

Analysis of Scientific Studies on Item Response Theory by Bibliometric Analysis Method

Gökhan Aksuⁱ

Adnan Menderes University

Cem Oktay Güzellerⁱⁱ

Akdeniz University

Abstract

The purpose of this study is to analyze the studies, which include Item Response Theory among the keywords, available in the Web of Science database between 1980-2018 through bibliometric analysis method. A total of 1,367 academic works has been analyzed. The authors, journals and countries having the highest number of studies in the field and their interrelations on the network in terms of collaboration have been determined through common citation analysis performed using Citespace II software. In addition, a word analysis was also conducted to determine most frequently used concepts in the field. As a result of the study it was found that the authors that have made the biggest contribution to the field are De Ayala, Embretson, Reckase, Reise and Chalmers; in addition, the countries making the biggest contribution are respectively US, Netherland, Canada, Spain and China. The number of citations that US got, which is the country that received the highest number of citations with 687 citations, is 7 times higher than Netherland, which is the second most cited country. Moreover, it was found that the journals that were mostly cited are respectively Psychometrika, Appl Psych Measurement, Item Response Theory, J Edu Measurement and Educ Psychol Measurement. As a result of the word analysis based on most repeated words, which was performed for the purpose of determining most popular subjects on the field, it was found that most frequently used words are item response theory, classical test theory, model, validation, reliability, validity and Rasch model

Key words: Item Response Theory, Citespace II, Bibliometric Analysis, Cite, Analysis

DOI: 10.29329/ijpe.2019. 189.4

ⁱ **Gökhan Aksu**, Instructor Dr., Adnan Menderes University, Vocational High School

Correspondence: gokhanaksu1923@gmail.com

ⁱⁱ **Cem Oktay Güzeller**, Prof. Dr., Akdeniz University, Tourism Faculty

INTRODUCTION

The foundation of social network analysis goes up to 1960's, but with the methodic and software improvements it became one of the most widely discussed domains again, especially in the last ten years (Barabasi, 2002; Watts, 2004). Social networks have become extremely popular in the past ten years with the significant studies conducted mostly in US after the second half of 20th century, accomplishing to attract the attention of scientists from various disciplines, including sociology, epidemiology, economics, computer engineering, telecommunication, communication and others (Dorogovtsev and Mendes, 2002; Wasserman and Faust, 2009).

Social network analysis is an interdisciplinary research area built on the theoretical bases obtained from sociology, anthropology, statistics, mathematics, information sciences, education, psychology and other disciplines over a long time period (Duijin and Vermunt, 2006). Even though the pioneers of social network analysis, such as Moreno (1946), Cartwright (1959) and Barnes (1954) were from origin sociology and social psychology, social network analysis has been renewed after a while using the combined experiences of sociology and graph theory (Mincer and Niewiadomska-Szynkiewicz, 2012). Graph theory, which is used to measure some characteristics of a network by determining its structural properties mathematically (Barnes and Harary, 1983), plays a very important role in social network analysis. In a more generic sense, graph theory is accepted as one of the best and most accurate ways of identifying and representing the structure of a relationship (or some relationships) that connects different members of a social group. Thus, the characteristics of the relationships among the individuals and organizations located in a network can be easily understood (Wasserman and Robbins, 2005). These connections sometimes show the relationships such as affinity and friendship; similarities such as being together and being in the same group; interactions between individuals such as collaboration and communication; or flow of resources such as knowledge and material. In this stage, the members of the network may take different shapes such as individuals, small groups, cities, etc. (Borgatti and Halgin, 2011).

Social network analysis aims to explain, visualize and understand the network structure obtained from the relationships among individuals, objects or units through statistical modelling. Many systems in the nature and in the technology are examples of social network (Crossley, Prell and Scott, 2009). Almost all activities, such as the e-mails that you received, contamination process of a disease, criminal activities, airline flight routes, etc., can be modelled as a social network (Jamali and Abolhassani, 2006). Although the concept of social network was born from the network idea in electric and transportation engineering, the origin of this concept comes from sociology (Freeman, 2004). After the definition of Wasserman and Faust (1994), social network data is seen as a set of actors and the social relationship system characterized by their social connections.

Social network analysis focus on the individuals and other social units, as well as the relationships between them (Martino and Spoto, 2006). Social network analysis is an inter-disciplinary research area aiming to predict the structure of the relationships between social entities and the impact of the relevant structure on other social phenomenon (Butts, 2008). Social network analysis method can be applied in many areas, from the analysis of cognitive processes inside the human mind to the analysis of the wars between countries (Wimmer and Min, 2006). Social network analysis may be a strong tool for psychologists for defining and modelling the relational connections where the behavior occurs and relational dimensions of this behavior. Social network analysis allows researchers to analyzes social structures that are naturally present, such as examining human behaviors (McPherson, Smith-Lovin and Cook, 2001). In addition, social network analysis is a valuable approach for examining the social structure that governs both intragroup and intergroup relationships and the processes lying under this structure. Each unit, called as social actor in a network, may be a person, a group, an organization, a firm, an author or a journal according to the type of the study. Nondirectional connections or directional flows between these units represent the relationships (Wasserman and Faust, 2009). The network structure possesses omni-directional or bi-directional connections among all individuals going from one to the other. If you have limited information about some of the actors in the network or there is no connection between them, the obtained network structure is called as an

ego-centric network (Borgatti, Everett and Johnson, 2013). With social network analysis, which aims to reveal the information present in different databases, researchers may interpret the data that has been produced by different sources more objectively, however they may fail to indicate indirect or intermediate connections between units or individuals (Wölfer, Faber and Hewstone, 2015). Therefore, it is possible to examine the structures at different levels of the social network and investigate their effects.

Visual representation of social networks is quite important in terms of understanding the data in the network and interpreting the results of the analysis more easily (Hogan, Carrasco and Wellman, 2007). Most of the software developed for this purpose have various modules for the visualization of the network. The discovery of the data at hand, the display of the nodes and connections in different designs are realized by visualizing them in different shapes according to their colors, dimensions and other advanced features. Bibliometric analysis is used to statistically and visually interpret the overall picture of a certain discipline and the results obtained by analyzing the scientific works published in this area.

Bibliometric analysis, which has been first emerged in 1917, became popular while famous English scientist Allen Richard has proposed the term “bibliometrics” instead of “statistical bibliography” in 1969 (Liao, Tang, Luo, Li, Chiclana and Zeng, 2018). The nature and direction of the scientific communication in scientific works can be statistically modeled based on the citations made to other sources and the bibliography of the work using bibliometric techniques. For example, bibliometric analysis is used for mapping the relationships between the journals and other scientific communication channels among cited works and determine the flow of the topics between disciplines (Borgman, 1999). In this way, it is possible to determine the cited articles; the scientists who cited them and their disciplines; the journals that are cited more frequently; and the impact of certain articles on subsequent researches through citation analysis (Tsay, 2011). According to the first definition made by Pritchard (1969), bibliometric analysis has been defined as “the application of statistical and mathematical methods to the books and other communication tools”. In bibliometric analysis, which has been defined by Van Leeuwen (2004) as quantitative measurement of qualitative characteristics, the number of citations made to an article is accepted as an indicator of the impact of the article on the scientific community. Bibliometric analysis provides the opportunity to reveal current situation through the analysis of the data belonging to the publications (Martinez, Cobo, Herrera and Herrera-Viedma, 2015). Bibliometrics is involved with the quantitative analysis of particular characteristics of the publications, such as author, subject, publication information, cited sources, etc. Quantitative analysis and statistical techniques are used in bibliometric analysis to define publication patterns in a certain area or in the literature (Abdi, Idris, Alguliyev and Aliguliyev, 2018). Based on the bibliometric data obtained in this way, the establishment of communication process in various disciplines can be investigated. In addition, researchers can identify the impact of a sole author or use bibliometric assessment methods to define the relationship between two or more authors or works (Jain, et al., 2015).

