳۳۶	-۰/۸۲۸۱	۰/۳۰۴	۰/۶۹۶	سؤال ۱۲۷	
۰/۲۸۱۱	۰/۳۰۵۹	-۰/۴۱۰۱	۰/۳۹۸۹	۰/۶۰۱۱	سؤال ۱۲۸	
-۰/۰۲۷۸	۰/۰۵۲	-۱/۰۶۷۲	۰/۲۵۵۹	۰/۷۴۴۱	سؤال ۱۲۹	
۰/۰۸۸	۰/۱۱۱۹	-۱/۰۸۹۳	۰/۲۵۱۷	۰/۷۴۸۳	سؤال ۱۳۰	
۰/۱۵۸۲	۰/۱۸۳۵	-۰/۷۱۲	۰/۳۲۹۱	۰/۶۷۰۹	سؤال ۱۳۱	
۰/۳۹۰۱	۰/۴۱۳۵	-۰/۰۳۹۱	۰/۴۹۰۲	۰/۵۰۹۸	سؤال ۱۳۲	
-۰/۱۳۰۴	۰/۱۵۲۵	-۱/۳۴۲۴	۰/۲۰۷۱	۰/۷۹۲۹	سؤال ۱۳۳	
۰/۰۱۳	۰/۰۳۷۱	-۱/۰۷۴۶	۰/۲۵۴۵	۰/۷۴۵۵	سؤال ۱۳۴	
۰/۲۲۷۵	۰/۲۴۹۶	-۱/۱۹۹۷	۰/۲۳۱۵	۰/۷۶۸۵	سؤال ۱۳۵	
۰/۱۴۱۶	۰/۱۶۷	-۰/۷۲۷۹	۰/۳۲۵۷	۰/۶۷۴۳	سؤال ۱۳۶	
۰/۲۵۴۸	۰/۲۸	-۰/۴۳۹۳	۰/۳۹۱۹	۰/۶۰۸۱	سؤال ۱۳۷	
۰/۰۴۲۹	۰/۰۶۸۵	-۰/۸۰۱۹	۰/۳۰۹۶	۰/۶۹۰۴	سؤال ۱۳۸	
۰/۰۳۰۳	۰/۰۵۴۹	-۰/۹۸۴۷	۰/۲۷۲	۰/۷۲۸	سؤال ۱۳۹	
۰/۰۱۷۹	۰/۰۴۲	-۱/۰۷۸۳	۰/۲۵۳۸	۰/۷۴۶۲	سؤال ۱۴۰	

جدول ۲. پارامتر سؤالات آزمون مبتنی بر مدل تک بعد سؤال پاسخ			
شماره سؤالات	ضریب حدس	ضریب دشواری	ضریب تمیز
سؤال ۲۱	۰	-۱/۷۰۶	۲/۷۶۲
سؤال ۲۲	۰	-۱/۲۹۸	۱/۳۸۹
سؤال ۲۳	۰/۲۷۲	۰/۵۴۱	۱/۵۵۳
سؤال ۲۴	۰/۲۸۲	۰/۴۶۶	۱/۸۴۴
سؤال ۲۵	۰	-۱/۱۱۲	۰/۹۱۹
سؤال ۲۶	۰/۲۳۴	۱/۳۸۴	۲/۰۶۶
سؤال ۲۷	۰/۲۱۱	۰/۵۸۲	۱/۲۴۵
سؤال ۲۸	۰/۲۵۵	۰/۹۷۹	۰/۹۰۵
سؤال ۲۹	۰/۱۸۸	۲/۰۸۹	۳/۷۲۹
سؤال ۳۰	۰/۲۰۷	-۰/۵۴۱	۱/۵۷۲
سؤال ۳۱	۰/۲۱۵	-۰/۱۰۶	۰/۸۷۶
سؤال ۳۲	۰/۲۶۴	۰/۷۶۷	۱/۹۰۶
سؤال ۳۳	۰	۰/۵۷	۰/۴۸۱
سؤال ۳۴	۰/۲۴۱	۰/۸۱۷	۲/۲۴۴
سؤال ۳۵	۰/۲۶۱	-۱/۴۸۶	۲/۳۹۳
سؤال ۳۶	۰/۰۰۱	-۱/۵۵۲	۱/۰۰۷
سؤال ۳۷	۰/۲۹۷	۰/۱۶۱	۱/۸۳۳
سؤال ۳۸	۰/۲۸۷	-۱/۳۵۳	۵/۶۵۳
سؤال ۳۹	۰/۱۰۱	-۰/۴۹۴	۱/۱۵۲
سؤال ۴۰	۰/۱۹۹	۲/۶۶	۰/۶۴

جدول ۲. پارامتر سؤالات آزمون مبتنی بر مدل تک بعد سؤال پاسخ			
شماره سؤالات	ضریب حدس	ضریب دشواری	ضریب تمیز
سؤال ۱	۰/۰۰۲	-۳/۱۲۱	۰/۹۲۵
سؤال ۲	۰/۰۰۳	-۳/۰۱۷	۱/۱۹۷
سؤال ۳	۰	-۲/۳۸	۱/۲۵۵
سؤال ۴	۰/۱۰۱	۱/۵۶	۱/۶۵۱
سؤال ۵	۰/۰۰۱	-۲/۹۷۳	۱/۴۱۵
سؤال ۶	۰	-۲/۷۶۲	۱/۸۸۷
سؤال ۷	۰/۰۷۱	-۱/۰۲۸	۱/۲۴۲
سؤال ۸	۰	-۲/۱۶۴	۲/۳۰۷
سؤال ۹	۰	-۲/۲۷۱	۱/۷۳۱
سؤال ۱۰	۰	-۱/۹۱	۱/۶۱۶
سؤال ۱۱	۰/۳۱۲	۰/۰۷	۱/۶۰۷
سؤال ۱۲	۰	-۲/۱۳۳	۱/۱۷۹
سؤال ۱۳	۰	-۲/۰۴۸	۱/۴۷۶
سؤال ۱۴	۰/۰۰۱	-۱/۹۷۳	۰/۷۳۸
سؤال ۱۵	۰/۲۳۵	-۰/۰۰۷	۱/۶۳۱
سؤال ۱۶	۰/۲۱۱	۱/۹۴۱	۲/۳۴۸
سؤال ۱۷	۰/۴۶۷	۰/۸۲۱	۲/۵۳۶
سؤال ۱۸	۰/۱۹۱	-۱/۴۷۴	۱/۴۵۴
سؤال ۱۹	۰	-۲/۳۶۳	۱/۱۷۵
سؤال ۲۰	۰/۲۲۳	۱/۱۹۸	۱/۶۷۷

