

Research Paper

Investigating the psychometric properties of english vocabulary size test using network data analysis

Nadia Soltani¹, Shahram Vahedi², Mansour Bayrami³, Mohammad Hosein Zarghami⁴

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

2. Professor, Department of Educational Sciences, University of Tabriz, Tabriz, Iran.

3. Professor, Department of Psychology, University of Tabriz, Tabriz, Iran.

4. Assistant Professor, Behavioral Sciences Research Center, Baqiyatallah University of Medical Sciences, Tehran, Iran.

Citation: Soltani, N., Vahedi, Sh., Bayrami, M., Zarghami, M.H. Investigating the psychometric properties of english vocabulary size test using network data analysis. J of Psychological Science. 2021; 20(103): 1131-1146.

URL: <https://psychologicalscience.ir/article-1-1131-fa.html>



ORCID



[10.52547/JPS.20.103.1131](https://doi.org/10.52547/JPS.20.103.1131)

ARTICLE INFO

ABSTRACT

Keywords:

Network Theory,
Question analysis,
Network parameters,
Bayesian Network

Background: Network science as a new field in psychological measurement provides a good platform for test analysis, but in Iran, research on the analysis of test questions using network data analysis has been neglected.

Aims: The aim of this study was to introduce network data analysis as a psychometric-mathematical technique and its use in the questionnaire analysis process.

Methods: This research is descriptive-analytical, the statistical population of all language learners was the top ten language teaching institutes in Tehran in 2020, from which 1556 language learners were selected by convenience sampling method and answering the English vocabulary size test (Nation and Waring, 1997). Data analysis was performed by qgraph software package under R software.

Results: The graph obtained from the communication network of questions based on the Fruchterman-Ringold algorithm was plotted, the parent-child relationship of the questions was determined based on the Bayesian network, and the descriptive features of the test and the indicators of network centrality were extracted.

Conclusion: Introducing the technique of analyzing network data and using it for questions analysis process is one of the achievements of this article. Using the advantages of network science such as the ability to simplify, manage, store and collect high-volume data and analyze test questions based on the map and indicators that are fully consistent with it, we can understand the internal relationship between different test questions, and by considering the importance of the questions in the network structure of the net relationship between the questions, it is possible to improve the testing methods. Similar studies and researches are proposed to compare the results of network data analysis with psychometric properties obtained based on conventional and classical methods.

Received: 30 Jan 2021

Accepted: 28 Feb 2021

Available: 23 Sep 2021

* **Corresponding Author:** Shahram Vahedi, Professor, Department of Educational Sciences, University of Tabriz, Tabriz, Iran.

E-mail: vahedi117@yahoo.com

Tel: (+98) 411-3392090

2476-5740/ © 2021 The Authors. This is an open access article under the CC BY-NC-ND license

(<https://creativecommons.org/licenses/by-nc/4.0/>).



Extended Abstract

Introduction

The use of network data analysis in the field of psychology and psychometrics is considered to be the latest methodology. Network methodology makes it possible to study several objectives simultaneously. Under the best of conditions, the measurement model provides an integrated texture that meets the requirements for testing at the same time (Araghi, Taheri, Joz Ramazani, Abas pour & Zarghami, 2017). The ability to integrate different variables through network data analysis can be used in test analysis and extraction of test indicators. Two reasons have led to the consideration of the network approach in the field of psychometric: a) in recent years, models and theories about latent variables have prevailed b) Unlike latent models, the network perspective has helped to develop dynamic systems (Borsboom, 2008). The purpose of question analysis and determination of test-construction indicators is to identify inappropriate questions and correct them, and finally to develop a measurement tool. In the network method, the question analysis is performed based on the map and parameters that correspond exactly to it, which is the result of the internal relationship between different test questions (Dörnyei, 2014). Here, network nodes refer to the test questions being studied, and the relationship between them is determined by interface lines called edges (Saarinen, Cater & Littman, 2020).

In this research, the benefits of network data analysis are used to analyze English vocabulary size test. Today, the importance and use of English in professional and educational affairs in the international arena is not hidden from anyone. Assessing language learners is an important part of teaching any foreign language. Inaccurate assessment of individuals' language knowledge can greatly affect the learning process. Therefore, designing the principles of making standardized tests can greatly guarantee the success of language teaching. The aim of this study was to introduce network data analysis as a psychometric-mathematical technique and its use in the questionnaire analysis process.

Method

This research is descriptive-analytical, the statistical population of all language learners was the top ten language teaching institutes in Tehran in 2020, from which 1556 language learners were selected by convenience sampling method and answering the English vocabulary size test (Nation and Waring, 1997). To draw the network map of vocabulary size test, Fruchterman-Ringold

algorithm and qgraph software package under R software were used and based on that, network parameters were extracted. To investigate the parent-child relationship, Bayesian network were drawn based on the Hilling Climbing (HC) algorithm using bnlearn and bnviewer software packages.

Results

In this study, the Cronbach's alpha reliability coefficient for all questions (140 questions) was 0.921. Nonlinear multifactorial analysis was performed under R software. We need a correlation matrix to map the network of relationships between vocabulary test questions. Since the test questions are dual values, the tetrachoric correlation matrix was chosen. The graph obtained from the communication network of questions based on the Fruchterman-Ringold algorithm was plotted (figure 1). Each graph has its own characteristics: A) Edge thickness (meaning the lines drawn between two nodes) indicates a significant relationship between the two questions based on correlation. B) The more this connection, the thicker the edge between the two nodes, the presence of the interface line means a meaningful connection. C) Indicates the color of the edges to communicate between variables. Green edges indicate a significant positive relationship between nodes, and red edges indicate a significant negative relationship. D) Central nodes are nodes that are located in the center of the graph, peripheral nodes are nodes that are located around. F) The proximity of nodes to each other indicates a high relationship between them. The parent-child relationship of the questions was determined based on the Bayesian network (figure 2), and the descriptive features of the test and the indicators of network centrality were extracted.

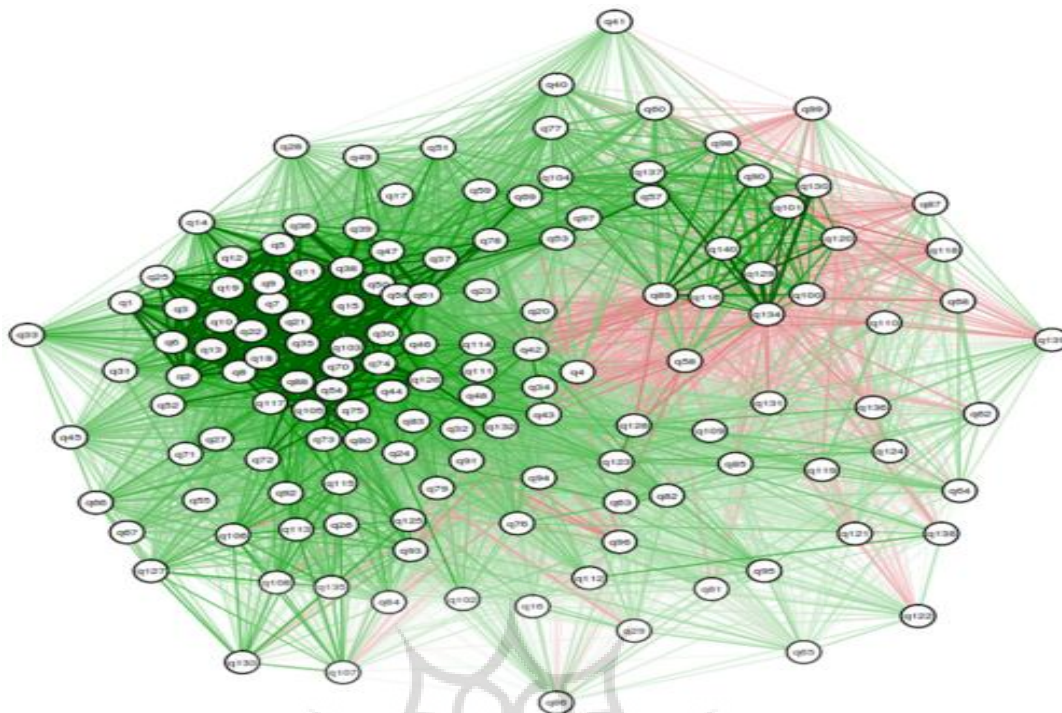


Figure 1: Network map of the relationship based on tetracoric correlation



Figure 2: Bayesian network with HC algorithm

Conclusion

Introducing the technique of analyzing network data and using it for questions analysis process is one of the achievements of this article. The use of graph theory-based methods is faster in finding inappropriate questions, and through it, the relationship matrix between the questions can be visualized quite clearly and a lot of information can be presented together. Using the advantages of network science such as the ability to simplify, manage, store and collect high-volume data and analyze test questions based on the map and indicators that are fully consistent with it, we can understand the internal relationship between

different test questions, and by considering the importance of the questions in the network structure of the net relationship between the questions, it is possible to improve the testing methods. Similar studies and researches are proposed to compare the results of network data analysis with psychometric properties obtained based on conventional and classical methods.