Regarding the history of measurement, there are various globally accepted approaches, including in chronological order Classical Test Theory (CTT), Generalizability Theory (GT) and then Item Response Theory (IRT) (Crocker and Algina, 1986). CTT applications are easier than other measurements theories and the operation load that they require is low, non-complex, thus they were widely used for years and they are still being used (Sünbül and Erkuş, 2013). Item Response Theory, which has been emerged to eliminate some limitations of CTT, is seen as a more advanced theory compared to CCT and it became more popular and preferable in the recent years (Reise, Ainsworth and Haviland, 2005). It is accepted that Item Response Theory provide solution to the problems encountered in test development, creating item pool, developing individual tests, determining item bias, weighting answer options and test equivalence (Hambelton and Swaminathan, 1985). The use of IRT-based models on the large-scale exams such as PISA shows the importance of this theory in terms of measurement and assessment. In PISA exam, especially in Science, Mathematics and Reading tests, plausible values are calculated through IRT-based posteriori distribution curves obtained from the answers that students have given to the other items of the questionnaire (OECD, 2017). For this

reason, IRT became one of the most discussed subjects in measurement and assessment area, especially in the recent years. This study aims to analyze all the articles published on this topic between 1980-2018 in Web of Science database and to determine the authors, journals and subject fields coming to the forefront. By this means, the authors and journals that have performed important works in item response theory were determined, and the popular subjects from past to present were identified. In addition, the countries of the researchers who have worked in this area were determined and the impact of country variable on the performed works was revealed. The study also focused on identifying the words that are frequently repeated in the academic works published in WoS database without putting any constraint, aiming to identify the trendy subjects in the item response theory field.

Problem statement of the research: What are the authors, journals, countries and subject fields that were effective in the works performed between 1980-2018?

The sub-problems addressed within the scope of the study are listed below:

1. Who are the most cited authors among the ones working on item response theory?
2. What are the clusters obtained according to the authors who were cited in item response theory?
3. What are the countries that worked on item response theory?
4. What are the most cited journals concerning item response theory?
5. What are the most current topics in item response theory?

METHODOLOGY

In this study, scientific works published in Item Response Theory area and included in Web of Science (WoS) database were analyzed through bibliometric analysis. Since the study aims to describe all characteristics of an existing situation, it is considered as a scanning model (Frankel and Wallen, 2006). It was also defined as cross-sectional type research because it involves descriptive analysis of the academic works published between 1980-2018, found in WoS database (Büyüköztürk, Kılıç-Çakmak, Erkan-Akgün, Karadeniz and Demirel, 2016).

Data Collection Process

As a result of the scanning made in Web of Science Core Collection database for the title “Item Response Theory”, the following data were obtained for 1,367 works published in the field: publication years, publication type, publication language, title, country of the authors and number of citations they have received from the sources scanned in Web of Science database. The academic works obtained after this operation were recorded into 3 different data files, each containing 500 scientific works. The distribution of the works according to years is shown in Table 1.

Table 1. Distribution of Publications by Years

Years	Frequency (N)	Percentage (%)
1980 – 1989	63	4.61
1990 – 1999	105	7.68
2000 – 2009	390	28.53
2010 – 2018	809	59.18
Total	1,367	100

The analysis of 1,367 academic works according to type showed that the majority of the works are articles. The distribution of the works according to type is shown in Table 2.

Table 2. Distribution of Publications According to Type

Type of Publication	f	%	Type of Publication	f	%
Article	1,029	75.27	Correction	11	0.80
Meeting summary	151	11.05	Software critics	5	0.37
Proceedings book	122	8.92	Book section	2	0.14
Book Review	55	4.02	Letter	2	0.14
Editorial text	15	1.10	Note	2	0.14
Critics	12	0.88	Discussion	1	0.07

After determining the academic works performed in item response theory area, data analysis was started.

Data Analysis

First of all, academic works obtained in electronic environment were analyzed in terms of their author, title, summary and source. Then, all bibliometric data, including the name of the authors and publications, title, source of the document, publication year, number of publications, number of citations, and type of the article were gathered together and saved as text document (with .txt extension). The data was transferred to Citespace II software and was analyzed to obtain results for the specified purposes. Citespace II is a Java application, which can be accessed free of charge, used to visualize and analyze the trends and models in the literature (Chen, 2006). Citespace II is a software that facilitates researchers' qualitative and quantitative work on scientific subject areas (Liu, Liu and Zheng, 2014).

The software especially focuses on finding the intellectual milestones and critical points in the development of a subject area or a discipline. CiteSpace software performs structural and temporal analysis of various networks obtained from scientific publications, such as collaboration networks, author associations and publication association (Synnestevedt, Chen and Holmes, 2005). The results obtained from Citespace II are in two different forms; cluster view and time-zone view. In time-zone view, the variation of common citations over time is visualized, whereas cluster view focuses on cluster divisions obtained from common citations over the defined time interval (Liu and Shen, 2013). The following steps were followed during the analysis of the data in Citespace software:

1. Based on bibliometric records, a new project file was formed.
2. Data files to be analyzed were loaded into the program.
3. The time period to be analyzed was defined by entering 1980 and 2018 as the starting and ending date.
4. For each time interval, threshold level was set as mostly cited 30 works.
5. "Cited references", "cited author" and "cited journal" options were activated for the nodes to be obtained as a result of the analysis.
6. Analysis results were separately reported in the form of cluster view and time zone view.

The names of some authors, journals and countries are represented by circle, line or color. The circle around the examined author, journal or country indicates citation history of a particular

reference, whereas the thickness of the circle shows the number of citations over a defined time period. The number of references increases as the circle gets bigger. The line between two circles indicates common reference sources, present in both citations. The thickness of this line shows the strength of the common citation, whereas its color indicates the time of the common citation (Liu and Shen, 2013).

FINDINGS

In the study, the findings concerning bibliometric analysis of the works about Item Response Theory included in Web of Science (WoS) database between 1980-2018 have been obtained first. In this regard, during common citation analysis of Item Response Theory, a total of 21,977 reference data belonging to 1,367 publications, in the form of article, discussion, book section and research note, have been analyzed. The modularity value of the obtained network was calculated as 0.833 and mean silhouette value as 0.294. According to mean silhouette value, which indicates the similarity of the elements in a cluster, it is observed that academic works included within the scope of the study are well-clustered. High modularity value means that there are strong connections among the nodes in the modules but the connection between the nodes of different modules are sparse (Yang, 2008). The distribution of the academic works performed on Item Response Theory between 1980-2018 according to years is shown in Figure1.