جدول ۲. پارامتر سؤالات آزمون مبتنی بر مدل تک بعد سؤال پاسخ

شماره سؤالات	ضریب حدس	ضریب دشواری	ضریب تمیز
سؤال ۸۷	۰/۲۷۱	-۳/۳۷۵	-۲/۱۸۵
سؤال ۸۸	۰/۱۱۹	-۰/۵۴۵	۳/۰۳
سؤال ۸۹	۰/۱۱۶	-۴/۶۴۶	-۰/۴۸۸
سؤال ۹۰	۰/۲۶۵	۲/۲۲۲	۳/۳۳
سؤال ۹۱	۰/۱۶۱	۰/۶۰۱	۰/۸۴
سؤال ۹۲	۰/۱۱۶	۱/۳۱۶	۱/۲۷۷
سؤال ۹۳	۰/۱۴۱	۱/۶۸۹	۱/۳۵۹
سؤال ۹۴	۰/۲۵۱	۱/۲۰۷	۲/۱۹۴
سؤال ۹۵	۰/۲۰۳	۱/۷۵	۳/۱۴۹
سؤال ۹۶	۰/۱۳	-۴/۵۲۵	-۰/۶۲۴
سؤال ۹۷	۰/۲۶۵	۱/۴۴۵	۱/۴۷۳
سؤال ۹۸	۰/۳۰۸	۲/۰۵۸	۲/۹۵۳
سؤال ۹۹	۰/۲۵۸	-۴/۰۸۷	-۱/۰۶
سؤال ۱۰۰	۰/۰۰۳	-۱۳/۳۸۶	-۰/۰۶۳
سؤال ۱۰۱	۰/۲۷۲	۲/۱۷۴	۴/۱۱۲
سؤال ۱۰۲	۰/۱۸۹	۲/۰۶۲	۱/۵۷۵
سؤال ۱۰۳	۰/۱۱۲	-۰/۹۷۴	۱/۹۵۱
سؤال ۱۰۴	۰/۳۳۷	۱/۴۰۴	۲/۲۶۹
سؤال ۱۰۵	۰/۱۲۳	-۰/۲۸	۲/۰۶۶
سؤال ۱۰۶	۰/۱۱۵	۱/۷۰۴	۱/۴۶۱
سؤال ۱۰۷	۰/۱۷۸	۲/۳۶۷	۱/۸۱۶
سؤال ۱۰۸	۰/۰۶۴	۲/۷۶۶	۰/۵۳
سؤال ۱۰۹	۰/۲۵۹	۱/۴۷۳	۱/۸۸۴
سؤال ۱۱۰	۰/۱۹	-۳/۵۵۵	-۱/۶۳۳
سؤال ۱۱۱	۰/۱۸۲	۰/۵۲۵	۱/۶۰۷
سؤال ۱۱۲	۰/۲۵۳	۲/۱۶۱	۰/۹۹۶
سؤال ۱۱۳	۰/۱۰۵	۱/۵۹۹	۱/۲۵۷
سؤال ۱۱۴	۰/۲۲۲	۰/۴۶۸	۱/۶۶۹
سؤال ۱۱۵	۰/۰۹۷	۱/۰۹۴	۱/۰۵۱
سؤال ۱۱۶	۰/۱۶۶	-۴/۴۵۷	-۰/۵۸۷
سؤال ۱۱۷	۰/۰۷۳	-۰/۲۷۶	۱/۶۲۷
سؤال ۱۱۸	۰/۲۶۶	-۴/۷۷۴	-۰/۷۵۸
سؤال ۱۱۹	۰/۱۸۴	۲/۱۹۲	۱/۵۳۵
سؤال ۱۲۰	۰/۲۷	۳/۱۷۵	۱/۸۴۵
سؤال ۱۲۱	۰/۲۴۸	۲/۳۲۷	۱/۸۹۶
سؤال ۱۲۲	۰/۲۷۶	۲/۵۵	۲/۲۵
سؤال ۱۲۳	۰/۲۲	۱/۲۷۹	۱/۳۱
سؤال ۱۲۴	۰/۲۵۶	۲/۰۹۴	۱/۶۴۹
سؤال ۱۲۵	۰/۱۳۴	۱/۶۳۲	۱/۳۷۳
سؤال ۱۲۶	۰/۲۱۶	-۰/۴۰۱	۱/۹۱۷
سؤال ۱۲۷	۰/۰۷۹	۲/۱۲۵	۰/۵۹۹
سؤال ۱۲۸	۰/۲۲۳	۱/۲۶۲	۱/۵۰۱
سؤال ۱۲۹	۰/۱۲۴	-۲۱/۷۶۹	-۰/۰۷۹
سؤال ۱۳۰	۰/۲۳۲	۲/۳۲۳	۳/۱۲۲
سؤال ۱۳۱	۰/۱۷۹	۲/۳۴۱	۰/۷۴۴
سؤال ۱۳۲	۰/۱۹۵	۰/۶۱۶	۱/۶۹۸

جدول ۲. پارامتر سؤالات آزمون مبتنی بر مدل تک بعد سؤال پاسخ

شماره سؤالات	ضریب حدس	ضریب دشواری	ضریب تمیز
سؤال ۴۱	۰/۲۹۳	۲/۶۸۱	۱/۵۱۹
سؤال ۴۲	۰/۲۱۳	۰/۸۶۶	۳/۴۳۹
سؤال ۴۳	۰/۳۱	۰/۷۸۲	۳/۸۱۳
سؤال ۴۴	۰/۱۵۵	۰/۳۷۵	۱/۸۳۳
سؤال ۴۵	۰/۱۶۲	۱/۷۰۸	۱/۰۰۸
سؤال ۴۶	۰/۲۳	-۰/۰۷۷	۲/۱۹۵
سؤال ۴۷	۰/۱۸	-۰/۴۶	۱/۳۴۵
سؤال ۴۸	۰/۲۵۴	۰/۶۳۵	۲/۳۳۴
سؤال ۴۹	۰/۱۵۲	۰/۸	۰/۷۷
سؤال ۵۰	۰/۲۶۶	-۰/۷۳	۲/۶۹۸
سؤال ۵۱	۰/۱۹۶	۱/۴۰۱	۰/۹۳۲
سؤال ۵۲	۰/۰۰۱	۰/۲۱۱	۰/۸۲۶
سؤال ۵۳	۰/۲۷۲	۱/۲۲۹	۲/۷۲۱
سؤال ۵۴	۰/۱۲۴	-۰/۲۱۵	۱/۶۷۲
سؤال ۵۵	۰/۳۱۹	۱/۲۸۸	۱/۳۵۸
سؤال ۵۶	۰/۱۸۸	-۰/۶۱	۲/۰۴۹
سؤال ۵۷	۰/۲۶۸	۱/۴۶۶	۳/۵۱۹
سؤال ۵۸	۰/۰۰۷	-۴/۳۰۶	-۰/۲۵۳
سؤال ۵۹	۰/۲۴۷	۱/۱۲۴	۱/۵۱۵
سؤال ۶۰	۰/۲۹۵	۲/۲۹۱	۴/۶۶
سؤال ۶۱	۰/۱۳۱	-۰/۸۵۹	۱/۹۵۸
سؤال ۶۲	۰/۲۴۳	۲/۰۰۶	۶/۸۹۶
سؤال ۶۳	۰/۲۷۷	۱/۳۰۵	۰/۰۷۹
سؤال ۶۴	۰/۲۴۷	۱/۸۳	۳/۹۰۹
سؤال ۶۵	۰/۱۷۴	۲/۰۰۴	۳/۸۸۵
سؤال ۶۶	۰/۱۸۲	۱/۸۲۲	۰/۹۴۸
سؤال ۶۷	۰/۱۳۲	۱/۶۹۱	۰/۷۵۱
سؤال ۶۸	۰/۲	-۴/۰۴	-۱/۰۷۵
سؤال ۶۹	۰/۰۱۷	-۰/۰۲	۰/۵۱۷
سؤال ۷۰	۰/۲۶	-۰/۷۵۳	۲/۳۹۴
سؤال ۷۱	۰/۰۷۹	۰/۹۲۱	۰/۹۱۲
سؤال ۷۲	۰/۱۵۲	۰/۵۴۸	۱/۱۳۲
سؤال ۷۳	۰/۲۳۳	۰/۱۱۲	۱/۴۴۸
سؤال ۷۴	۰/۲۵۵	-۰/۷۶۴	۲/۴۶۱
سؤال ۷۵	۰/۱۶۲	۰/۱۵	۲/۴۲۲
سؤال ۷۶	۰/۲۰۷	۱/۳۷۹	۱/۶۲۴
سؤال ۷۷	۰/۲۷۸	۱/۹۴۵	۰/۷۷
سؤال ۷۸	۰/۲۲۷	۰/۷۵	۱/۵۸۸
سؤال ۷۹	۰/۲۷۹	۰/۸۱۶	۱/۶۰۴
سؤال ۸۰	۰/۱۶۱	۰/۴۲۹	۱/۹۱۴
سؤال ۸۱	۰/۱۹۱	۲/۹۳۸	۰/۷۸۴
سؤال ۸۲	۰/۲۳۹	۱/۴۴۱	۲/۴۸۱
سؤال ۸۳	۰/۱۹۷	۰/۴۰۵	۱/۵۸۲
سؤال ۸۴	۰/۱۳۲	۲/۳۷۷	۰/۹۹۶
سؤال ۸۵	۰/۳۳۱	۱/۴۱۷	۲/۳۷۹
سؤال ۸۶	۰/۰۶۱	۷/۳۴۴	۰/۱۷۹