Ethical Considerations

Compliance with ethical guidelines: This article is extracted from PhD dissertation of the first author in the major of educational psychology from Tabriz University and the date of

approval of the proposal is 11/06/2019. Before the test, while giving the necessary explanations to the participants about the importance and purpose of the research, they were given the opportunity to ask their questions in the field of research, as well as before the test to not be forced and voluntary to participate in the research and It was also pointed out that the personal information of the volunteers was kept confidential and not used elsewhere, and in order to gain the trust and anonymity of the participants, their names and surnames were refused.

Funding: This study was conducted as a PhD thesis with no financial support.

Authors' contribution: The first author is the main researcher of this research. The second and third authors are the supervisors and forth auther was the advisor.

Conflict of interest: The authers decler no conflict of interest for this study.

Acknowledgments: I would like to appreciate the supervisors, advisor and the participants.



بررسی ویژگی‌های روان‌سنجی آزمون خزانه لغات انگلیسی با استفاده از تحلیل داده‌های شبکه‌ای

نادیا سلطانی^۱، شهرام واحدی^{۲*}، منصور بیرامی^۳، محمد حسین ضرغامی^۴

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

۲. استاد، گروه علوم تربیتی، دانشگاه تبریز، تبریز، ایران.

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

۴. استادیار، مرکز تحقیقات علوم رفتاری، دانشگاه علوم پزشکی بقیه الله، تهران، ایران.

چکیده

مشخصات مقاله

زمینه: علم شبکه به عنوان یک حوزه نوین در اندازه‌گیری روانشناختی بستر مناسبی را برای انجام تحلیل‌های آزمون‌سازی فراهم می‌آورد، اما در ایران پژوهش در خصوص تحلیل سوالات آزمون با استفاده از تحلیل داده‌های شبکه‌ای مغفول مانده است.

هدف: این پژوهش با هدف معرفی تحلیل داده‌های شبکه‌ای به عنوان یک تکنیک روان‌سنجی-ریاضیاتی و استفاده از آن در فرآیند تحلیل سوالات پرسشنامه انجام شد.

روش: این پژوهش از نوع توصیفی-تحلیلی است، جامعه آماری کلیه زبان آموزان ده موسسه برتر آموزش زبان شهر تهران در سال ۱۳۹۸ بود که از بین آن‌ها ۱۵۵۶ زبان آموز به روش نمونه‌گیری در دسترس انتخاب شدند و به آزمون خزانه لغات انگلیسی (نیشن و وارینگ، ۱۹۹۷) پاسخ دادند. تجزیه و تحلیل داده‌ها توسط بسته نرم‌افزاری qgraph تحت نرم افزار R انجام شد.

یافته‌ها: گراف حاصل از شبکه از تباطی سوالات مبتنی بر الگوریتم فراچترمن-رینگولد ترسیم گردید، رابطه والد-فرزند سوالات براساس شبکه نیز مشخص شد و ویژگی‌های توصیفی آزمون و شاخص‌های مرکزیت شبکه استخراج گردید.

نتیجه‌گیری: معرفی تکنیک تحلیل داده‌های شبکه‌ای و استفاده از آن به منظور فرآیند تحلیل سوالات، یکی از دستاوردهای این پژوهش است. با استفاده از مزایای علم شبکه مانند قابلیت ساده‌سازی، مدیریت، ذخیره و مجموعه‌سازی داده‌های با حجم بالا و تحلیل سوالات آزمون براساس نقشه و شاخص‌های کاملاً منطبق بر آن، می‌توان به ارتباط درونی سوالات مختلف آزمون پی برد و با در نظر گرفتن اهمیت سوالات در ساختار شبکه ارتباط خالص بین سوالات، می‌توان به ارتقاء روش‌های آزمون‌سازی پرداخت. انجام مطالعات مشابه و پژوهش‌هایی به منظور مقایسه نتایج حاصل از تحلیل داده‌های شبکه‌ای با ویژگی‌های روان‌سنجی که بر اساس شیوه‌های مرسوم و کلاسیک بدست می‌آیند پیشنهاد می‌گردد.

کلیدواژه‌ها:

نظریه شبکه،

تحلیل سوال،

شاخص‌های شبکه،

شبکه‌بیزین

دریافت شده: ۱۳۹۹/۱۱/۱۱

پذیرفته شده: ۱۳۹۹/۱۲/۱۰

منتشر شده: ۱۴۰۰/۰۷/۰۱

* نویسنده مسئول: شهرام واحدی، استاد، گروه علوم تربیتی، دانشگاه تبریز، تبریز، ایران.

رایانامه: vahedi117@yahoo.com

تلفن: ۰۴۱۱-۳۳۹۲۰۹۰

مقدمه

یکی از مسایل اصلی آزمون‌سازی^۱ روانی بررسی نظری و تجربی^۲ چگونگی الگوی روابط در سایه مدل‌های علمی است. مدل‌های اندازه‌گیری روانشناختی که تاکنون گسترش یافته‌اند، مبتنی بر دو رویکرد تفسیری‌اند. تفسیرهای بازتابی^۳ که در آنها صفت اندازه‌گیری شده به عنوان علت مشترک مشاهدات قلمداد می‌شود و تفسیرهای تجمعی یا تراکمی^۴ که در آنها صفت اندازه‌گیری شده حاصل اثر مشترک مشاهدات در نظر گرفته می‌شود. نظریه کلاسیک^۵ مبتنی بر تفسیرهای تراکمی است و نظریه سؤال پاسخ^۶ به عنوان یک نظریه بازتابی محسوب می‌شود (برسبوم، ۲۰۰۸). در کنار این مدل‌ها، مدل شبکه که صفت، ویژگی یا پدیده مورد مطالعه را حاصل ارتباط بین اجزای تشکیل دهنده آن می‌داند نیز اخیراً توسط گروهی از پژوهشگران روانشناسی دانشگاه آمستردام هلند، پایه‌گذاری و گسترش یافته است. اندازه‌گیری روانشناختی بر اساس این رویکرد، به سنجش روابط بین اجزا و شاخص‌های شبکه می‌پردازد (برسبوم و کرامر، ۲۰۱۳).

تحلیل شبکه‌ای پتانسیل‌هایی دارد که آن را در حد یک روش‌شناسی خاص ارتقاء می‌دهد. روش‌شناسی که می‌توان واقعیت‌گرایی انتقادی را به عنوان مبنای هستی‌شناسی آن در نظر گرفت (باخ‌هانسن، ۲۰۱۴). هستی‌شناسی که در برگیرنده نمایش، نامگذاری رسمی، تعریف طبقات، ویژگی‌ها و روابط بین مفاهیم و داده‌ها و موجودیت‌هایی هستند که صورت واقعی به یک یا همه حوزه‌ها می‌بخشد. در رئالیسم انتقادی آن چیزی واقعی است که به نحو علی مؤثر باشد (توانا، ۱۳۹۳).

در ادبیات پیشین، روانشناسی تربیتی به عنوان شاخه‌ای از علم قلمداد می‌شد که علاوه بر کاربری نظریات روانشناسی در حوزه تعلیم و تربیت به اندازه‌گیری روانشناختی 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) (رست و گلمبوک، ۱۹۸۹). از این رو درگیر شدن روانشناسی تربیتی با رویکردهای بین‌رشته‌ای مدرن می‌تواند افق‌های تازه‌ای را به روی این حوزه از علم روانشناسی تربیتی بگشاید. مطالعاتی وجود داشته است که

به استفاده از مدل‌های مختلف شبکه در حوزه سنجش تربیتی پردازند، به عنوان مثال می‌توان به تحقیقات آلموند، میسلوی، استنبرگ، یان و ویلیامسون (۲۰۱۵)، کالبرتسون (۲۰۱۵) و میلان، لابودا، پریز دیلا کروز و گارسیا (۲۰۱۰) اشاره کرد که از مدل‌های بیزین شبکه در ارزیابی‌های آموزشی استفاده کرده‌اند.