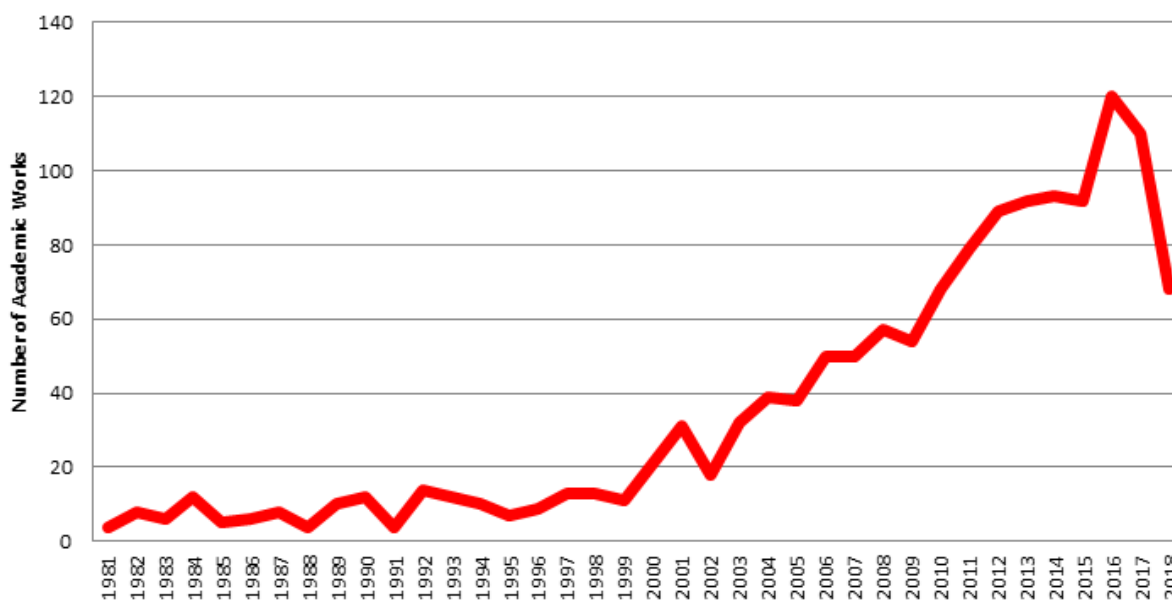


Figure 1. Distribution of the works in WoS database according to years

Regarding Figure 1, it can be seen that there is an increase on the number of works performed on IRT especially in 2000's. It was found that after 2000, the number of performed works was constantly increased, with a monotonous increase up to present.

Publications' Common Citation Network

Figure 1 shows the network of the works with 15 or more citations. Since the threshold value was set as 15, less cited works were not included in the network.

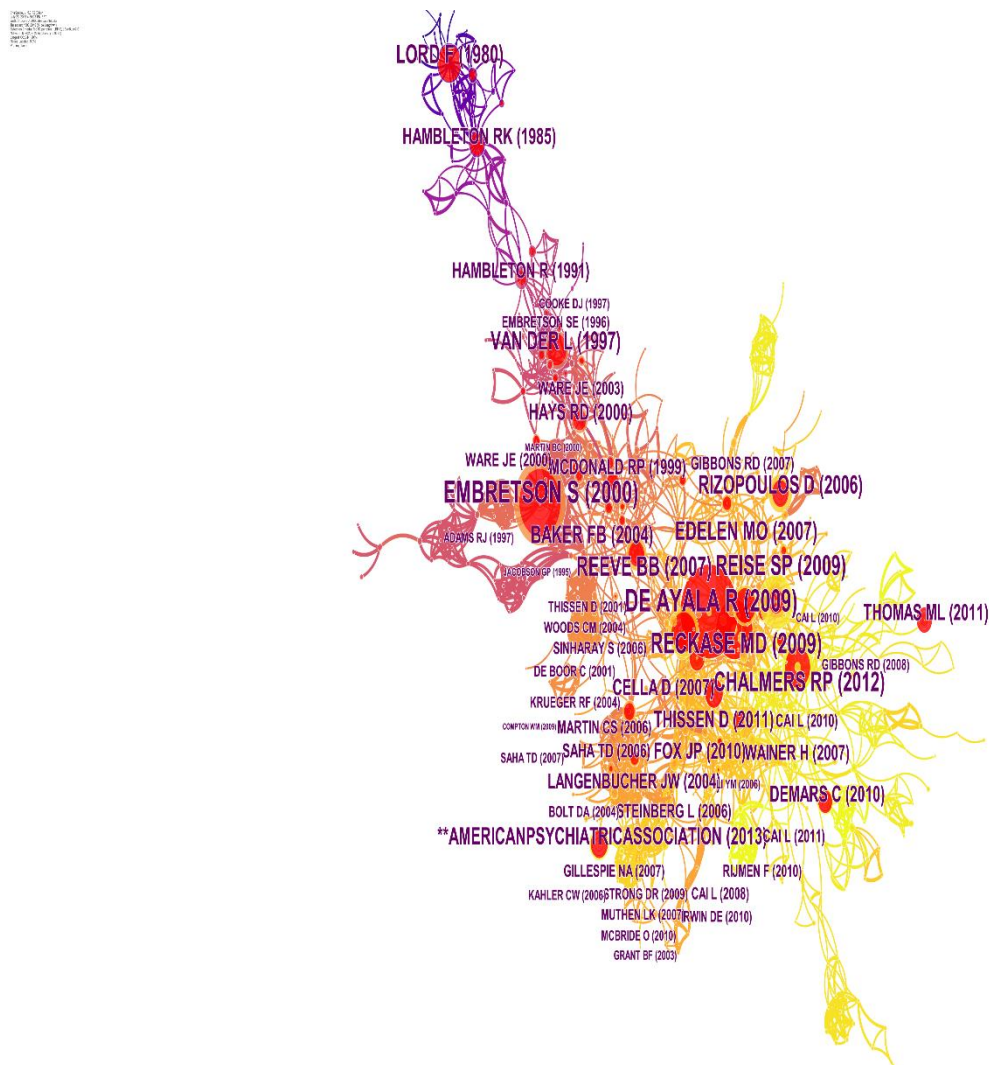


Figure 1. Network Structure of the Publications with Common Citation

The network shown in Figure1 has a total of 802 nodes and 2,295 connections. The density of the network was found to be 0.0071, Modularity value is 0.8336 and Mean silhouette value is 0.3004. 1,346 academic works were divided into 123 clusters. The citation values of top 10 works and their clusters are displayed in Table 1.











Table 1. Top 10 sources with highest citation

Number of Citation	Source	Cluster#
71	de Ayala R, 2009, THEORY PRACTICE ITEM, 0, 0	3
64	Embretson S, 2000, ITEM RESPONSE THEORY, 0, 0	11
45	Reckase MD, 2009, STAT SOC BEHAV SC, 0, 1	3
41	Reise SP, 2009, ANNU REV CLIN PSYCHO, 5, 27	3
38	Chalmers RP, 2012, J STAT SOFTW, 48, 1	12
37	Reeve BB, 2007, MED CARE, 45, 0	9
30	Edelen MO, 2007, QUAL LIFE RES, 16, 5	3
30	Lord F, 1980, APPLICATIONS ITEM RE, 0, 0	4

29	van der L, 1997, HDB MODERN ITEM RESP, 0, 0	2
29	Rizopoulos D, 2006, J STAT SOFTW, 17, 1	3

According to the data in Table 1, the most cited source is the work performed by de Ayala, R. (2009), belonging to cluster 3, with 71 citations. Second most cited work is the work performed by Embretson, S. (2000), belonging to cluster 11, with 64 citations. Third most cited work is the work performed by Reckase, M.D. (2009) with 45 citations. It can be seen that this work and the works of de Ayala, R. (2009), Reise, S.P. (2009), Edelen, M.O (2007) Van der L. (1997) and Rizopoulos (2006) belong to the same cluster. Hence, it can be seen from Figure 1 that these works are positioned very close to each other on the 2-dimensional map. Centrality values could not be calculated due to the intensity of the network. But as can be seen from the network, the works performed by de Ayala, Reckase, Embretson, Reise and Chalmers are the fundamental works that get highest citation in item response theory area. In addition, the results of burst analysis, which have been performed to see the most popular years of the works performed by different researchers, are shown in Table 2. Top 10 bursts among 55 citation bursts obtained from the analysis are shown in the Table.