برای ترسیم شبکه بالا، 114×113 تقسیم بر ۲ یعنی ۶۴۴۱ آزمون آماری همبستگی تفکیکی با تعدیل آلفا انجام شده است که از این میان تنها یال‌هایی بین دو گره در شکل ترسیم شده‌اند که نشان‌دهنده همبستگی‌های تفکیکی معنادار در سطح ۹۹ درصد اطمینان می‌باشند. گرافی که به منظور مطالعه ساختار سؤالات پرسشنامه ترسیم می‌شود دارای ویژگی‌های مختلفی است: الف) ضخامت یال‌ها (منظور خطوط ترسیم شده بین دو گره) نشان‌دهنده ارتباط معنادار بین دو سؤال بر اساس همبستگی است. اگر بین دو متغیر که در ادبیات شبکه به آن گره گفته می‌شود یک خط (یال) رسم شده باشد، نشان می‌دهد که بین این دو ارتباط وجود دارد، عدم وجود خط رابط بین دو متغیر نشان‌دهنده فقدان ارتباط معنادار بین آنها است. ب) هر چه این ارتباط بیشتر باشد، یال بین دو نشانه قطورتر می‌شود؛ همانطور که گفته شد، وجود خط رابط به معنای ارتباط معنادار است که هر چه قطر یال نازک‌تر باشد، شدت ارتباط معنادار کمتر است. ج) رنگ یال‌ها جهت ارتباط بین متغیرها را نشان می‌دهد. یال‌های سبز رنگ نشان‌دهنده ارتباط مثبت معنادار بین گره‌ها است و رنگ قرمز یال، به معنای ارتباط منفی معنادار است. د) گره‌های مرکزی نیز گره‌هایی هستند که در مرکز گراف جای گرفته‌اند و گره‌های پیرامونی گره‌هایی می‌باشند که در اطراف قرار می‌گیرند. الگوریتم استفاده شده برای ترسیم شبکه مورد نظر گره‌های مؤثرتر را در مرکز شبکه قرار می‌دهد. و) همسایگی گره‌ها به یکدیگر نشان‌دهنده ارتباط بالا بین آنها است. به عبارتی فاصله اقلیدسی در گراف بالا به معنای ارتباط بیشتر مفهومی بین آنها است. از این ویژگی گراف گاهی با هدف بعدیابی نیز استفاده می‌شود (عراقی و همکاران، ۱۳۹۶).

بر اساس این شبکه، شاخص‌های مرکزیت شبکه بدست آمد که در جدول ۳ درج شده است. بر این اساس سؤالات ۱۱۸، ۳۹، ۲۴، ۹۸، ۶۹، ۴۹، ۲۲، ۱۳۲، ۱۳۳ و ۱۲ به ترتیب دارای بیشترین مقدار بینیت می‌باشند. سؤالات ۱۱۸، ۶۹، ۳۷ و ۹۸، ۴۹، ۱۳۳، ۱۱۳، ۳۹، ۶۷ و ۴۱، جزء ده سؤال با مقدار نزدیکی بالا و سؤالات ۱۱۸، ۴۹، ۱۱۵، ۹۰، ۱۳۱، ۲۲، ۱۱، ۳۷، ۴۳ و ۴۱ به ترتیب ده سؤالی هستند که بیشترین مقدار استحکام را دارند. در شاخص تأثیر مورد انتظار نیز سؤالات ۹۵، ۱۰۹، ۹۸، ۱۰۶، ۵۶، ۲۱، ۵۰، ۲۴، ۵۷ و ۳۸ سؤالاتی هستند که بیشترین تأثیرگذاری مورد انتظار را بر شبکه مورد مطالعه دارند.

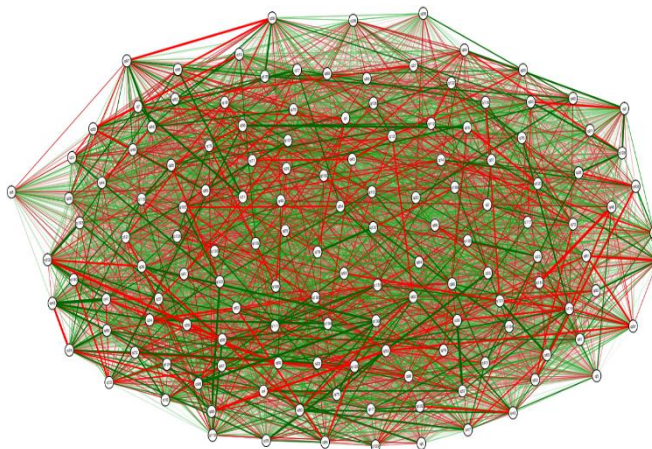
جدول ۲. پارامتر سؤالات آزمون مبتنی بر مدل تک بعد سؤال پاسخ

شماره سؤالات	ضریب حدس	ضریب دشواری	ضریب تمیز
سؤال ۱۳۳	۰/۱۸۴	۲/۱۷۴	۳/۷۰۱
سؤال ۱۳۴	۰/۱۶۴	-۱۰/۸۶	-۰/۱۹۵
سؤال ۱۳۵	۰/۱۳۶	۱/۹۸۲	۱/۵۰۸
سؤال ۱۳۶	۰/۲۶۹	۱/۸۹۵	۲/۲۹۷
سؤال ۱۳۷	۰/۳۱۲	۱/۶۲۱	۲/۲۷۳
سؤال ۱۳۸	۰/۰۷۶	۹/۲۱۴	۰/۱۱۹
سؤال ۱۳۹	۰/۲۵۴	۲/۹۰۱	۱/۸۳۳
سؤال ۱۴۰	۰/۲۱۶	-۶/۹۱۵	-۰/۴۴۳

سؤالات دارای ضریب تمیز منفی به صورت تصادفی پاسخ داده شده‌اند و نه براساس سطح توانایی زبانی شرکت‌کنندگان. به طور کلی سؤالات انتهایی به دلیل دشوار بودن مقدار حدس بیشتری دارند.