یک مدل اندازه‌گیری خوب چارچوب مرجعی را برای طراحی آزمون و مشکلات عملی دیگر ایجاد می‌کند. در بهترین شرایط مدل اندازه‌گیری بافت یکپارچه‌ای را فراهم می‌کند که الزامات ساخت آزمون را به صورت همزمان در بستر خود داراست. علم شبکه و نظریه گراف بستر مناسبی را برای انجام تحلیل‌های آزمون‌سازی فراهم می‌آورد. کاربرد این علم در حوزه‌های مختلف از هنر تا علوم پایه، علوم تجربی و علوم‌نومین در مقالات و پروژه‌های مختلف منتشر شده است. روش‌شناسی شبکه امکان مطالعه همزمان چندین هدف را میسر می‌سازد. این موضوع روان‌سنجان را ترغیب می‌کند تا در سنجش سازه‌های انسانی از تحلیل شبکه که ابزاری برای نشان دادن عناصر یک سیستم و ارتباط درونی آن است، استفاده نمایند (عراقی، طاهری، جزء رضانی و همکاران، ۱۳۹۶).

به یک گراف 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^V و به هر عضو E یک یال^G می‌گوییم (فورتوناتو، 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، اشیمین و برسبوم، ۲۰۱۲).

1. testing
2. empirical
3. reflective interpret
4. formative interpret

5. classic theory
6. item response theory
7. node
8. edge

گره‌های شبکه همان سؤالات آزمون مورد مطالعه خواهند بود و ارتباط بین آنها خطوط رابط گراف ترسیم شده‌اند. شاخص‌های شبکه با نمایش بصری آن انطباق کامل دارد و این قابلیت قدرت دیداری‌سازی نتایج و ارائه گزارش‌های ساده و در عین حال جامعی از تحلیل آزمون‌سازی سؤالات ارائه می‌دهد. شاخص‌های مبتنی بر نظریه کلاسیک اندازه‌گیری و نظریه سؤال پاسخ مبتنی بر روش‌های تجریدی ریاضیاتی هستند که علاوه بر انتقاد فلسفی اساسی که به آنها گرفته شده است، پیچیدگی بالایی دارند و قدرت تحلیل سریع و گزارش‌دهی را دشوار می‌سازند. اگر چه رویکرد شبکه‌ای از ریاضیات قوی برخوردار است اما انطباق ریاضیات آن با نمایش بصری شبکه، امکان درک درست ریاضیات را به همراه دارد (سرین، کاتر و لیتمن، ۲۰۲۰).

در اینجا از مزایای تحلیل داده‌های شبکه‌ای در جهت تحلیل سؤالات آزمون خزانه لغات زبان انگلیسی استفاده می‌شود. امروزه اهمیت و کاربرد زبان انگلیسی در امور شغلی و تحصیلی در عرصه بین‌الملل بر هیچ‌کس پوشیده نیست. با توجه به نیازهای متخصصین حوزه‌های مختلف به ابزار زبان انگلیسی می‌بایست در جهت ارتقا کیفی دانش عمومی و تخصصی زبان انگلیسی با دیدگاه اجرایی و بلند مدت برنامه‌ریزی شود. سنجش آموخته‌های زبان آموزان بخش مهمی از آموزش هر زبان خارجی محسوب می‌شود. ارزیابی غیر دقیق معلومات زبانی افراد می‌تواند فرآیند آموزش را بشدت تحت تأثیر قرار دهد. بنابراین طراحی اصول ساخت آزمون‌های استاندارد می‌تواند موفقیت آموزش زبان را تا حد زیادی تضمین کند. معرفی تحلیل داده‌های شبکه‌ای به عنوان یک روش در گسترش مفاهیم آزمون‌سازی و رسوب دانش در این حوزه را می‌توان به عنوان مهمترین هدف این پژوهش در نظر گرفت. بنابراین این پژوهش با هدف کاربرد علم شبکه در ترسیم شبکه ارتباطی سؤالات آزمون خزانه لغات انگلیسی نیشن و وارینگ (۱۹۹۷) انجام شد.

روش

الف) طرح پژوهش و شرکت کنندگان: این پژوهش تحلیلی - توصیفی و از نظر هدف بنیادی - کاربردی است. ابزار مورد استفاده در این پژوهش، آزمون خزانه لغات انگلیسی نیشن و وارینگ (۱۹۹۷) می‌باشد و تمرکز اصلی این پژوهش بر تحلیل سؤالات آزمون است. اجرای مختلف این

تحلیل شبکه در حوزه روانشناسی و اندازه‌گیری روانی می‌تواند تلاش بین‌رشته‌ای قلمداد شود که در آن نظریات روانشناسی با روش‌شناسی کمی آماری و ریاضی شکل گرفته و گسترش می‌یابد و در کنار آن، برای چالش‌های تحلیلی در این حوزه پاسخ مناسب ارائه می‌دهد. تحلیل شبکه‌ای، روش‌شناسی علمی و تکنیکی آماری در توصیف، تحلیل و بررسی روابط تجربی بین موضوعات مورد پژوهش - که در ادبیات تحلیل شبکه به آنها گره می‌گویند - و ساختار این روابط است. این گره‌ها می‌توانند افراد، اشیاء، گروه‌ها، مفاهیم، مشاغل، سؤالات، عوامل، ابعاد، خوشه‌ها و طبقات مختلف باشند. این تحلیل دارای این قابلیت است که واقعیت بیرونی - که به صورت شبکه است - را به طور منظم و فشرده تشریح کند و امکان چکیده‌سازی، ذخیره‌سازی، تکثیر، ترکیب و یکپارچه‌سازی، بازیابی و دیداری‌سازی اطلاعات را فراهم آورد (چلبی، ۱۳۷۵). در حال حاضر پژوهش‌های مربوط به تحلیل شبکه چهار هدف اصلی را دنبال می‌کنند که مهمترین آن درک یکپارچگی و انسجام پدیده‌هایی است که متأثر از متغیرها یا مؤلفه‌های مختلف می‌باشند (بورسوم و کرامر، ۲۰۱۳).

قابلیت یکپارچه‌سازی و ادغام متغیرهای مختلف از طریق این تحلیل را می‌توان در تحلیل‌های آزمون‌سازی سؤالات و استخراج شاخص‌های آزمون‌سازی استفاده کرد. رویکرد تحلیل شبکه‌ای به عنوان روش‌شناسی و تکنیک منعطف، خاستگاه ریاضیاتی دارد که می‌تواند در حوزه آزمون‌سازی بکار گرفته شود. 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 برمی‌گیرد (بورسوم، ۲۰۰۸).

هدف تحلیل سؤال و تعیین شاخص‌های آزمون‌سازی سؤالات، 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 براساس نقشه و شاخص‌های کاملاً منطبق بر آن انجام می‌شود که نتیجه ارتباط درونی سؤالات مختلف آزمون است (دورنوی، ۲۰۱۴). در اینجا

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). برای تعیین پارامتر سؤالات، آزمون روی گروهی از شرکت‌کنندگان اجرا شد. جامعه آماری کلیه زبان آموزان ده موسسه برتر آموزش زبان شهر تهران در سال ۱۳۹۸ بود که از بین آن‌ها ۱۵۵۶ زبان آموز به روش نمونه‌گیری در دسترس انتخاب شدند و به آزمون خزانه لغات انگلیسی (نیشن و وارینگ، ۱۹۹۷) پاسخ دادند. لازم به ذکر است در پژوهش عراقی و همکاران (۱۳۹۶) که به تحلیل سؤالات آزمون با استفاده از علم شبکه پرداخته‌اند حجم نمونه ۱۰۰۰ نفر در نظر گرفته شده است، همچنین ریکاس (۲۰۰۹) جهت استفاده از تحلیل عاملی غیرخطی با اطلاعات کامل، نمونه‌هایی با حجم بسیار پایین‌تر پیشنهاد نموده لذا در این پژوهش سعی گردید جهت اطمینان بیشتر حجم نمونه نسبت به این مطالعات بطور قابل توجهی بالاتر باشد. قبل از اجرای آزمون ضمن ارائه توضیحات لازم به شرکت‌کنندگان در خصوص اهمیت و هدف پژوهش، این فرصت به آن‌ها داده شد که سؤالات خود را در زمینه پژوهش مورد نظر مطرح نمایند، همچنین قبل از اجرای آزمون به عدم اجبار و داوطلبانه بودن شرکت در پژوهش و نیز به محرمانه ماندن اطلاعات شخصی داوطلبان و عدم استفاده از این اطلاعات در جای دیگر، اشاره شد و به منظور جلب اعتماد و غیر قابل شناسایی بودن شرکت‌کنندگان، از اخذ نام و نام خانوادگی آنها خودداری گردید. جهت ترسیم نقشه شبکه ارتباط بین سؤالات آزمون خزانه لغات از الگوریتم فراچترمن-رینگولد و بسته نرم افزاری qgraph تحت نرم‌افزار R استفاده شد و بر اساس آن شاخص‌های شبکه استخراج گردید. جهت بررسی رابطه والد-فرزندی سؤالات شبکه بیز بر اساس الگوریتم تپه نوردی یا HC با استفاده از بسته‌های نرم‌افزاری bnlearn و bnviewer ترسیم شد.