Table 2. Citation Burst Values of the Publications Cited by the Works in the Database

Burst	Source	Cluster#	Start	End	1980-2018
28.75	Embretson S.	11	2002	2008	
20.20	Lord F.	4	1980	1988	
14.85	van der L.	2	1998	2005	
13.73	Chalmers RP.	12	2016	2018	
13.27	Hambleton RK.	4	1987	1993	
13.24	de Ayala R.	3	2012	2018	
11.20	Hambleton R.	2	1992	1999	
10.50	Rizopoulos D.	3	2010	2014	
9.29	Reeve BB.	9	2010	2015	
8.82	Baker FB.	7	2006	2012	

As can be seen from Table 2, the work with the highest citation burst value is the work performed by Embretson S. (2000), having a value of 28.75, belonging to cluster 11, between 2002 and 2008. It is followed by the work performed by Lord, F. (1980) with 20.20 citation burst value. It can be seen that this work attracted attention until 1988, but it did not get much citations in the following years. The works with the third highest burst value with 14.85 has been performed by Van der L. (1997), which was popular from 1998 to 2005 but it lost its importance in the following years. The works performed by Chalmers (2012) has the fourth highest citation burst value and it was very popular between 2016-2018. The results displayed in Table 2 give information about the popularity of the works according to years.

After this operation, cluster analysis was performed to determine the clusters using common citation network. The topic clusters obtained from the network are displayed in Figure 2.

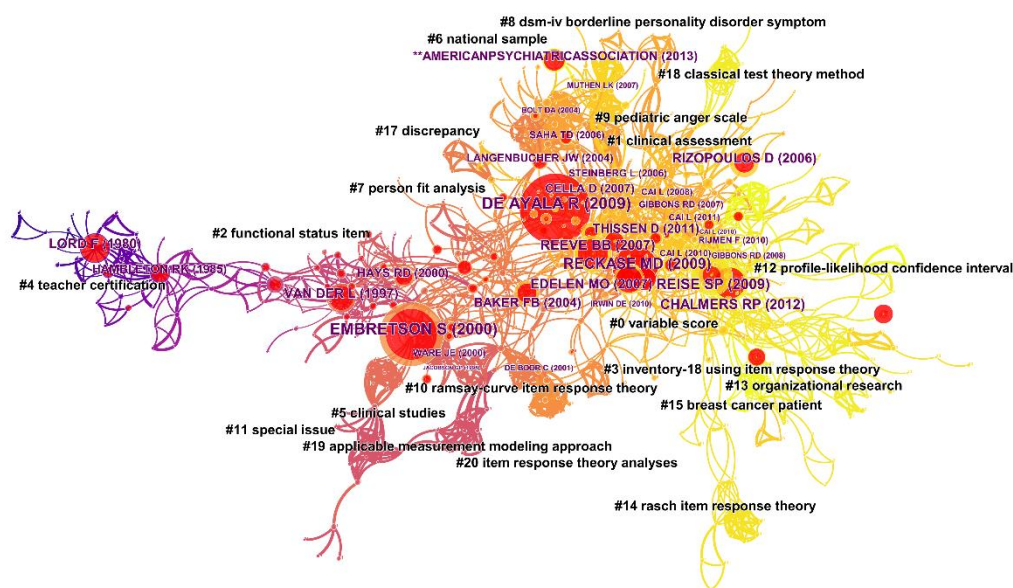


Figure 2. Clusters Obtained Based on the Works that Received Common Citations

As can be seen from Figure 2, the works performed by Lord, F. (2000) and Hambleton, R.K. (1985) belong to the cluster number 4, called as teacher certification. Similarly, the work called American Psychiatric Association (2013) belong to the cluster number 6, called as national sample. The names of top 7 clusters obtained from Citespace software using different clustering algorithms, their mean silhouette values and cluster sizes are displayed in Table 3.

Table 3. Statistics of the Clusters Obtained as a Result of Common Citation Analysis

Cluster	Size	Mean Silhouette	TFIDF	LLR	MI	Citation Year
0	67	0,752	item response theory	variable score	psychometric properties, short form	2009
1	66	0,735	item response theory	clinical assessment	psychometric properties, short form	2007
2	51	0,900	item response theory	functional status item	self-report measure, adult attachment	1994
3	46	0,838	item response theory	Inventory, using item response theory	psychometric properties, short form, self-report measure	2007
4	41	0,966	item response theory	teacher certification	psychometric properties, short form, self-report measure	1983
5	40	0,850	item response theory	clinical studies	psychometric properties, short form, self-report measure	2000
6	38	0,911	item response theory analysis	national sample	psychometric properties, short form, self-report measure	2004
7	37	0,737	item response theory	person fit analysis	psychometric properties, short form, self-report measure	2001

TFIDF: Term Frequency by Inversed Document Frequency, LLR: Log-Likelihood Ratio, MI: Mutual Information

Mean silhouette value, which varies between -1 and +1, shows the extent of an object belonging to its own cluster and high silhouette value means that the object is strongly matched with its own cluster, and weakly with the neighbor cluster. Accordingly, it can be seen that cluster 6, 4, 2, 5, 3, 0, 7 and 1 have more homogenous structures compared to the others. In addition, LLR was found to be the most suitable method in determining the names of the clusters (Rousseeuw, 1987).

Author Common Citation Network

Authors with 50 or more citations are displayed in the network that represents the important authors worked on item response theory. The network consists of 361 nodes and 2,401 connections. The density of the network was computed as 0.0369, modularity values as 0.3898 and mean silhouette value as 0.2315. It was found that there are 29 clusters in the author's network. The authors having top 10 highest number of citations are displayed in Table 4.

Table 4. Mostly Cited Authors

Number of Citations	Source	Cluster number#
421	Lord FM, 1981, SO, 0, 0	0
392	Thissen D, 1984, SO, 0, 0	0
380	Samejima F, 1987, SO, 0, 0	0
314	Embretson S, 1986, SO, 0, 0	2
291	Hambleton RK, 1982, SO, 0, 0	0
281	Bock RD, 1982, SO, 0, 0	4
231	Reise SP, 1993, SO, 0, 0	4
220	Baker FB, 1984, SO, 0, 0	0
201	Hambleton R, 1992, SO, 0, 0	0
186	Rasch G, 1986, SO, 0, 0	3

According to Table, mostly cited three authors are Lord (1981), Thissen (1984) and Samejima (1987) and they are in the same cluster. Embretson (1986), who received 314 citation is in cluster 2, whereas Bock (1982) and Reise (1993) who got 281 and 231 citations respectively are in cluster 4. Log Likelihood Ratio method was used for naming the clusters. The network structure, which shows the results more understandably, is shown in Figure 3.



Figure 3. Author Common Citation Network

The yellow circles in the network, indicate the centrality of the actors, whereas red circles indicate citation bursts. Centrality degrees, which measure the significance of mostly cited authors in the network and show the actors located at the center, are displayed in Table 5 (Borgatti, 2006).