ورودی تحلیل‌های شبکه ماتریس مجاورت است. این ماتریس به عنوان ورودی، نقش تعیین‌کننده‌ای را در تعیین حدود شبکه دارد. از آنجا که هدف نهایی این پژوهش کشف رابطه خالص بین سؤالات برای استخراج الگوی آزمون تطبیقی است، ماتریس مجاورت براساس همبستگی تفکیکی بین سؤالات استخراج شده است. چون تمام سؤالات دو ارزشی است، همبستگی بین سؤالات براساس ضریب تتراکوریک است و ارتباط خالص بین آنها بر اساس همبستگی تتراکوریک تفکیکی انجام شد.

نقشه ارتباط بین سؤالات براساس الگوریتم فراچترمن - رینگولد و توسط بسته نرم‌افزاری qgraph تحت نرم‌افزار R ترسیم شد. شبکه ارتباط خالص بین سؤالات در زیر نشان داده شده است.



شکل ۱. شبکه ارتباط خالص (همبستگی تفکیکی تتراکوریک) معنادار بین سؤالات آزمون حجم لغات

جدول ۳. شاخص‌های مرکزیت سؤالات در شبکه سؤالات آزمون حجم لغات				
شماره سؤالات	بینیت	نزدیکی	استحکام	تأثیر مورد انتظار
۱۳۶	۶۲	۰/۰۰۱۸۰۳۱۰۵	۲۷/۷۰۰۳۴	۱/۳۰۷۹۳۹۳
۱۳۷	۴۵	۰/۰۰۱۸۴۸۰۴۷	۲۹/۴۵۰۰۴	۲/۶۹۰۱۱۸۸۲
۱۳۸	۴۴	۰/۰۰۱۸۰۹۷۰۸	۲۷/۵۹۰۷۳	-۰/۵۴۱۷۲۹۲
۱۳۹	۵۲	۰/۰۰۱۸۶۲۲۸۳	۲۸/۶۳۲۸۱	-۱/۷۵۲۲۲۷۷
۱۴۰	۴۰	۰/۰۰۱۸۰۱۷۶۷	۲۷/۶۸۷۶۷	-۰/۴۸۶۳۳۳۷

این شاخص‌ها با هدف بهبود انتخاب سؤال اول در الگوریتم انطباقی بدست آمده‌اند. از بین این شاخص‌ها، شاخص نزدیکی بیشترین انطباق را با هدف این پژوهش جهت بهبود انتخاب سؤال اول دارد. شاخص نزدیکی بیانگر این است که یک گره چگونه به صورت غیرمستقیم با سایر گره‌ها در ارتباط است. شاخص نزدیکی نشان می‌دهد که به طور متوسط کدام سؤالات نزدیکی بیشتری به کل سؤالات دارند. یعنی مجموع فاصله کدام سؤالات از همه سؤالات کمتر است و به لحاظ ریاضیاتی، میانگین کوتاه‌ترین مسیر هر سؤال از سایر سؤالات است. بنابراین از شاخص‌های مرکزیت شبکه، شاخص نزدیکی با هدف انتخاب سؤال اول برگزیده می‌شود. بر این اساس الگوریتم پیشنهادی دارای گام‌های زیر می‌باشد:

۱. انتخاب سؤال نخست: مهمترین بخش در بهبود الگوریتم‌های انطباقی مبتنی بر نظریه سؤال - پاسخ که در این پژوهش بر آن اهتمام ورزیده شده است، بهبود انتخاب سؤال اول است. از آنجایی که فرض بر این است که اطلاعات پیشین از آزمون‌شوندگان در اختیار نیست، بهترین نقطه شروع بر اساس شاخص‌های شبکه، سؤالی است که مقدار نزدیکی بیشتری دارد. همان‌طور که ذکر شد نزدیکی به لحاظ ریاضیاتی، میانگین کوتاه‌ترین مسیر هر سؤال از سایر سؤالات است. بنابراین سؤالات با نزدیکی بالا می‌توانند به عنوان معرف تمام سؤالات لحاظ شوند. بنابراین در گام اول به دلیل عدم اطلاع از توانایی افراد شرکت‌کننده و دشواری متفاوت، یک سؤال به صورت تصادفی از بین ده سؤال دارای بیشترین مقدار شاخص نزدیکی انتخاب می‌شود. سؤالات ۱۱۸، ۶۹، ۳۷ و ۹۸، ۴۹، ۱۳۳، ۱۱۳، ۳۹، ۶۷ و ۴۱، جزء ده سؤال با مقدار نزدیکی بالا می‌باشند.

۲. برای ارائه سؤال بعد، اگر فرد به سؤال غلط پاسخ داده باشد، سؤالی انتخاب می‌شود که در وهله نخست با سؤال قبلی در شبکه سؤالات در ارتباط باشد. از بین تمام سؤالاتی که با سؤال انتخاب شده قبلی در ارتباط است سؤالی مورد نظر است که به طور همزمان