(ب) ابزار

در این مطالعه تأکید بر روش‌شناسی و تکنیک است که به عنوان نمونه از آزمون ۱۴۰ سؤالی خزانه لغات انگلیسی که توسط نیشن و وارینگ (۱۹۹۷) تهیه شده استفاده می‌شود. به عبارتی هدف بررسی ویژگی‌های روان‌سنجی سؤالات آزمون با استفاده از علم شبکه است که می‌تواند در خصوص آزمون‌های دیگر نیز بکار گرفته شود. البته آزمون نیشن و وارینگ ویژگی‌های خاصی

دارد که در انتخاب پژوهشگر بی‌تأثیر نبوده که از آن جمله می‌توان به موارد زیر اشاره کرد: این آزمون تمام لغاتی را شامل می‌شود که فرد برای درک لغات استفاده می‌کند و ربطی به سایر بخش‌های زبان مانند گوش دادن، درک مطلب و غیره ندارد. از این آزمون با اهداف مختلفی استفاده می‌شود، مثلاً برای تدوین برنامه‌ریزی و طراحی درسی، برای سنجش قبل از ورود به دوره یادگیری خواندن زبان انگلیسی و غیره. این آزمون هم برای افراد انگلیسی زبان و هم برای افرادی که انگلیسی زبان دوم آنها است، استفاده می‌شود. این آزمون از سطوح مختلف دشواری لغات تشکیل شده است 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 (BNC)^۱ گرفته شده است که در برگیرنده ۱۰۰ میلیون لغت است و در بخش‌های مختلف ملل انگلیسی زبان وجود دارد. مجموعه لغات BNC 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. این آزمون چیزی را می‌سنجد که فرض سنجش آن وجود دارد، بگلار اعتقاد دارد که این آزمون یک عامل غالب را می‌سنجد و عوامل دیگر نقش چندانی در عملکرد آزمون ندارند. انتظار تمایز بین افراد مختلف با سطوح مختلف زبانی بر اساس این آزمون برآورده می‌شود و قدرت بالایی در تمایز سطوح مختلف زبانی از طریق این آزمون وجود دارد. به همین دلیل ابزار بسیار مناسبی برای سنجش میزان رشد زبانی، زبان‌آموزان محسوب می‌شود. اجراهای این آزمون در موقعیت‌های مختلف نتایج با ثبات و پایایی بالایی را نشان می‌دهد. اجرای مختلف این آزمون توسط بگلار 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). تفسیر و نمره‌دهی آزمون آسان و بصورت دو ارزشی (غلط یا صحیح) می‌باشد، سؤالات آزمون روشن و بدون ابهام‌اند. حداقل نمره کسب شده در این آزمون می‌تواند صفر و حداکثر ۱۴۰ باشد. 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).

1. British National Corpus

یافته‌ها

شرکت کنندگان در این پژوهش را هم زنان (۵۳ درصد) و هم مردان (۴۳ درصد) تشکیل می‌دهند، بقیه افراد جنسیت خود را مشخص نکرده‌اند. این افراد دارای سطح تحصیلات متفاوت بودند بیشترین افراد شرکت کننده در این پژوهش دارای تحصیلات کارشناسی ارشد (۳۴ درصد) و کارشناسی (۳۱ درصد) و مابقی زیر دیپلم (۱۶ درصد)، دیپلم (۱۲ درصد) و دکتری (۷ درصد) می‌باشند. به لحاظ اقتصادی ۶۰ درصد وضعیت اقتصادی خود را متوسط، ۳۰ درصد خوب، ۹ درصد فقیر و یک درصد مرفه قلمداد کرده‌اند. از بین این افراد ۹۱ درصد زبان اصلی آنها فارسی بوده است و سایرین به زبان‌های دیگر صحبت می‌کردند. از این بین ۸۸ درصد راست دست و ۱۲ درصد چپ دست بودند. در تحلیل داده‌ها ابتدا سؤالات آزمون اجرا شد. یکی از ویژگی‌های روانسنجی یک پرسشنامه مقدار پایایی آن است. در این پژوهش مقدار ضریب پایایی آلفای کرونباخ برای کل سؤالات (۱۴۰ سؤال) ۰/۹۲۱ بدست آمد. این موضوع نشان می‌دهد که همه سؤالات با یکدیگر همگن‌اند و از نگاه این شاخص توصیفی نیازی به حذف هیچ یک از سؤالات نیست.

با توجه به نتایج بدست آمده از ضریب آلفای کرونباخ و همگنی بالای سؤالات، و ادعای نیشن و وارینگ (۱۹۹۷) مبنی بر تک بعدی بودن این پرسشنامه، پژوهشگر پیشنهاد تحلیل عاملی غیرخطی با اطلاعات کامل را می‌دهد که محدودیت‌های تحلیل عاملی سنتی را ندارد. استفاده از تحلیل عاملی غیر خطی با اطلاعات کامل از لحاظ روانسنجی به منظور آزمون‌سازی در حوزه علوم انسانی و علوم رفتاری مناسب‌تر و دقیق‌تر می‌باشد و انطباق بیشتری با ماهیت موضوعات مربوط به انسان دارد. به بیان دیگر استفاده از این شیوه تحلیل عاملی، روایی آماری بالاتری نسبت به شیوه‌های سنتی تحلیل عاملی دارد. شیوه‌های کنونی (سنتی) تحلیل عاملی خطی به فهم مفاهیم زیربنایی و کشف ابعاد (یا سازه‌های اصلی) سؤالات مورد مطالعه کمک می‌نمایند و از ارزش و اعتبار مناسب برخوردارند، اما دارای محدودیت‌هایی نیز می‌باشند که این محدودیت‌ها ممکن است مانع دستیابی کامل به ابعاد زیربنایی داده‌ها گردد. بعضی از این محدودیت‌ها عبارتند از:

۱. استفاده از تحلیل عاملی به شیوه‌های سنتی مستلزم پیوسته بودن نمرات است. نمرات بدست آمده از این مطالعه دو ارزشی است و مقیاس آن حداکثر رتبه‌ای است.
 ۲. رابطه ابعاد با نمرات در تحلیل عاملی سنتی خطی فرض می‌شود. در صورتی که روابط در حوزه انسانی به دلیل تأثیر متغیرهای زیاد در پاسخ‌دهی خطی نیست.
 ۳. در تحلیل عاملی سنتی از کلیه اطلاعاتی که از پرسشنامه بدست می‌آید، استفاده نمی‌شود.
- این عوامل سبب می‌شود تا نتوان به روایی آماری نتایج بدست آمده از تحلیل‌های سنتی اعتماد نمود (ریکاس، ۲۰۰۹). تحلیل چند عاملی غیر خطی تحت نرم‌افزار R انجام می‌شود. شاخص‌های برازش این مدل در جدول ۱ نشان داده شده است.

جدول ۱. شاخص‌های برازش مدل تک عاملی مبتنی بر تحلیل عاملی غیر خطی

شاخص تناسب	مقدار معناداری
لگاریتم بیشینه درستمایی	-۱۰۶۷۸۳.۹
اطلاعات آکائیکه	۲۱۴۱۲۷.۹
اطلاعات بیزی	۲۱۵۶۰۳
اصلاح شده اطلاعات آکائیکه	۲۱۴۲۶۴.۴
اصلاح شده اطلاعات بیزی	۲۱۴۷۱۳.۵
G2 (1e+10)	۱۹۲۷۲۸.۲
ریشه خطای مجذور میانگین	p.value=1
	۰

مقدار آزمون معناداری G2 برابر ۱۹۲۷۲/۲ است که با توجه به ارزش p بدست آمده (p.value= 1) نشان می‌دهد که بین مدل تجربی که بر اساس داده‌ها بدست آمده و مدل مورد انتظار تک بعدی اختلاف معناداری وجود ندارد. مقدار ارزش p کامل نشان دهنده انطباق قابل قبول و بالای مدل نظری و مدل تجربی است. مقدار ریشه خطای مجذور میانگین (RMSEA= 0) نیز انطباق کامل مدل تجربی و مدل نظری را تأیید می‌کند. به بیان دیگر شاخص‌های برازش نشان می‌دهند که مدل تک عاملی برای این داده‌ها مطلوب و مناسب است. برای ترسیم نقشه شبکه ارتباط بین سؤالات آزمون حجم لغات به ماتریس همبستگی نیاز داریم. از آنجایی که سؤالات آزمون دو ارزشی هستند ماتریس همبستگی تراکوریک برای این کار انتخاب می‌شود. به دلیل حجم بالای این ماتریس (۱۴۰*۱۴۰) از درج آن در اینجا خودداری می‌شود.