Table 5. Centrality Degrees of Cited Authors and Cluster Numbers

Centrality	Reference	Cluster number #
0.16	Birnbaum A, 1982, SO, 0, 0	5
0.12	Bock RD, 1982, SO, 0, 0	4
0.11	Thissen D, 1984, SO, 0, 0	0
0.10	Reise SP, 1993, SO, 0, 0	4
0.10	Mcdonald RP, 1990, SO, 0, 0	1
0.09	Mislevy RJ, 1989, SO, 0, 0	4
0.08	Andrich D, 1986, SO, 0, 0	5
0.07	Holland PW, 1987, SO, 0, 0	2
0.07	Reckase MD, 1985, SO, 0, 0	0
0.06	**American Psychiatric Association, 1997, SO, 0	2

Regarding the centrality of the authors, it can be seen that Birnbaum (1982) is the author with the biggest impact in the network with 0.16. It was found that this author and Andrich (1986) are in cluster 5. The author with the second highest centrality is Bock (1982) with 0.12 and he is in the cluster 4 along with Mislevy (1989). Thissen (1984), who has the third biggest centrality with 0.11, was found to belong to initial cluster 0 along with Reckase (1985).

Citation burst were identified for 95 authors covered in WoS database, who worked on item response theory. Citation burst of top 10 authors and the years of the bursts are displayed in Table 6.

Table 6. Citation Burst Statistics of the Authors

Burst	Source	Cluster#	Start	End	1980-2018
37.71	Lord F, 1981,	0	1980	1997	
23.75	Cai L, 2012,	1	2014	2018	
18.56	de Ayala R, 2012	3	2012	2018	
16.41	Chalmers RP, 2016,	1	2016	2018	
15.36	Hulin CL, 1982,	0	1982	2001	
15.03	Wingersky MS, 1983,	0	1983	2000	
14.70	Reeve BB, 2010	2	2010	2015	
13.93	Edelen MO, 2009,	2	2014	2018	
12.59	Lord FM, 1981	0	1980	1992	
12.01	Rizopoulos D, 2010,	1	2010	2018	

According to Table 6, the author with the highest citation burst value is Lord, F. (1981), with 37.71. The work, whose effective period was between 1980-1997, belongs to initial cluster 0 and it had no citation burst after 1997. It can be seen that the author with the second highest citation burst value is Cai, L. (2012), with 23.75 and he received numerous citations between 2014-2018. Similarly, it can be seen that the author with the third highest citation burst value is De Ayala, R. (2012), with

18.56 and he got the citation burst in recent years. The review of Table 6 as a whole, reveals that Chalmers, R.P. (2016), Reeve, B.B. (2010), Edelen, M.O. (2009) and Rizopoulos (2010) got citation bursts after 2010, whereas Hulin, C.L. (1982), Wingersky, M.S. (1983) and Lord, F.M. (1981) have been popular from 1980's to 2000's, afterwards they lost their popularity as can be seen from the lack of citation.

Country Collaborations

Country collaborations were also analyzed within the scope of the study in order to reveal the countries that have been contributed to the academic works performed in item response theory. It was found that the network obtained as a result of the analysis consists of 29 nodes and 122 connections. The density of country collaboration network is 0.3005, its modularity value is 0.2212 and mean silhouette value 0.3359. The total number of citations received by the countries and the years of these citations are displayed in Table 7.

Table 7. Number of Citations for Country Collaborations and Centrality Values

Number of Citations	Centrality	Source Country	Year
687	0.69	US	1982
97	0.17	NETHERLANDS	1987
57	0.09	CANADA	1992
49	0.12	SPAIN	2001
43	0.03	PEOPLES R CHINA	2008
41	0.08	ENGLAND	2004
36	0.00	BRAZIL	2012
35	0.16	FRANCE	2004
30	0.00	JAPAN	2001
30	0.28	AUSTRALIA	2005
29	0.03	ITALY	2012
28	0.04	GERMANY	2003

Regarding Table 7, the country with the highest citation in item response theory area is US with 687 citations; the county with highest centrality is also US with a value of 0.69. Accordingly, it can be seen that US is well ahead as the leader country. It was found that the country with the second highest citation is Netherland with 97 citations, which also have the second highest centrality with 0.17. The third most cited country Canada has 57 citations and 0.09 centrality value, whereas the centrality value of Spain, which received 49 citations, is 0.12 and centrality value of Australia, which get 30 citations, was found to be 0.28. Moreover, centrality values of Japan and Brazil were computed as 0.00 and it was concluded that both of them are not active in IRT field. The network structure formed for country collaborations for easier interpretation are displayed in Figure 4.

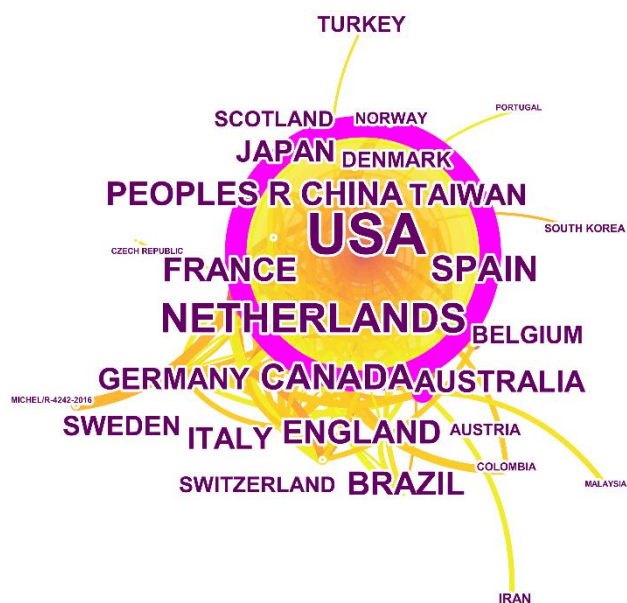


Figure 4. Country Collaborations

As can be seen from Figure 4'te, US makes the highest contribution to the field in terms of response theory, followed by Netherland, Canada, Spain and China. Citation burst values of 6 countries for which citation burst was detected and effective years of the bursts are shown in Table 7.

Table 7. Citation Burst Statistics according to Countries

Burst	Source	Start	End	1980-2018
5.99	ITALY	2014	2018	
5.97	PEOPLES R CHINA	2015	2018	
5.08	BRAZIL	2012	2018	
4.08	FRANCE	2014	2015	
3.88	TAIWAN	2004	2009	
3.73	DENMARK	2003	2006	

As can be seen from Table 6, the country with the highest citation bursts value was found to be Italy with 5.99. The work that covers 2014-2018 period received considerable citations in recent years. The country with the second highest citation bursts value was found to be China with 5.97, which is effective between 2015-2018. The country with the third highest citation bursts value is Brazil with 5.08, which was effective between 2014-2015.

Journal Common Citation Network

Journal common citation network was formed based on the journals in which the academic works covered in WoS database have been published. The network obtained at the end of the analysis include the journals with 100 or more citations. There are 312 nodes and 1,988 connections in the network. The density the network was found to be 0.041, Modularity value is 0. 4372 and Mean

silhouette value is 0. 2193. The network was divided into 29 clusters in terms of common characteristics of the journals through LLR estimation method. The network structure of the journals that published significant works made on IRT is displayed in Figure 5.