جدول ۳. شاخص‌های مرکزیت سؤالات در شبکه سؤالات آزمون حجم لغات				
شماره سؤالات	بینیت	نزدیکی	استحکام	تأثیر مورد انتظار
۹۱	۲۶	۰/۰۰۱۷۵۴۳۱۹	۲۶/۸۶۰۲۲	-۱/۶۷۶۸۶۸۲
۹۲	۳۶	۰/۰۰۱۷۳۸۸۲۶	۲۶/۹۰۰۱۲	۰/۷۱۷۶۸۴۴۴
۹۳	۳۴	۰/۰۰۱۸۲۹۶۲۷	۲۸/۶۱۲۵۷	-۰/۷۳۰۷۴۱۱
۹۴	۵۴	۰/۰۰۱۷۷۲۲۸۳	۲۵/۴۷۱۴۴	-۲/۳۱۵۷۶۰۱
۹۵	۳۳	۰/۰۰۱۸۱۶۰۴۲	۲۸/۰۳۸۴۷	۴/۸۲۷۴۹۲۶۲
۹۶	۳۹	۰/۰۰۱۸۱۴۶۹۴	۲۷/۵۵۰۹۹	۲/۰۹۳۷۴۶۵۹
۹۷	۴۶	۰/۰۰۱۸۰۸۵۲۳	۲۶/۸۱۰۲۵	۲/۰۶۱۵۹۳۱۷
۹۸	۸۳	۰/۰۰۱۹۰۲۱۹۵	۲۸/۲۰۱۴۹	۴/۴۵۹۹۳۸۶۶
۹۹	۳۸	۰/۰۰۱۷۹۱۷۶۹	۲۸/۳۸۰۰۷	-۱/۴۲۴۰۷۶۲
۱۰۰	۵۵	۰/۰۰۱۸۵۰۸۷۴	۲۹/۳۲۶۸۱	-۱/۷۸۱۵۰۲۵
۱۰۱	۳۷	۰/۰۰۱۷۹۱۳۹	۲۸/۴۱۲۷۵	۰/۱۴۹۳۹۲۹۴
۱۰۲	۴۶	۰/۰۰۱۸۲۶۵	۲۸/۰۹۷۸۷	۱/۱۶۶۷۷۸۶۷
۱۰۳	۳۹	۰/۰۰۱۷۸۱۵۰۲	۲۷/۶۰۹۱۹	-۰/۲۰۵۱۶۷۷
۱۰۴	۵۸	۰/۰۰۱۸۴۶۵۷۵	۲۸/۹۴۰۱۴	۲/۳۳۸۹۸۵۴۲
۱۰۵	۶۷	۰/۰۰۱۸۵۶۵۹۷	۲۹/۳۸۱۴۳	۰/۱۰۱۴۹۳۵۶
۱۰۶	۲۵	۰/۰۰۱۷۶۳۷۹	۲۷/۰۸۰۶۹	۴/۳۲۶۴۴۱۸۷
۱۰۷	۳۰	۰/۰۰۱۸۰۲۳۴۲	۲۸/۲۱۰۱۶	۲/۴۲۲۲۵۵۶۱
۱۰۸	۴۸	۰/۰۰۱۷۷۲۰۸۹	۲۶/۶۳۸۹۱	-۱/۵۹۵۳۳۵۲
۱۰۹	۵۷	۰/۰۰۱۷۹۵۲۱۶	۲۷/۶۹۹۴۸	۴/۴۸۲۰۹۲۳۶
۱۱۰	۶۰	۰/۰۰۱۸۲۶۷۰۳	۲۶/۴۹۱۴۴	-۰/۰۱۱۷۰۳۴
۱۱۱	۳۸	۰/۰۰۱۸۱۶۴۹۲	۲۶/۸۵۱۰۴	۰/۷۷۴۵۹۵۷۴
۱۱۲	۴۳	۰/۰۰۱۷۸۳۱۴۵	۲۶/۵۳۸۰۲	-۱/۳۷۹۷۶۲۷
۱۱۳	۵۱	۰/۰۰۱۸۸۰۱۱۱	۲۹/۲۹۵۶۶	-۰/۱۷۸۵۲۲۲
۱۱۴	۴۵	۰/۰۰۱۸۳۸۳۴۲	۲۸/۲۲۴۵۹	۰/۰۸۵۳۶۲۲۶
۱۱۵	۶۰	۰/۰۰۱۸۴۸۹۸۸	۳۰/۴۰۷۳۱	-۱/۹۹۲۵۶۲۵
۱۱۶	۳۶	۰/۰۰۱۷۹۲۳۵	۲۷/۴۷۰۵۱	۱/۸۱۲۹۵۷۹۲
۱۱۷	۳۸	۰/۰۰۱۷۹۸۴۷۲	۲۷/۱۳۳۴۲	-۱/۰۰۶۱۱۳۴
۱۱۸	۱۰۳	۰/۰۰۱۹۳۵۴۵۹	۳۰/۷۲۶۳۶	-۱/۱۳۱۶۴۹۸
۱۱۹	۴۶	۰/۰۰۱۷۸۰۱۷۹	۲۸/۱۴۶۵۵	۳/۳۶۳۰۷۸۳۳
۱۲۰	۳۴	۰/۰۰۱۷۴۷۱۴۲	۲۵/۲۱۴۶۶	-۰/۱۵۶۲۶۹۹
۱۲۱	۳۹	۰/۰۰۱۸۶۱۴۰۴	۲۸/۲۵۱۱۳	۱/۳۶۵۷۷۱۴۴
۱۲۲	۳۳	۰/۰۰۱۷۵۱۵۴	۲۷/۵۴۷۵۵	-۰/۳۰۰۰۶۹۵
۱۲۳	۵۲	۰/۰۰۱۸۴۳۰۵	۲۸/۱۹۰۶۲	۲/۳۵۷۲۶۵۴
۱۲۴	۲۴	۰/۰۰۱۷۱۴۳۲۷	۲۶/۲۱۲۱۹	۲/۶۱۰۷۲۴۷۳
۱۲۵	۳۹	۰/۰۰۱۷۹۸۵۲۶	۲۹/۴۹۷۰۵	-۲/۰۳۲۱۰۳۷
۱۲۶	۶۰	۰/۰۰۱۸۵۴۲۹۴	۲۷/۸۳۲۹۲	۱/۸۹۲۲۲۴۰۸
۱۲۷	۵۱	۰/۰۰۱۸۳۴۴۴۴	۲۷/۸۶۶۴۹	-۰/۲۱۵۹۶۹۴
۱۲۸	۲۹	۰/۰۰۱۷۳۲۳۲۸	۲۵/۵۴۷۳۶	۲/۷۷۸۵۲۶۱۱
۱۲۹	۳۷	۰/۰۰۱۷۹۸۵۴۵	۲۷/۳۱۱۷۲	-۰/۹۵۵۷۴۵۹
۱۳۰	۲۹	۰/۰۰۱۸۴۲۸۰۳	۲۹/۲۲۸۶۵	۰/۱۳۸۷۸۸۷۹
۱۳۱	۵۹	۰/۰۰۱۸۶۶۹۱۳	۳۰/۳۲۵۸۶	-۰/۸۱۷۲۷۴
۱۳۲	۷۶	۰/۰۰۱۸۱۶۷۶۵	۲۸/۴۳۷۷۱	-۲/۱۷۶۱۵۵۸
۱۳۳	۷۵	۰/۰۰۱۸۸۲۵۵۶	۲۹/۵۴۴۵۷	۱/۳۱۴۳۰۴۸۲
۱۳۴	۳۲	۰/۰۰۱۷۵۲۰۱۶	۲۷/۱۱۷۰۳	۰/۷۵۵۰۸۶۸۴
۱۳۵	۵۲	۰/۰۰۱۷۷۹۶	۲۷/۵۹۵۸۸	۱/۸۴۶۶۸۱۹

فرد بعد از چهار سؤال درست، یک سؤال را غلط پاسخ می‌دهد. بنابراین در حال حاضر دو مقدار، تنای اولیه و الگوی پاسخ با حداقل واریانس (با اولین پاسخ غیر همسان) داریم. روش نیوتن - رافسون به منظور اصلاح توانایی اولیه استفاده می‌شود. در این روش مقدار تابع در مرحله $t+1$ برابر است با مقدار تابع در مرحله t منهای مشتق اول تابع تقسیم بر مشتق دوم تابع:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

۶. بعد از بدست آوردن مقدار تنای اصلاح شده $[\theta_j]_{t+1}$ ، لازم است این مقدار در معادله یک جایگزین $[\theta_j]$ شود. و مجدداً مقدار معادله محاسبه شود. این قدر این کار انجام شود که اختلاف بین $[\theta_j]_{t+1}$ از $[\theta_j]_t$ از یک مقدار مشخصی مانند $0/001$ کمتر شود. بعد از دستیابی به این نقطه برش، $[\theta_j]_{t+1}$ نهایی به عنوان توانایی فرد در نظر گرفته می‌شود.

۷. انتخاب سؤال بعد: انتخاب سؤال بعد بر اساس روشی که در گام دوم ذکر شده است، صورت می‌پذیرد.