نقشه شبکه ارتباط بین سؤالات آزمون حجم لغات براساس الگوریتم فراچترمن-رینگولد ترسیم شد. هر گراف ویژگی‌های خاص خود را دارد که به آنها پرداخته خواهد شد.

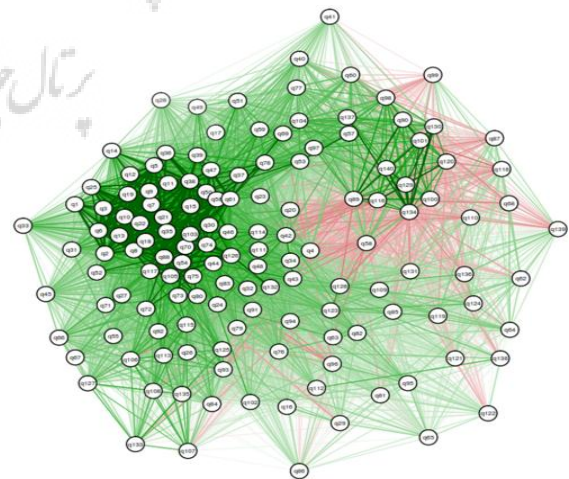
الف. ضخامت یال‌ها (منظور خطوط ترسیم شده بین دو گره) نشان دهنده ارتباط معنادار بین دو سؤال بر اساس همبستگی است. اگر بین دو متغیر که در ادبیات شبکه به آن گره گفته می‌شود یک خط (یال) رسم شده باشد نشان می‌دهد که بین این دو ارتباط وجود دارد، عدم وجود خط رابط بین دو متغیر نشان دهنده فقدان ارتباط معنادار بین آنها است.

ب) هر چه این ارتباط بیشتر باشد، یال بین دو گره قطورتر می‌شود، وجود خط رابط به معنای ارتباط معنادار است. هر چه قطر یال نازک‌تر باشد، شدت ارتباط معنادار کمتر است.

ج. رنگ یال‌ها جهت ارتباط بین متغیرها را نشان می‌دهد. یال‌های سبز رنگ نشان دهنده ارتباط مثبت معنادار بین گره‌ها است و رنگ قرمز یال، به معنای ارتباط منفی معنادار است.

د) گره‌های مرکزی گره‌هایی هستند که در مرکز گراف جای گرفته‌اند، 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 با یکدیگر نشان دهنده 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 استفاده می‌شود (عراقی و همکاران، ۱۳۹۶).



شکل ۱. نقشه شبکه ارتباط بین سؤالات آزمون حجم لغات بر اساس همبستگی تراکوریک

همان‌طور که همبستگی‌های ساده تراکوریک نشان می‌دهد، شبکه ارتباط بین نشانگان از ارتباط بالایی برخوردار است. همبستگی‌های بالا و مثبت به چگال شدن نقشه در بخش چپ نقشه منجر شده است. اکثر ارتباطات بین سؤالات معنادار است و این نشان دهنده انسجام معنادار شبکه سؤالات بر اساس همبستگی‌های ساده تراکوریک است. بر اساس این شبکه و شبکه همبستگی خالص بین سؤالات (به عبارتی در ارتباط بین هر دو سؤال اثر سایر سؤالات حذف می‌شود و بسیاری از همبستگی‌های مثبت، منفی می‌شوند که نشان‌دهنده بعد یا ابعاد مشترک بین سؤالات است) شاخص‌های مرکزیت شبکه بدست می‌آید.

برای دستیابی به اهمیت هر کدام از سؤالات در شبکه ارتباطی آنها، نیاز است تا از شاخص‌های مرکزیت شبکه استفاده شود. در اینجا چهار شاخص مختلف بینیت، نزدیکی، استحکام و تأثیر مورد انتظار در نظر گرفته شده‌اند و مقادیر آنها برای هر کدام از سؤالات در جداول مربوطه لیست شده‌اند.

شاخص بینیت به اهمیت ارتباطی یک گره و نقش آن در ارتباط بین گره‌های مختلف اشاره دارد. به عبارتی هر گره به طور میانگین چند بار در مسیر ارتباطی بین دو گره شبکه قرار می‌گیرد. همان‌طور که از نام آن پیدا است شاخص بینیت، سر راه بودن بین گره‌های مختلف را نشان می‌دهد. گره‌ای که میزان بینیت آن بیشتر باشد، نقش انتقال اطلاعات بیشتری در کل شبکه بر عهده دارد. حذف این گره‌ها، به اختلال در شبکه یا از هم پاشیدگی مفهوم جاری در شبکه منجر می‌شوند. در جدول ۱ و ۲ ده سؤالی که بیشترین و کمترین مقدار بینیت را دارند، ارائه شده است.

جدول ۲. ده سؤال با بیشترین مقدار بینیت

رتبه	شماره سؤالات	بینیت
رتبه ۱	q118	۱۰۳
رتبه ۲	q39	۸۶
رتبه ۳	q24	۸۴
رتبه ۴	q98	۸۳
رتبه ۵	q69	۸۲
رتبه ۶	q49	۷۸
رتبه ۷	q22	۷۶
رتبه ۸	q132	۷۶
رتبه ۹	q133	۷۵
رتبه ۱۰	q12	۷۲

می‌سنجند. به عبارت دیگر راه بهینه‌دستیابی به مفهوم ارائه شده توسط کل سؤالات، سؤالاتی هستند که مقدار نزدیکی بالاتری دارند.

جدول ۴. ده سؤال با بیشترین مقدار نزدیکی

رتبه	شماره سؤالات	نزدیکی
رتبه ۱	q118	۰/۰۰۱۹۳۵۵
رتبه ۲	q69	۰/۰۰۱۹۳۰۵
رتبه ۳	q37	۰/۰۰۱۹۰۴۷
رتبه ۴	q98	۰/۰۰۱۹۰۲۲
رتبه ۵	q49	۰/۰۰۱۸۹۵۶
رتبه ۶	q133	۰/۰۰۱۸۸۲۶
رتبه ۷	q113	۰/۰۰۱۸۸۰۱
رتبه ۸	q39	۰/۰۰۱۸۷۹۸
رتبه ۹	q67	۰/۰۰۱۸۷۸۴
رتبه ۱۰	q41	۰/۰۰۱۸۷۷۸

جدول ۵. ده سؤال با کمترین مقدار نزدیکی

رتبه	شماره سؤالات	نزدیکی
رتبه ۱۳۱	q36	۰/۰۰۱۷۲۳۸
رتبه ۱۳۲	q9	۰/۰۰۱۷۱۶
رتبه ۱۳۳	q35	۰/۰۰۱۷۱۴۴
رتبه ۱۳۴	q124	۰/۰۰۱۷۱۴۳
رتبه ۱۳۵	q33	۰/۰۰۱۶۹۸۷
رتبه ۱۳۶	q5	۰/۰۰۰۷۲۹
رتبه ۱۳۷	q8	۰/۰۰۱۶۱۲۵
رتبه ۱۳۸	q21	۰/۰۰۱۵۸۲۴
رتبه ۱۳۹	q6	۰/۰۰۱۵۸۲۴
رتبه ۱۴۰	q38	۰/۰۰۱۵۸۲

سؤالات ذکر شده در بالا سؤالات مرکزی‌تر در شبکه سؤالات حجم لغات زبان انگلیسی می‌باشند. به عبارتی حذف این سؤالات بیشترین آسیب را بر انسجام درونی شبکه سؤالات می‌گذارد. از آنجایی که هر سؤال یک لغت را می‌پرسد، لغات دارای بالاترین بینیت می‌توانند مسیر یادگیری کل شبکه لغات را نیز نشان دهند. به عبارت دیگر ده سؤالی که می‌توانند به خزانه لغات بیشتر کمک نمایند، و ارتباط بین آنها را تبیین نمایند در لیست بالا مشاهده می‌شود.