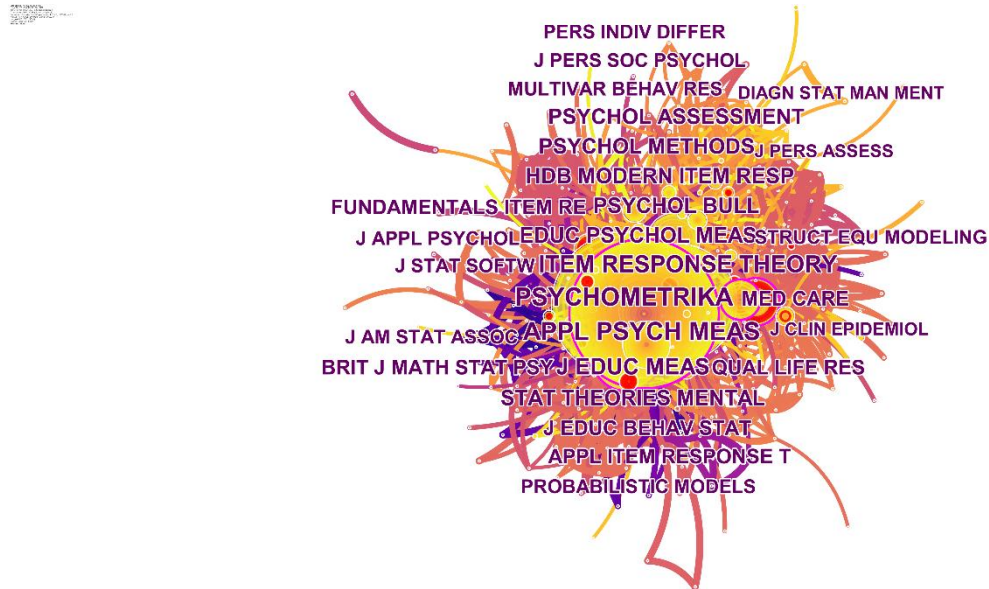


Figure 5. Journal Common Citation Network

Regarding Figure 5, it can be seen that the journals that come to forefront are Psychometrika, Appl PM, Item Response Theory and Educational Measurement. The citations that the journals have received and descriptive statistics of the clusters based on these citations are shown in Table 8.

Table 8. Mostly Cited Journals

Number of Citations	Centrality	Source	Cluster number#
671	0.09	Psychometrika, 1982, PSYCHOMETRIKA, 0, 0	0
669	0.12	Appl PM, 1982, APPL PSYCH MEASUREMENT, 0, 0	0
539	0.09	Item RT, 1984, ITEM RESPONSE THEORY, 0, 0	0
368	0.15	J EM, 1982, J EDUC MEAS, 0, 0	0
331	0.10	Educ PM, 1982, EDUC PSYCHOL MEAS, 0, 0	0
292	0.11	Psychol B, 1986, PSYCHOL BULL, 0, 0	2
260	0.03	Stat TM, 1989, STAT THEORIES MENTAL, 0, 0	0
258	0.04	Hdb MIR, 1997, HDB MODERN ITEM RESP, 0, 0	0
249	0.05	Psychol M, 1999, PSYCHOL METHODS, 0, 0	4
245	0.07	Psychol A, 1997, PSYCHOL ASSESSMENT, 0, 0	2
220	0.11	Med C, 2000, MED CARE, 0, 0	1
219	0.12	Qual LRES, 2000, QUAL LIFE RES, 0, 0	1
208	0.09	Brit JMSPSY, 1983, BRIT J MATH STAT PSY, 0, 0	0
162	0.10	Multivar BRES, 1999, MULTIVAR BEHAV RES, 0, 0	4

According to Table 8, it can be seen that the five journals with highest citations are respectively Psychometrika (N=671), Appl PM (N=669), Item Responce Theory (N=539), J EM (N=368) and Educational measurement (N=331). According to centrality, the journals that make most effective works in the field are J EM (m=0.15), Appl PM (m=0.12), Qual LRES (m=0.12), Psychol

Bulletin ($m=0.11$) and Med C ($m=0.11$). Top five journals with highest citations are in the same cluster whereas Med C and Qual LRES journals, which have very similar centrality values, are in cluster 1. In addition, the analysis results showing top 10 journals with the highest citation bursts among the 96 journals, which were identified as a result of the citation burst analysis, and the years of citations are displayed in Table 9.

Table 9. Citation Burst Values of the Journals Cited by the Works Covered in the Dataset

Burst	Source	Cluster	Start	End	1980-2018
46.11	Applications IRE, 1981. (Applications Item Response)	0	1980	1997	
28.95	Statistical T, 1981. (Statistical Theories)	0	1980	1997	
25.92	J SS, 2010. (Journal of Statistical Software)	0	2014	2018	
20.97	THESIS, 2015.	0	2015	2018	
19.53	Multilog UG, 1993. (Multilog Users Guide)		1993	2009	
17.00	R LENVSC, 2011. (R Language and Environment Statistical Company)	0	2013	2018	
16.53	Theory PI, 2010. (Theory Practice Item)	0	2012	2018	
14.35	Struct EQU, 2004. (Structural Equaling Modeling).	4	2015	2018	
12.54	J ES, 1982. (Journal of Educational Statistics)	0	1982	1996	
12.44	PSYCHOMETRIKA, 1982.	0	1982	1996	

Regarding Table 9, it can be seen that the biggest citation burst is performed by Applications IRE with 46.11 citation burst value. This journal, which was quite popular between 1980-1997, belongs to cluster zero (#0). The journal with the second highest citation burst is Statistical Theories, with 28.95 citation burst value. This journal has been quite popular between 1980-1997 and it is in the same cluster with Applications IRE. It was found that J Stat Soft, which has 25.92 citation burst value has been popular between 2014-2018 and it belongs to the same cluster with the top 2 journals. Similarly, the journal called Thesis was popular in 2015 and afterwards with 20.97 citation burst value, and it is in the same cluster with the top three journals having the highest citation burst.

Word analysis

A word analysis was performed to identify the words that were frequently used in the works concerning item response theory without dividing them into subcategories such as author, journal, and country. As a result of the analysis, which has been performed based on the frequency of the words without putting any restriction, the words repeated 20 or more times were identified. The network that has been formed accordingly has 378 nodes and 1,999 connections. The density of the network was found to be 0.0281, Modularity value is 0.456 and Mean silhouette value is 0.4992. The network was

divided into a total of 13 clusters. The citation values of top 10 works and their clusters are displayed in Table 1. The word cloud of mostly repeated words is shown in Figure 6.

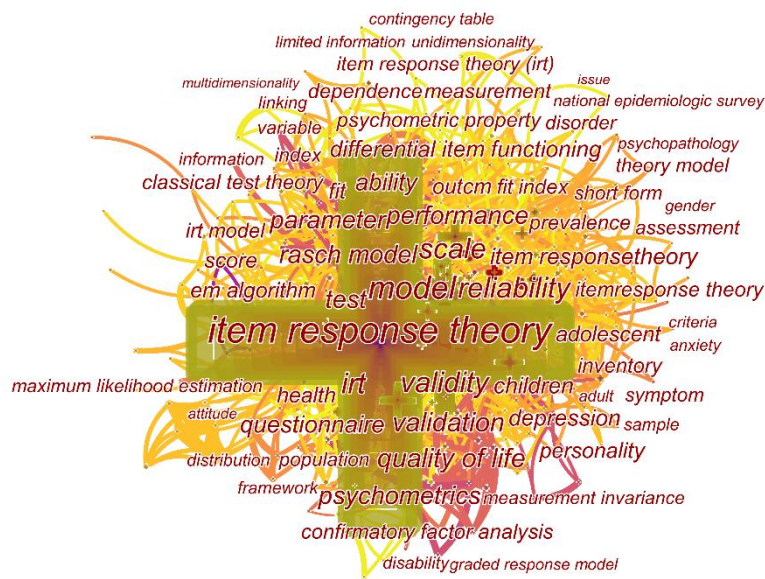


Figure 6. Word Cloud of Mostly Repeated Words

The number of repetitions of each word, the years of repetition and centrality values of the words are displayed in Table 10, allowing statistical interpretation of visual results.