۸. بعد از مشخص شدن سؤال بعدی، سؤال برای فرد ارائه می‌شود، فرد یا پاسخ صحیح می‌دهد یا پاسخ غلط. در هر صورت، لازم است مجدداً الگوی پاسخ فرد مشخص شود. به عنوان مثال فردی که الگوی پاسخ (۱۱۱۱۰) را قبل از مرحله هفتم داشته است، در صورتی که سؤال بعدی را درست پاسخ دهد، الگوی (۱۱۱۱۰۱) را خواهد داشت و در صورتی که غلط پاسخ دهد، الگوی (۱۱۱۱۰۰) را دارا است. بر اساس چنین الگوی پاسخی و بر اساس معادلات مرحله ۵، مجدداً توانایی فرد محاسبه می‌شود.

۹. بعد از محاسبه مجدد توانایی فرد، نیاز است که اختلاف بین توانایی‌ها محاسبه شود. در صورتی که این اختلاف کمتر از $0/001$ باشد، فرآیند متوقف شده و تنای مرحله $t+1$ به عنوان تنای نهایی در نظر گرفته می‌شود در غیر این صورت الگوریتم وارد مرحله هشتم می‌شود.

بحث و نتیجه‌گیری

هدف از اجرای این پژوهش ارائه الگوی آزمون انطباقی بر اساس علم شبکه بود. همسو با تحقیقات کیونگ و یانگ (۲۰۱۲)،

شدت یال و درجه دشواری آن کمتر باشد. شدت یا قدرت یال‌ها مقدار همبستگی تفکیکی بین دو به دو سؤالات است. درجه دشواری سؤالات قبلاً تحت نظریه سؤال - پاسخ برآورد شده‌اند. به عنوان مثال اگر اولین سؤال ۶۷ باشد، سؤالات ۸۵ و ۱۳۴ که رابطه خالص بیشتری با این سؤال دارند، انتخاب می‌شوند. درجه دشواری سؤال ۶۷ برابر $1/691$ ، سؤال ۸۵ برابر $1/417$ و سؤال ۱۳۴ برابر $1/86$ می‌باشد. از آنجایی که هر دو سؤال درجه دشواری کمتر از سؤال ۶۷ دارند، سؤالی برای ارائه انتخاب می‌شود که دارای مقدار قدرت یا شدت بیشتر یال باشد. در این مثال بعد از غلط زدن سؤال ۶۷ سؤال ۱۳۴ با شدت یال $0/1$ انتخاب می‌شود.

اگر آزمودنی به اولین سؤال صحیح پاسخ دهد، سؤالی انتخاب می‌شود که اولاً قدرت یال آن بیشتر باشد و ثانیاً درجه دشواری بیشتری نیز داشته باشد. در مثال بالا اگر به سؤال ۶۷ صحیح پاسخ داده شود، چون سؤالی با درجه دشواری بیشتر با این سؤال در ارتباط نیست، سؤالی به عنوان سؤال دوم انتخاب می‌شود که فقط مقدار درجه دشواری آن بیشتر است به عبارتی اولویت قدرت یا شدت یال جای خود را به اولویت درجه دشواری بیشتر می‌دهد.

۳. فرآیند مطرح شده در گام دوم تا جایی ادامه می‌یابد که اولین الگوی دارای واریانس ظاهر شود. به عبارتی به الگوی پاسخ آزمودنی‌ها در برگزیده حداقل یک سؤال درست یا یک سؤال غلط نیاز است، تا امکان برآورد توانایی تحت نظریه سؤال - پاسخ میسر شود. به عنوان مثال به الگوی پاسخ ۱۱۱۱۱۱۰ یا ۰۰۰۱ برسیم.

۴. بعد از دستیابی به الگوی پاسخ با حداقل واریانس، یک مقدار تنای (مبتنی بر نظریه سؤال - پاسخ) اولیه برای فرد لازم است که در نظر گرفته شود و این مقدار برابر $0/851$ در نظر گرفته می‌شود. این مقدار میانگین توانایی افراد شرکت‌کننده در این پژوهش است که تحت مدل سه پارامتری سؤال - پاسخ بدست آمده است.

۵. براساس مقدار پارامتر توانایی اولیه فرد یعنی $0/851$ و الگوی دارای واریانس پاسخ شرکت‌کننده، مقدار توانایی برآورد شده وی تحت نظریه سؤال-پاسخ و روش نیوتن - رافسون اصلاح می‌شود. منظور از الگوی پاسخ ترتیب مقادیر صفر و یک است که صفر به پاسخ‌های غلط و یک به پاسخ‌های درست داده می‌شود. مثلاً الگوی (۰۰۰۰۰۰۱) یا (۱۱۱۱۰) که در الگوی پاسخ اول، فرد بعد از شش سؤال غلط به یک سؤال درست پاسخ داده و در الگوی دوم

در این باره برای آزمون انطباقی ضرورت دارد. فقدان این اطلاعات در ادبیات آزمون‌های انطباقی سنتی سبب می‌شود تا سنجش انطباقی بر اساس ساده‌ترین سؤال صورت پذیرد (هان، ۲۰۱۲). به بیان دیگر در الگوهای سنتی انطباقی، اطلاعات پیشین منفک از فرآیند آزمون انطباقی می‌باشند، یک بخش کاملاً جدا از بخش‌ها و گام‌های دیگر آزمون انطباقی. اطلاعات پیشین در حوزه آزمون‌های پیشرفت تحصیلی مثلاً می‌تواند بر اساس نمرات درسی قبلی، معدل آخرین ترم تحصیلی، نظرات معلمان یا اساتید و... بدست آید. مسئله بعدی ارتباط برقرار کردن بین اطلاعات پیشین و انتخاب سؤالات در آزمون انطباقی است. به عنوان مثال زمانی که اطلاعات پیشین سابقه پیشرفت تحصیلی دانش‌آموز در قالب نمره معدل است، لازم است این نمره به یک سؤال پیوند بخورد و این سؤال به عنوان سؤال اول برای آن سطح از توانایی که معدل سمبلی از آن است، انتخاب شود. در سنجش‌های انطباقی سنتی که بعضی از آنها ذکر شد، بدون توجه به اینکه این اطلاعات کمی یا کیفی هستند لازم است نهایتاً در انتخاب اولین سؤال دخیل باشند؛ که نیازمند تصمیم‌گیری در هم‌سازی اطلاعات پیشین با ویژگی‌های سؤالات است که در اکثر مواقع این هم‌سازی از اعتبار بالایی برخوردار نیست. در شیوه‌های سنتی ذکر شده، برای موقعیت‌های پیچیده‌تر راهکار عملی وجود ندارد و بیشتر جنبه نظری دارند. به عنوان مثال در سنجش هرمی که به سنجش‌های انطباقی متوالی شهرت دارند و به عنوان روش‌های انطباقی کامل قبل از ظهور نظریه سؤال - پاسخ شناخته می‌شود، زمانی که الگوهای پاسخ غیر منتظره‌اند، کارکرد خود را از دست می‌دهند (گوماخ، ۲۰۱۲). علاوه بر این در الگوی هرمی به عنوان کامل‌ترین الگوی سنتی انطباقی، انتخاب سؤال در هر مرحله تنها وابسته به انتخاب صحیح یا غلط آزمودنی است؛ اگر آزمودنی صحیح پاسخ دهد سؤال دشوارتر و در صورتی که غلط پاسخ دهد سؤال ساده‌تر به فرد داده می‌شود یعنی تنها ملاک برای انتخاب سؤال، پارامتر دشواری سؤال است. این پارامتر در مدل‌های سنتی انطباقی بر اساس روش‌های کلاسیک اندازه‌گیری تعریف می‌شود. نکته قابل توجه در الگوی انطباقی این پژوهش با نظریه سؤال - پاسخ، فقدان تکیه بر انتخاب اولین سؤال براساس توانایی آزمون شونده است. به عبارتی در انتخاب اولین سؤال زمانی که اطلاعات از آزمون شونده در دست نیست، بهترین استراتژی تکیه بر قوت خود گره‌ها (در اینجا سؤالات) است (بینگ و هووانگ، ۲۰۱۰) که