جدول ۳. ده سؤال با کمترین مقدار بینیت

رتبه	شماره سؤالات	بینیت
رتبه ۱۳۱	q27	۲۵
رتبه ۱۳۲	q106	۲۵
رتبه ۱۳۳	q17	۲۴
رتبه ۱۳۴	q124	۲۴
رتبه ۱۳۵	q5	۲۲
رتبه ۱۳۶	q35	۲۰
رتبه ۱۳۷	q8	۸
رتبه ۱۳۸	q38	۷
رتبه ۱۳۹	q6	۵
رتبه ۱۴۰	q21	۴

گره‌های جدول ۳ در شبکه سؤالات، گره‌های پیرامونی محسوب می‌شوند و از مرکز شبکه دور هستند. به عبارتی حذف این سؤالات تأثیری زیادی در آزمون حجم لغات ندارند. به عنوان مثال سؤال ۲۱، سؤالی است که بعد از حذف اثر سایر سؤالات از روی ارتباطات آن، نقش تأثیرگذاری در آزمون خزانه لغات نسبت به سایر سؤالات ندارد.

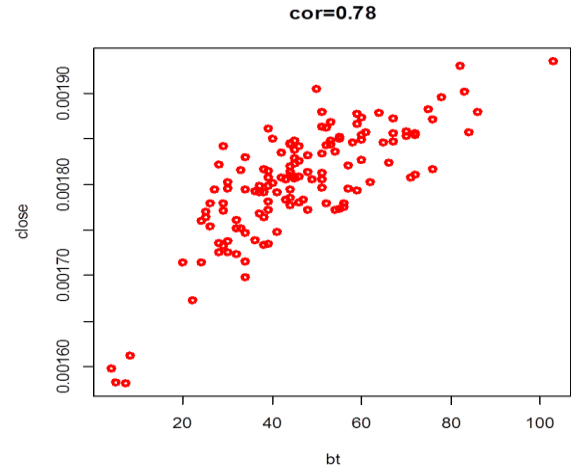
شاخص نزدیکی بیانگر این است که یک گره چگونه به صورت غیر مستقیم با سایر گره‌ها در ارتباط است. شاخص نزدیکی نشان می‌دهد که به طور متوسط کدام سؤالات نزدیکی بیشتری به کل سؤالات دارند. یعنی مجموع فاصله کدام سؤالات از همه سؤالات کمتر است. به لحاظ ریاضیاتی، میانگین کوتاه‌ترین مسیر هر سؤال از سایر سؤالات است.

سؤالات ۱۱۸، ۶۹، ۳۷ و ۹۸، ۴۹، ۱۳۳، ۱۱۳، ۳۹، ۶۷ و ۴۱، جزء ده سؤال با مقدار نزدیکی بالا می‌باشند. این ده سؤال می‌توانند نمایندگان خوبی از سؤالات حجم لغات باشند که می‌توانند حجم لغات را بسنجند. به بیان دقیق‌تر، از دید شرکت کنندگان در مطالعه، سؤالات با مقدار نزدیکی بالاتر، نمایندگان خوبی برای مفهومی هستند که سایر همه سؤالات

سؤالات ۳۶، ۹، ۳۵، ۱۲۴، ۸، ۲۱، ۵، ۳۳، ۶ و ۳۸ سؤالاتی هستند که کمترین مقدار نزدیکی را دارند. این سؤالات از مفهوم مشترکی که از سایر سؤالات درک می‌شود، جدا افتاده‌اند. عموماً سؤالاتی که دارای بینیت بالا هستند دارای نزدیکی بالا نیز می‌باشند و برعکس. در نمودار زیر مقدار همبستگی بین این دو شاخص و پراکنندگی مقادیر مربوطه نشان داده شده است.

جدول ۷. ده سؤال با کمترین مقدار استحکام

رتبه	شماره سؤالات	استحکام
رتبه ۱۳۱	q۳۵	۲۵/۲۷۴۰۱
رتبه ۱۳۲	q۱۲۰	۲۵/۲۱۴۶۶
رتبه ۱۳۳	q۳۳	۲۴/۶۴۲۹۶
رتبه ۱۳۴	q۲	۲۴/۵۰۱۶۳
رتبه ۱۳۵	q۹	۲۳/۷۲۱۰۷
رتبه ۱۳۶	q۵	۲۲/۱۸۲۷۸
رتبه ۱۳۷	q۸	۲۱/۸۸۱۰۶
رتبه ۱۳۸	q۳۸	۲۱/۶۴۳۳۳
رتبه ۱۳۹	q۶	۲۱/۵۸۹۲۴
رتبه ۱۴۰	q۲۱	۲۱/۳۸۶۳



شکل ۲. مقدار همبستگی بین شاخص نزدیکی و بینیت و پراکندگی مقادیر

شاخص دیگری به نام استحکام وجود دارد که به قدرت ارتباط مستقیم یک گره با سایر گره‌ها اشاره دارد. به لحاظ ریاضیاتی این شاخص مجموع وزن (شدت مطلق) رابطه هر گره را با سایر گره‌هایی که با آنها در ارتباط است مشخص می‌کند.

سؤالات ۲۱، ۶، ۳۸، ۸، ۵، ۹، ۲، ۳۳، ۱۲۰، ۳۵ کمترین مقدار استحکام را دارند. از آنجایی که در اینجا ارتباط متوسط خالص مورد نظر است، این سؤالات ارتباط مستقیم ضعیف با سایر سؤالات دارند.

شاخص چهارم، شاخص تأثیر مورد انتظار است. برای شناسایی بهتر گره‌های تأثیرگذار در شبکه زمانی که یال‌های شبکه وزن دارند (مانند شبکه موجود در این پژوهش)، لازم است تا بین یال‌های مثبت و منفی تمایز قایل شویم. بر خلاف سایر شاخص‌های مرکزی، هدف این شاخص سنجش ماهیت و قدرت کلی و تأثیر تجمعی (cumulative influence) درون شبکه می‌باشد و بنابراین انتظار می‌رود که این شاخص به نقش گره‌ها و امید تأثیر آنها در فعال‌سازی، پایداری و بهبود شبکه اشاره داشته باشد (روبین، میلنر، مک نالی، ۲۰۱۶).

جدول ۶. ده سؤال با بیشترین مقدار استحکام

رتبه	شماره سؤالات	استحکام
رتبه ۱	q۱۱۸	۳۰/۷۲۶۳۶
رتبه ۲	q۴۹	۳۰/۶۱۶۱۷
رتبه ۳	q۱۱۵	۳۰/۴۰۷۳۱
رتبه ۴	q۹۰	۳۰/۳۶۲۱۶
رتبه ۵	q۱۳۱	۳۰/۳۲۵۸۶
رتبه ۶	q۲۲	۳۰/۰۹۴۷۹
رتبه ۷	q۱۱	۳۰/۰۶۷۳۳
رتبه ۸	q۳۷	۲۹/۹۷۴۲۹
رتبه ۹	q۴۳	۲۹/۷۳۵۴۸
رتبه ۱۰	q۴۱	۲۹/۶۵۵۵۸

سؤالات ۱۱۸، ۴۹، ۱۱۵، ۹۰، ۱۳۱، ۲۲، ۱۱، ۳۷، ۴۳ و ۴۱ به ترتیب ده سؤالی هستند که بیشترین مقدار استحکام را دارند. به عبارتی ارتباط مستقیم این سؤالات با سایر سؤالات به صورت میانگین بیشترین بوده است. منظور از ارتباط مستقیم همان مقدار قدر مطلق همبستگی‌های تفکیکی تراکوریک است. در اینجا تعداد ارتباطات اهمیت ندارد.

جدول ۸. سؤالات با بیشترین مقدار تأثیر مورد انتظار

رتبه	شماره سؤالات	تأثیر مورد انتظار
رتبه ۱	q۹۵	۴/۸۲۷۴۹۲۶۲
رتبه ۲	q۱۰۹	۴/۴۸۲۰۹۲۳۶
رتبه ۳	q۹۸	۴/۴۵۹۹۳۸۶۶
رتبه ۴	q۱۰۶	۴/۳۲۶۴۴۱۸۷
رتبه ۵	q۵۶	۴/۲۶۸۶۵۷۱۲
رتبه ۶	q۲۱	۴/۱۱۹۱۹۲۸۸
رتبه ۷	q۵۰	۳/۹۰۱۳۸۱۵
رتبه ۸	q۲۴	۳/۸۷۰۴۶۲۹۸
رتبه ۹	q۵۷	۳/۷۱۳۱۳۶۸
رتبه ۱۰	q۳۸	۳/۵۴۰۶۱۲۱۲

سؤالات ۹۵، ۱۰۹، ۹۸، ۱۰۶، ۵۶، ۲۱، ۵۰، ۲۴، ۵۷ و ۳۸ ده سؤالی هستند که بیشترین تأثیر گذاری را بر شبکه مورد مطالعه دارند.