Table 10. Statistics about Mostly Repeated Words

Number of Citations	Centrality	Word	Year
436	0.08	item response theory	1994
192	0.05	model	1993
107	0.06	validity	1999
95	0.09	reliability	2000
93	0.06	IRT	2003
91	0.03	scale	2003
67	0.03	validation	2004
65	0.04	quality of life	2003
59	0.08	Rasch model	2001
59	0.08	test	2000
56	0.06	performance	1996
52	0.09	parameter	1999
52	0.04	psychometrics	2005
47	0.08	questionnaire	2000
43	0.03	ability	2000
39	0.02	children	1999
37	0.03	item response theory	2001
35	0.10	differential item functioning	2000
34	0.03	depression	2007
33	0.05	EM algorithm	2001
31	0.07	fit	1999
27	0.07	confirmatory factor analysis	2008
23	0.07	inventory	1999
23	0.06	measurement	2006

23	0.05	classical test theory	2005
18	0.06	maximum likelihood estimation	1998
13	0.06	distribution	2005

Regarding Table 10, it can be seen that the words mostly repeated in the academic works concerning item response theory are respectively; item response theory, model, validity, reliability, IRT, scale and validation. Accordingly, the review of most frequently used words revealed that item response theory is used to determine the validity and reliability of the models built for measurement purposes, as well as their relationship with classical test theory. It can be seen from subsequent words that IRT is used for the validation of the scales and tests and Rasch model was considerably repeated in this stage. Since usage frequencies of the words “scale”, “questionnaire” and “psychometric” are very close to each other, it can be concluded that IRT-based methods are employed for determining psychometric characteristics of measurement tools. Similarly, the frequency values of “ability”, “children” and “item response theory” words are following each other, which was taken as an indicator that item response theory is used in the skill estimations of the students. A total of 16 citation bursts were identified, top 10 words with highest citation bursts and their years of popularity are displayed in Table 11.

Table 11. Citation Burst Values of the Words Cited by the Works Covered in the Dataset

Word	Year	Burst	Start	End	1980-2018
model	1980	7.7895	1980	2002	
IRT	1980	6.2823	2003	2007	
reliability	1980	6.0093	2000	2008	
test	1980	5.9689	2000	2009	
validity	1980	5.05	2007	2008	
ability	1980	4.6252	2003	2007	
validation	1980	4.5118	2016	2018	
Rasch model	1980	4.4726	2001	2004	
outcome	1980	4.1689	2014	2015	
national epidemiologic survey	1980	4.1437	2009	2011	

Regarding Table 11, it can be seen that the word with the highest citation burst value is model and it was the most popular word in item response theory between 1980-2002. The word with the second highest citation burst was found to be IRT, which was considerably repeated in WoS database between 2003-2007. The word reliability, which has been emerged first in 1980's, has the third highest citation burst value and it was popular between 2000-2009. The word test that has the fourth highest citation burst value has been emerged in 1980, but it became very popular between 2000-2009. Similarly, the word validity that has been emerged in 1980's, made a citation burst between 2007-2008. The word ability was very popular between 2003-2007. As can be seen from the table, the popularity of most frequently used words in the subject of item response theory lasted very short and the popular concepts in the field varied according to years.

RESULTS, DISCUSSION AND SUGGESTIONS

This study was performed to determine the outstanding authors, journals, countries and subject areas of the works performed in Item Response Theory area through bibliometric analysis of the

academic works published in WoS database. 1,367 academic works published between 1980-2018 have been analyzed through Citespace II software and the outcomes were reported both graphically and statistically.

As a result of the findings obtained in the study, the authors who get the highest number of citations in the field are De Ayala, Embretson, Reckase, Reise and Chalmers, whereas the authors that created biggest citation bursts in the field are respectively Embretson, Lord and Van der L. A common citation analysis was performed based on the mostly cited authors and a total of 7 clusters were obtained, named as variable score, clinical assessment, functional status item, inventory, teacher certification, clinical studies, and national sample. Regarding the countries that made contribution to the field, it was found that US, Netherland, Canada, Spain and China are the countries that performed the highest number of works, respectively. Regarding citation burst values of the countries, the biggest bursts were created by Italy, China and Brazil. The top five organizations that made the highest number of works in terms of journal were found to be Psychometrika, Appl Psych Measurement, Item Response Theory, J Edu Measurement and Educ Psychol Measurement. Regarding citation burst values of the journals, Application Item Response, Statistical Theories and J Statistical Software are at the top three. Moreover, as a result of the word analysis conducted on the works made in the field, it was found that mostly repeated words are item response theory, classical test theory, model, validating, reliability, validity and Rasch model.

In a similar study conducted by Glanzel (2012), the researcher has determined 4 subject areas and has compared the number of clusters emerged in two time periods, 1999-2003 and 2004-2008. The number of clusters were found to be different in two time periods; country collaborations were also analyzed in these 4 different subject areas to reveal the countries that have made the highest number of collaborations in each field. As a result of the study it was found that US is the country with the highest contribution in all four subject areas.

In another similar study conducted by Liu and Shen (2013), the change of academic works concerning idioms according to time, countries and universities have been examined through bibliometric analysis method. Citespace was used in the study where mostly cited authors and the clusters based on them have been determined. As a result of the study, it was found that the academic works performed since 1960 until present were divided into a total of 7 clusters.

In the study conducted by Martinez, Cobo, Herrera and Herrera-Viedma (2015), the works published in 25 journals between 1930-2012 have been scanned in WoS and Journal Citation Reports databases and they have been analyzed through bibliometric analysis method. As a result of the study, in which Science Mapping software has been used, a total of 8 clusters were identified, including children (mostly worked topic), social services, health services, violence, women, HIV/AIDS, social service specialists and education. Moreover, the analyses have been reperformed in order to see the variation of the clusters in three different time periods, namely 1930-1989, 1990-2002 and 2003-2012. As a result, it was found that mostly cited subject fields have been varied in different time periods.

Yalçın and Yayla (2016) have analyzed 543 academic works on Technological Pedagogical Content Knowledge, published between 2008-2015 in WoS and Scopus through bibliometric analysis method using Citespace II software. Most cited authors, clusters obtained based on these authors, most cited articles and clusters obtained based these articles have been determined. As a result of the study it was found that the studies performed in this area have been increased from past to present. In addition, the authors, journals and countries that realized the highest number of works have been reported through burst analysis according to time.

It can be seen that studies involving bibliometric analysis method became popular in recent years. In a similar study performed by Zhang, Huang, Quing, Li and Tian (2017) the academic works about remote sensing, published in WoS, Scopus and Google Scholar databases between 2010-2015 have been examined through bibliometric analysis method. In this study, they have compared the works related to remote sensing with the works performed in other areas, and they have identified the

countries and institutes that have mostly contributed to the field. In addition, they have attempted to determine popular topics of the field through word query.

Güzeller and Çeliker (2017) have analyzed 703 academic works in gastronomy area between 1970-2017 using Citespace II software. As a result of the analysis of the works in WoS database in gastronomy subject field, it was found that US plays a key role in country collaborations, the journal and author with the highest citation burst is Journal of Culinary Science & Technology, and Herve This. In addition, it was found that the works that directed the field have been realized between 2003-2005.