دورنوی (۲۰۱۴)، پلاجتر و ولل (۲۰۱۵)، کالبرستون (۲۰۱۶)، اسپکمپ و همکاران (۲۰۱۲) و مارکات و پنمان (۲۰۱۹) نتایج پژوهش حاضر حاکی از آن است که تحلیل داده‌های شبکه‌ای به عنوان یک حوزه نوین در اندازه‌گیری روانشناختی، در روش قابلیت‌هایی دارد که می‌تواند در شخصی‌سازی آزمون استفاده شود به نحوی که با استفاده از تحلیل داده‌های شبکه، می‌توان توانایی آزمودنی‌ها را بطور مؤثر و کارآمد ارزیابی کرد و دقت و سرعت در برآورد توانایی را بهبود بخشید. همچنین مطابق با پژوهش عراقی و همکاران (۱۳۹۶) نتایج مطالعه حاضر نیز حاکی از آن است که می‌توان از تکنیک تحلیل داده‌های شبکه‌ای به خوبی در حوزه تحلیل سؤالات چه بصورت مستقل و چه در کنار شیوه‌های مرسوم تحلیل سؤال استفاده نمود. مزایای کاربرد چنین تکنیکی علاوه بر سادگی فهم، تفسیر و شناسایی سؤالات مناسب و نامناسب، نقشه ارتباطی بین سؤالات است که می‌تواند در تحلیل ابعاد زیربنایی سؤالات استفاده شود و تمام نتایج تحلیلی را بصورت یکپارچه و یکجا ارائه دهد. بنابراین می‌توان از تحلیل داده‌های شبکه در حوزه سنجش و آموزش بعنوان روشی دقیق، کم‌هزینه و آسان در ساخت آزمون‌های انطباقی استفاده نمود. الگوریتم انطباقی پیشنهاد شده در این پژوهش نه تنها به کوتاه شدن آزمون (۱۴۰ سؤالی) کمک می‌کند و مانع ورود بیشتر خطاها به نتایج می‌شود، بلکه نسبت به آزمون ثابت یعنی آزمون ۱۴۰ سؤالی دقت برآوردهای بهتری دارد، چرا که در سنجش توانایی افراد در هر مرحله از نظریه مدرن اندازه‌گیری بهره می‌گیرد. اگر آزمونی بتواند در زمان کمتر مانع وارد شدن خطاهای بیشتر مانند خستگی، تداخل، یکنواختی و... غیره شود، بر اساس اصل صرفه جویی در علم این روش ارجح است. الگوریتم یا الگوی پیشنهادی در این مطالعه یک الگوی چندسطحی است که از اطلاعات پیشین ممتحنین برای انتخاب بهینه سؤال بعد (انطباقی شدن) استفاده می‌کند. در شیوه‌های سنتی انطباقی مانند الگوی قطعه‌بندی شده، الگوی دو مرحله‌ای یا دو بخشی، انتخاب شخصی یا خودگزینی، سنجش تناوبی، سنجش هرمی و سنجش انطباقی طبقه‌دار، از اطلاعات پیشین برای برآورد یا انتخاب سؤال اولیه استفاده می‌شد یا حداقل فرضی برای شروع وجود داشت (ثورندایک، ۱۹۱۷؛ ترجمه هومن، ۱۳۹۵). نقطه شروع آزمون انطباقی وابسته به اطلاعاتی است که از آزمون شونده وجود دارد، به طور کلی یا این اطلاعات موجود است و یا موجود نیست، تصمیم

الگوریتم حاضر (تهیه یک برنامه کامپیوتری) به دلیل محدودیت زمانی و هزینه مالی است. خوشبختانه با وجود چنین محدودیت‌هایی پژوهشگران موفق به جمع‌آوری داده در حجم بالا (حدود ۱۶۰۰ نفر) و انجام تحلیل‌های سطح بالا شده‌اند. با این وجود برای اجرای عملی و ساخت نرم‌افزاری محدودیت‌ها حاکم است که می‌تواند در قالب پیشنهاد در تحقیقات بعدی به انجام برسد. مدل ارائه شده در این پژوهش قابلیت دارد تا تحت وب درآمده و بتواند با استفاده از شبکه اینترنت به کاربران در حوزه سنجش توانمندی حجم لغات کمک نماید. از آنجایی که آزمون نیشن و وارینگ یک آزمون معتبر جهانی در حوزه اندازه‌گیری حجم لغات است، مؤسسات مختلف زبان‌آموزی، مدارس و دانشگاه‌ها می‌توانند از این الگوریتم در جهت تعیین سطح زبان‌آموزان و شناسایی میزان حجم لغات دانش‌آموزان و دانشجویان به عنوان یک آزمون تشخیصی استفاده نمایند. همچنین در این پژوهش تأکید بر روش‌شناسی است و همان‌طور که بورسوم (۲۰۰۸) مطرح نموده پدیده‌های روانشناختی دارای ماهیت شبکه‌ای هستند. بنابراین می‌توان از تکنیک تحلیل داده‌های شبکه‌ای در بررسی سایر پدیده‌های روانشناختی چه در مطالعات بنیادی و چه در مطالعات کاربردی، استفاده نمود.

منابع

- ثورندایک، رابرت ال (۱۹۱۷). *روانشنجی کاربردی*. ترجمه: حیدرعلی هومن (۱۳۹۵). چاپ پنجم. تهران: انتشارات دانشگاه تهران.
- ضرغامی، محمدحسین؛ قائمی، فرحناز و قائمی، فاطمه (۱۳۹۲). برآورد استعداد افراد در فعال‌سازی ژن‌ها. *فصلنامه ژنتیک در هزاره سوم*، ۱۱(۱)، ۲۹۷۹-۲۹۷۱.
- شریفی‌یگانه، نگار؛ فلسفی‌نژاد، محمدرضا؛ دلاور، علی؛ فرخی، نور علی و جمالی، احسان (۱۳۹۵). تعیین مقایسه‌پذیری برآورد پارامتر توانایی در سنجش انطباقی کامپیوتری و مداد - کاغذی. *فصلنامه مطالعات اندازه‌گیری و ارزشیابی آموزشی*، ۶(۱۴)، ۲۳۴-۲۰۳.
- عراقی، لادن؛ طاهری، آزاده؛ جزء رضایی، زهرا؛ عباسپور، افروز و ضرغامی، محمدحسین (۱۳۹۶). آزمون کاربرد تحلیل داده‌های شبکه‌ای در تحلیل سؤال. *فصلنامه اندازه‌گیری تربیتی*، ۷(۲۸)، ۲۴۰-۲۰۷.
- Bennett, R. E. (2002). Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *The Journal of Technology, Learning and Assessment*, 1(1), 1-23.
- Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika*, 71(3), 425-440