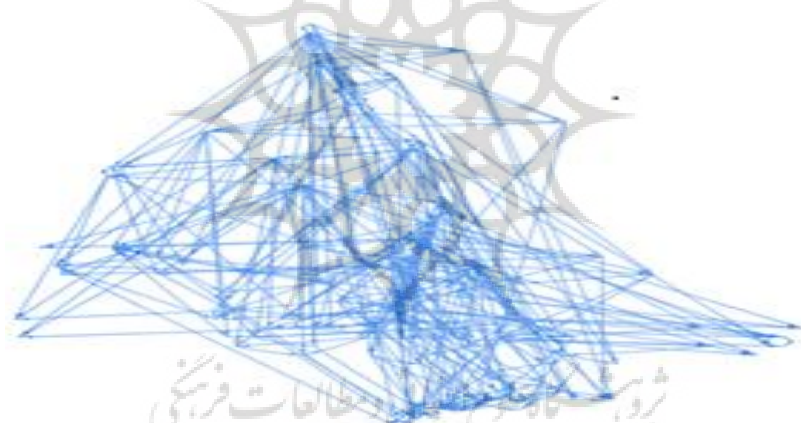
جدول ۹. سؤالات با کمترین مقدار تأثیر مورد انتظار

رتبه	شماره سؤالات	تأثیر مورد انتظار
رتبه ۱	رتبه ۱۳۱	q۳۰
رتبه ۲	رتبه ۱۳۲	q۱۱۵
رتبه ۳	رتبه ۱۳۳	q۱۲۵
رتبه ۴	رتبه ۱۳۴	q۱۳۲
رتبه ۵	رتبه ۱۳۵	q۳۳
رتبه ۶	رتبه ۱۳۶	q۹۴
رتبه ۷	رتبه ۱۳۷	q۳۶
رتبه ۸	رتبه ۱۳۸	q۸۹
رتبه ۹	رتبه ۱۳۹	q۴۰
رتبه ۱۰	رتبه ۱۴۰	q۲۵

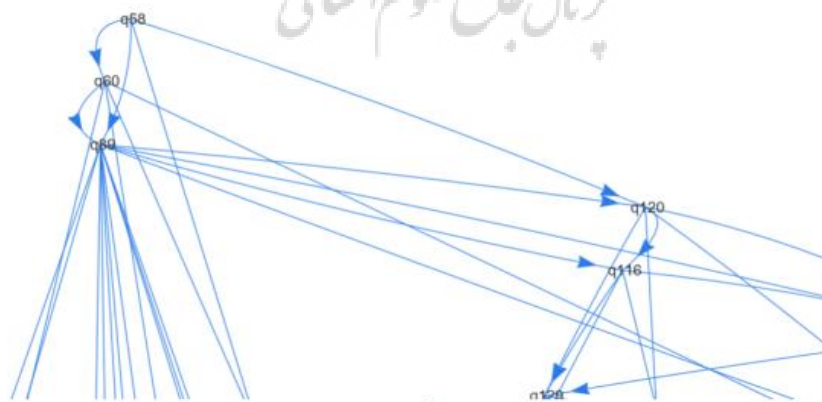
سؤالات ۲۵، ۴۰، ۸۹، ۳۳، ۹۴، ۳۶، ۱۳۲، ۱۲۵، ۱۱۵، ۳۰ سؤالاتی هستند که کمترین تأثیر را بر شبکه سؤالات مؤثر بر آزمون خزانه لغات دارند. به

عبارتی در پایداری شبکه و فعال سازی شبکه کمترین نقش را نسبت به سایر سؤالات دارند.

نقشه شبکه بیزی سؤالات به صورت کلی بر اساس الگوریتم تپه نوردی یا HC با استفاده از بسته‌های نرم‌افزاری bnlearn و bnviewer ترسیم شد. از آنجایی که این نقشه دینامیک است، فضای کاغذی امکان بررسی دقیق آن را به لحاظ دیداری، ندارد. در اینجا شبکه، بر اساس قواعد بیزین ترسیم شده است به همین دلیل یال‌ها دارای جهت‌اند و دوری در شبکه مشاهده نمی‌شود. لازم به ذکر است که در سؤالات بالا شبکه‌ها بر اساس الگوریتم فراچترمن-رینگولد که غیر بیزینی است ترسیم شده بودند تا نقشه کلی و اهمیت سؤالات در مفهوم خزانه لغات را نشان دهند. شبکه بیز رابطه والد-فرزندی سؤالات مختلف را مشخص می‌کند. به عبارتی مشخص می‌کند که والدین هر سؤال چه سؤال یا سؤالاتی می‌باشد. از آنجایی که قابلیت این تصویر با استفاده از کامپیوتر ممکن است در زیر بریده‌ای از رابطه والد-فرزندی بعضی از سؤالات نشان داده می‌شود.



شکل ۳. شبکه بیزین با الگوریتم HC در ارتباط کلی بین سؤالات آزمون حجم لغات



شکل ۴. بزرگنمایی بخشی از شبکه بیزین ارتباط بین سؤالات آزمون حجم لغات

در این شکل سؤال ۵۸ والد سؤالات ۶۰، ۸۹ و ۱۲۰ است. و سؤال ۱۲۰ فرزند سؤالات ۸۹ و ۵۸ است. در این مدل والد - فرزندی کلیه سؤالات مشخص می‌شود. به‌عنوان مثال سؤال ۵۸ والدی است که فرزند سؤال دیگری نیست. والد سؤال ۶۰ سؤال ۵۸ است، والد سؤال ۸۹ سؤالات ۵۸ و ۶۰ است و به همین ترتیب والد مستقیم سؤال ۱۲۲ سؤالات ۱۱۲، ۱۱۶ و ۱۳۴ می‌باشند. این مدل دارای ۱۴۰ گره و ۳۹۲ آرک یا یال جهت‌دار است که با الگوریتم HC بدست آمده‌اند. سایر مشخصات مدل در زیر ارائه شده است.

nodes:	140
arcs:	392
undirected arcs:	0
directed arcs:	392
average markov blanket size:	9.37
average neighbourhood size:	5.60
average branching factor:	2.80
learning algorithm:	Hill-Climbing
score:	BIC (disc.)
penalization coefficient:	3.674937
tests used in the learning procedure:	1432673
optimized:	TRUE

بحث و نتیجه‌گیری

هدف از پژوهش حاضر کاربرد علم شبکه در ترسیم شبکه ارتباطی سؤالات آزمون خزانه لغات انگلیسی نیشن و وارینگ (۱۹۹۷) و معرفی تحلیل داده‌های شبکه‌ای و رسوب دانش در این حوزه بود. هم ارتباط بین سؤالات از طریق گراف ارتباطی آنها، نیازمند دانش فنی زیاد درباره نظریات اندازه‌گیری و روش‌های تحلیل سؤال که در حال حاضر مرسوم است نمی‌باشد و می‌تواند در تحلیل سؤالات آزمون به خوبی استفاده شود. شبکه‌بیزی به خوبی رابطه علی را نمایش می‌دهد، دستگاه ریاضیاتی شبکه‌بیزی به خوبی گسترش یافته است و نرم‌افزارهای خوبی در این حوزه بوجود آمده است (لیتوینکو و لیتوینکو، ۲۰۱۷).