محور اصلی انتخاب سؤال اول در این مطالعه می‌باشد. در الگوهای سنتی تنها یک حدس یا تخمین توانایی مفروض آزمودنی سبب انتخاب اولین سؤال می‌شود و بعد از اجرای اولین سؤال، پاسخ آزمودنی به عنوان اطلاعات پیشین قلمداد می‌شود. اما از آنجایی که سؤالات در یک شبکه به هم پیوسته قرار دارند و بین سؤالات ارتباط (در این مطالعه ارتباط خالص) برقرار است، پاسخ فرد (یعنی صحیح یا غلط) برای انتخاب نزدیک‌ترین سؤال به اولین سؤال، تعیین کننده است. نزدیک‌ترین سؤال، سؤالی است که نه تنها به لحاظ محتوای مورد مطالعه همخوان با اولین سؤال است، بلکه استراتژی روش‌های سنتی انطباقی را به بهترین صورت در نظر می‌گیرد. به عبارتی به جای تمرکز صرف بر دشواری سؤال با تعریف نظریه کلاسیک در روش‌های سنتی، پارامتر دشواری محاسبه شده بر اساس نظریه سؤال - پاسخ را مبنا قرار می‌دهد که مشکلات نظری و عملی کمتری در آزمون‌سازی انطباقی دارد و قابلیت انطباقی‌سازی آن بیشتر است. رفتار الگوی شبکه مطرح شده در این پژوهش در مقایسه با روش‌های انطباقی موجود در ادبیات پژوهش، به گونه‌ای است که سؤالات براساس ابعاد زیربنایی، رویه‌ها، وجوه یا هر فاکتور معلوم یا نامعلوم که شباهت بیشتری به یکدیگر دارند در رابطه نزدیک‌تری قرار می‌گیرند (مارکات و پنمان، ۲۰۱۹) و سؤالاتی که با اهداف آزمون (در اینجا حجم لغات انگلیسی آزمون‌شوندگان) همخوان نیستند در شبکه نقش محوری ندارند. انتخاب سؤالات در الگوی پیشنهادی، تنها بر اساس پارامتر دشواری سؤالات نیست بلکه ارتباط خالص بین سؤالات تعریف شده براساس احتمالات شرطی که نشان‌دهنده استقلال و وابستگی است، نیز اهمیت دارد. سؤالاتی که به عنوان سؤال اول انتخاب می‌شوند که در سنجش توانایی نقش برجسته‌تری دارند و به عنوان سؤالات مرکزی شمرده می‌شوند. این سؤالات مرکزی سؤالاتی هستند که بار محتوایی کل شبکه سؤالات را بر دوش می‌کشند و شاخص‌های نسبتاً مناسبی از محتوایی هستند که آزمون به طور کلی می‌سنجد (ولل، ۲۰۰۴). به عبارتی الگوریتم حاضر اهمیت شبکه ارتباطی بین سؤالات را در موضوع انطباقی‌سازی مدنظر قرار داده است. این موضوع نه تنها در انتخاب سؤال اول بلکه در کل فرآیند انتخاب سؤال تأثیر می‌گذارد.

یکی از محدودیت‌های این پژوهش که می‌تواند به عنوان پیشنهادی برای پژوهش‌های آینده باشد، فقدان عملی‌سازی

- adaptive testing. *Int J Educ Technol High Educ*, 14(2), 1-21.
- Plajner, M., & Vomlel, J. (2015). Bayesian Network Models for Adaptive Testing. Technical report, ArXiv: 1511.08488.
- Rust, J., Golombok, S. (2009). *Modern Psychometrics. The Science of Psychological Assessment*, (3rd ed.). New York: Routledge.
- Saarinen, S., Cater, E., Littman, M. L. (2020). Applying prerequisite structure inference to adaptive testing. *In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge* (pp. 422-427).
- Tay, L., Meade, A. W., Cao, M. (2015). An overview and practical guide to IRT measurement equivalence analysis. *Organizational Research Methods*, 18(1), 3-46.
- Van der Linden, W. J., & Glas, C. A. W. (2000). *Computer adaptive testing: Theory and practice*. Norwell, MA: Kluwer Academic Publishers.
- Vomlel, J. (2004). Bayesian Networks in Educational Testing. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 12(supp01), 83-100.
- Ying, M. H., & Huang, H. H. (2010). Item selection strategic via social network analyze. *International Journal of Digital Content Technology and its Applications*, 5(3), 727-732.
- Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, 64(9), 1089-1108.
- Culbertson, M.J. (2015). Bayesian Networks in Educational Assessment: The State of the Field. *Applied Psychological Measurement*, 40(1) 3-21.
- Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47(1), 80-91.
- Edwards, J.R. (2011). The fallacy of formative measurement. *Organizational Research Methods*, 14(2), 370-388.
- Epskamp, S., Cramer, A. O. J., Waldrop, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software*, 48(4), 1-18.
- Fayers, P. M. (2007). Applying item response theory and computer adaptive testing: the challenges for health outcomes assessment. *Quality of Life Research*, 16(1), 187-194.
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3-5), 75-174.
- Gmoch, R. (2012). Computerized Adaptive Testing in Poland/Komputerowe Testowanie Adaptacyjne W Polsce. *Chemistry-Didactics-Ecology-Metrology*, 17(1-2), 61-68.
- Han, K. T. (2012). An efficiency balanced information criterion for item selection in computerized adaptive testing. *Journal of Educational Measurement*, 49(3), 225-246.
- Ilhan, M., & Guler, N. (2018). A Comparison of Difficulty Indices Calculated for Open-Ended Items According to Classical Test Theory and Many Facets Rasch Model. *Eurasian Journal of Educational Research*, 75(6), 99-114.
- Kalender, I. (2011). *Effects of different computerized adaptive testing strategies on recovery of ability* (Unpublished doctoral dissertation). Middle East Technical University, Ankara.
- Kyung, S.K., Yong, S.C. (2012). Bayesian Network Approach to Computerized Adaptive Testing. *Internet Journal of Smart Home*, 6(3), 75-82.
- Lord, F.M., & Novick, M.R. (1968) *Statistical Theories of Mental Test Scores*. Addison-Wesley, Menlo Park.
- Marcot, B. G., & Penman, T. D. (2019). Advances in Bayesian network modelling: Integration of modelling technologies. *Environmental modelling & software*, 111, 386-393.
- Nation, I.S.P., & Beglar, D. (2007) A Vocabulary Size Test. *The Language Teacher*, 31(7), 9-13.
- Nation, I.S.P., & Waring, R. (1997). Vocabulary Size, text coverage and word lists. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary: Description, acquisition and pedagogy* (pp. 6-19). Cambridge: Cambridge University Press.
- Oppl, S., Reisinger, F., Eckmaier, A., & Helm, Ch. (2017). A flexible online platform for computerized