شبکه‌های بیزی یک مدل احتمالاتی گرافیکی است که مجموعه‌ای از متغیرهای تصادفی و وابستگی‌های شرطی آنها را از طریق گرافی جهت‌دار و بدون دور نشان می‌دهد (بنگال، ۲۰۰۷). از این مدل‌ها برای تصمیم‌گیری در شرایطی که عدم اطمینان بر آن حاکم است استفاده می‌شود. توانایی این مدل‌ها صرفاً به استدلال در شرایط نامطمئن که در آن صدها متغیر وجود دارد، نیست بلکه برای کمک به افراد با هدف فهم حوزه مورد نظر نیز استفاده می‌شوند و دلیل آن توانایی شبکه در نشان دادن وابستگی‌های بین

متغیرها است. در آزمون‌های روانی سؤالات مختلف آزمون به عنوان متغیرهای مختلف می‌باشند که شبکه‌بیزی قابلیت شناسایی وابستگی بین سؤالات را دارد. از طریق روش d-separation در این شبکه‌ها می‌توان گزاره‌های استقلال شرطی را بدست آورد. هر یال در شبکه‌بیزی ترسیم شده از شدت یکسان برخوردار نیست، هر یال دارای قدرت یا شدت مشخص است. خطوط بین گره‌های مختلف اطلاعات مناسبی را به منظور تحلیل محتوای سؤالات پرسشنامه ارائه می‌دهند. نقشه ارتباطی بین سؤالات می‌تواند درک بهتری از آنها را فراهم آورد و نزدیکی و دوری سؤالات از یکدیگر را می‌توان به لحاظ محتوایی بررسی کرد. استفاده از شیوه‌های مبتنی بر نظریه گراف سرعت بیشتری در پیدا کردن سؤالات نامناسب دارد و می‌توان از طریق آن ماتریس ارتباط بین سؤالات را بصورت کاملاً واضح بصری نمود و اطلاعات زیادی را بصورت یکجا ارائه نمود. از آنجایی که ورودی‌های شبکه ماتریس‌های مجاورت هستند می‌توان از طریق همبستگی‌های تفکیکی و نیمه تفکیکی به درک درستی از اهمیت سؤالات مختلف رسید. استفاده از الگوریتم فراچترمن - رینگولد در ترسیم شبکه ارتباط بین سؤالات مختلف به گرافی منجر می‌شود که در آن علاوه بر اینکه فواصل بین سؤالات مختلف معنادار است قطر یال‌ها، ابعاد زیربنایی نقشه، ساختار و خوشه‌های تشکیل شده نیز می‌توانند در تحلیل سؤالات استفاده شوند و امکان بعدیابی را حداقل در فضای دو بعدی فراهم آورد. از جمله محدودیت‌های پژوهشی می‌توان به کمبود مطالعات پژوهشی در ارتباط با موضوع این مقاله، فقدان منابع فارسی در حوزه مورد پژوهش و دشواری دسترسی به مقالات و اسناد خارجی و مشکلات اجرایی پروژه‌های دانشگاهی که مربوط به فرهنگ بروکراتیک دانشگاهی است اشاره کرد. نتایج بدست آمده از گراف‌های بالا قابل مقایسه با شاخص‌ها و پارامترهای بدست آمده از طریق روش‌های مرسوم تحلیل سؤال می‌باشند. استفاده از نظریات کلاسیک و سؤال پاسخ اندازه‌گیری در کنار تحلیل عاملی جزو روش‌های مرسوم تحلیل سؤال به شمار می‌رود. با توجه به مغفول ماندن علم شبکه در حوزه آزمون‌سازی در ایران انجام مطالعات مشابه و پژوهش‌هایی به منظور مقایسه گراف سؤالات پرسشنامه حاضر و نتایج حاصل از تحلیل داده‌های شبکه‌ای با ویژگی‌های روانسنجی که بر اساس شیوه‌های مرسوم و کلاسیک بدست می‌آیند پیشنهاد می‌شود.

ملاحظات اخلاقی

پیروی از اصول اخلاق پژوهش: این مقاله برگرفته از رساله دکتری نویسنده اول در رشته روانشناسی تربیتی از دانشگاه تبریز و تاریخ تصویب پروپزال ۱۳۹۸/۰۳/۲۱ است. قبل از اجرای آزمون ضمن ارائه توضیحات لازم به شرکت کنندگان در خصوص اهمیت و هدف پژوهش، این فرصت به آنها داده شد که سؤالات خود را در زمینه پژوهش مورد نظر مطرح نمایند، همچنین قبل از اجرای آزمون به عدم اجبار و داوطلبانه بودن شرکت در پژوهش و نیز به محرمانه ماندن اطلاعات شخصی داوطلبان و عدم استفاده از این اطلاعات در جای دیگر، اشاره شد و به منظور جلب اعتماد و غیر قابل شناسایی بودن شرکت کنندگان، از اخذ نام و نام خانوادگی آنها خودداری گردید.

حامی مالی: این پژوهش در قالب رساله دکتری و بدون حمایت مالی می‌باشد.

نقش هر یک از نویسندگان: نویسنده اول محقق اصلی این پژوهش است. نویسندگان دوم و سوم استادان راهنما و نویسنده چهارم استاد مشاور رساله می‌باشند.

تضاد منافع: نویسندگان هیچ تضاد منافی در رابطه با این پژوهش اعلام نمی‌نمایند.

تشکر و قدردانی: بدین وسیله از اساتید راهنما و مشاور و کلیه شرکت کنندگان در این پژوهش تشکر و قدردانی می‌گردد.



پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی

References

- Almond, R. G., & Mislevy, R. J. Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian Networks in Educational Assessment. Statistics For Social & Behavioral Sciences*. New York, NY. DOI: 10.1007/978-1-4939-2125-6. [\[link\]](#)
- Araghi, L., Taheri, A., Joz Ramazani, Z., Abas Pour, A., & Zarghami, M.H. (2017) Study of applying network data analysis in item analysis. *Quarterly of Educational Measurement*. 7(28), 207-240. DOI: 10.22054/jem.2017.2787.1086. [\[Link\]](#)
- Ben-Gal, I. (2007) Bayesian networks. In: Ruggeri F, Faltin F, Kenett R (eds) *Encyclopedia of statistics in quality & reliability*. Wiley, London. DOI: 10.1002/9780470061572.eqr089. [\[link\]](#)
- Borsboom, D. (2008). Psychometric perspectives on diagnostic systems. *Journal of Clinical Psychology*, 64, 1089–1108. DOI: 10.1002/jclp.20503. [\[link\]](#)
- Borsboom, D., & Cramer, A.Q. (2013). Network Analysis: An Integrative Approach to the structure of Psychopathology. *Annual review of Clinical Psychology*, 9, 91-121. DOI: 10.1146/annurev-clinpsy-050212-185608. [\[link\]](#)
- Buch-Hansen, H. (2014). Social network analysis and critical realism. *Journal for the Theory of Social Behaviour*, 44(3), 306–325. DOI: 10.1111/jtsb.12044. [\[link\]](#)
- Chalabi, M. (1995). Network Analysis in Sociology. *Journal of Social Science*, 3(5-6), 9-48. DOI: 10.22054/qjss.1995.5152. [\[Link\]](#)
- Culbertson, M.J. (2015). Bayesian Networks in Educational Assessment: The State of the Field. *Applied Psychological Measurement*, 40(1) 3–21. DOI: 10.1177/0146621615590401. [\[link\]](#)
- Dörnyei, Z. (2014). Researching complex dynamic systems: ‘Retrodictive qualitative modelling’ in the language classroom. DOI: 10.1017/S0261444811000516. [\[link\]](#)
- 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. DOI: 10.18637/jss.v048.i04. [\[link\]](#)
- Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3-5), 75-174. DOI: 10.1016/j.physrep.2009.11.002. [\[Link\]](#)
- Litvinenko, A., & Litvinenko, N. (2017). Application of Bayesian Networks for Estimation of Individual Psychological Characteristics. *Przegląd Elektrotechniczny*. 1(5). DOI: 10.15199/48.2019.05.23. [\[Link\]](#)
- Millán, E., Laboda, T., Pérez-de-la-Cruz, J. L., & García, F. (2010). Bayesian networks for student model engineering. *Computers and Education, Language Teaching*, 47(1), 80-91.55(4): 1663-1683. DOI: 10.1016/j.compedu.2010.07.010. [\[Link\]](#)
- Nation, I.S.P., Beglar, D. (2007) A Vocabulary Size Test. *The Language Teacher*, 31(7), 9-13. [\[Link\]](#)
- 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. DOI:10.4236/ojml.2015.53023. [\[Link\]](#)
- Reckase MD (2009). *Multidimensional Item Response Theory*. Springer-Verlag, New York. DOI:10.1007/978-0-387-89976-3. [\[Link\]](#)
- Robinaugh, D. J., Millner, A. J., & McNally, R. J. (2016). Identifying highly influential nodes in the complicated grief network. *Journal of Abnormal Psychology*, 125(6), 747–757. DOI: 10.1037/abn0000181. [\[Link\]](#)
- Rust, J., Golombok, S. (2009). *Modern Psychometrics. The Science of Psychological Assessment*, (3rd ed.). New York: Routledge. [\[Link\]](#)
- Saarinén, 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 (LAK '20). Association for Computing Machinery, New York, NY, USA, 422–427. DOI: 10.1145/3375462.3375541. [\[Link\]](#)
- Tavana, M. A. (2015). The Paradigm of Critical Realism Interdisciplinary Methodology. *Journal of Interdisciplinary Studies in the Humanities*. 7(1), 27-56. DOI: 10.7508/isih.2015.25.002. [\[Link\]](#)