The data set used in this study has been formed based on 1,367 academic works indexed in WoS database, between 1980-2018. It is believed that the most extensive data set has been used compared to the works in which similar methods had been used, aiming to reveal the general status of the field. It is thought that this study will set an example for future studies in terms of the performance of the analysis. In addition, it can be seen as a pioneer in realizing the works based on bibliometric analysis method, in education and social science area, using different databases. In addition, with the network structure obtained for common work and common author and country collaborations, the connections between the authors, works and countries that are named as node have been visually presented. In this way, the big picture of the area has been revealed. The collaboration structure belonging to outstanding works and authors of item response theory should be considered as a guide that will form the start for future researches.

REFERENCES

- Abdi, A., Idris, N., Alguliyev, R. M. & Aliguliyev, R. M. (2018). Bibliometric Analysis of IP&M Journal (1980–2015), *Journal of Scientometric Research*, 7 (1), 54-62.
- Barabási, A. L. (2002) *Linked: How Everything Is Connected to Everything Else*. New York: Basic Books.
- Barnes, I.A (1954). Class and committees in a Norwegian island parish. *Human Relations*, 7, 39-58.
- Barnes, J. A. & Harary, F. (1983). Graph theory in network analysis. *Social Network*, 5, 235-244.
- Borgatti, S.B., Everett, M.G. & Johnson, J.C. (2013). *Analyzing Social Networks*. UK: Sage Publications.
- Borgatti, S., & Halgin, D. S. (2011). On network theory. *Organization Science*, 22, 1168–1181.
- Borgman, C. L. (1999). Books, bytes, and behavior: Rethinking scholarly communication for a global information infrastructure. *Information Services & Use*, 19, 117-121.
- Butts, C. T. (2008). Social network analysis: A methodological introduction, *Asian Journal of Social Psychology*, 11, 13-41.
- Cartwright, D. (ed.) (1959). *Studies in Social Power*, Ann Arbor, MI: Institute for Social Research.
- Chen C. (2005). *CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature*. JASIST.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory* (First edition). California: Thomson Learning.
- Crossley, N., Prell, C. & Scott, J. (2009). Social Network Analysis: Introduction to Special Edition, *Methodological Innovations Online*, 4, 1-7.

- Dorogovtsev, S. N. & Mendes, J. F. F. (2002). Evolution of networks, *Advances Phys*51 (4), 1079–1187.
- Duijin, M.A.J. & Vermunt, J. K. (2006). What Is Special About Social Network Analysis?, *Methodology* 2006, 2 (1), 2-6.
- Freeman, L. C. (2004) *The Development of Social Network Analysis: A Study in the Sociology of Science*. Vancouver: Empirical Press.
- Glanzel, W. (2012). Bibliometric methods for detecting and analysing emerging research topics, *El profesional de la información*, 21 (2), 194-201.
- Güzeller, C. O. & Çeliker, N. (2017). Geçmişten Günümüze Gastronomi Bilimi: Bibliyometrik Bir Analiz, *Journal of Tourism and Gastronomy Studies*, 5 (2), 88-102.
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Hogan, B., Carrasco, J. A. & Wellman, B. (2007). Visualizing Personal Networks: Working with Participant-Aided Sociograms, *Field Methods* 19 (2), 116-144.
- Jain S, Basavaraj P, Singla A, Singh K, Kundu H, Vashishtha V. (2015). Bibliometric analysis of Journal of Clinical and Diagnostic Research (Dentistry Section; 2007-2014). *Journal of Clinical and Diagnostic Research*. 9 (4), 47-51.
- Jamali, M. & Abolhassani, H. (2006). *Different Aspects of Social Network Analysis*, in: IEEE/WIC/ACM International Conference on Web Intelligence, 66–72.
- Liao, H., Tang, M., Luo, L., Li, C., Chiclana, F. ve Zeng, X.J. (2018), A Bibliometric Analysis and Visualization of Medical Big Data Research, *Sustainability*, 10 (166), 1-18.
- Liu, D.D., Liu, S.L. & Zhang, J. H. (2014). Visualization analysis of research hotspots based on CiteSpace II: taking medical devices as an example, *Medical Devices: Evidence and Research*, 7, 357-361.
- Liu, W. & Shen, H. (2013). CiteSpace II: Idiom Studies Development Trends, *Journal of Arts and Humanities (JAH)*, 2 (2), 85-97.
- Martinez, M.A., Cobo, M.J., Herrera, M. & Herrera-Viedma, E. (2015). Analyzing the Scientific evolution of Social Work Using Science Mapping, *Research on Social Work Practice*, 25(2), 257-277.
- Martino, F. & Spoto, A. (2006). Social Network Analysis: A brief theoretical review and further perspectives in the study of Information Technology, *PsychNology Journal*, 4 (1), 53-86.
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27, 415–444.
- Moreno, J.L. (1946). Sociogram and sociomatrix: A note to the paper by Forsyth and Katz. *Sociometry*, 9, 348-349.
- Mincer, M. & Niewiadomska-Szynkiewicz, E. (2012). Application of Social Network Analysis to the Investigation of Interpersonal Connections, *Journal of Telecommunications and information technology*, 2, 83-91.
- OECD, (2017). *PISA 2015 Technical Report: Scaling PISA data*, OECD Publishing, Paris.

- Pritchard, A. (1969). Statistical Bibliography or Bibliometrics. *Journal of Documentation*, 25 (4), 348–349.
- Reise, S. P., Ainsworth, A. T. & Haviland, M. G. (2005). Item Response theory. Fundamentals, applications, and promise in psychological research. *Current Directions in Psychological Science*, 14 (2), 95-101.
- Sünbül, Ö. ve Erkuş, A. (2013). Madde Parametrelerinin Değişmezliğinin Çeşitli Boyutluluk Özelliği Gösteren Yapılarda Madde Tepki Kuramına Göre İncelenmesi, *Mersin Üniversitesi Eğitim Fakültesi Dergisi*, 9 (2), 378-398.
- Synnestvedt M, Chen C, Holmes J. (2005). CiteSpace II: visualization and knowledge discovery in bibliographic databases, *AMIA AnnuSymp Proc 2005*, 724-728.
- Tsay, M. Y. (2011). A Bibliometric Analysis on the Journal of Information Science, *Journal of Library and Information Science Research*, 5 (2), 1-28.
- Van Leeuwen T. N. (2004). *Second generation bibliometric indicators: the improvement of existing and development of new bibliometric indicators for research and journal performance assessment procedures [dissertation]*. Leiden.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. New York: Cambridge University Press.
- Wasserman, S. & Faust, K. (2009) *Social Network Analysis: Methods and Applications*, USA: Cambridge University Press.
- Wasserman, S., & Robbins, G. (2005). An introduction to random graphs, dependence graphs, and p^* . In P. J. Carrington, J. Scott, & S. Wasserman, (eds.). *Models and Methods in Social Network Analysis*. Cambridge: Cambridge University Press.
- Watts, D. J. (2004) *Six Degrees*. New York: W. W. Norton.
- Wimmer, A. & Min, B. (2006). From empire to nation-state: Explaining wars in the modern world, 1816–2001. *American Sociological Review*, 71 (6), 867–897.
- Wölfer, R., Faber, N. S. & Hewstone, M. (2015). Social Network Analysis in the Science of Groups: Cross-Sectional and Longitudinal Applications for Studying Intra and Intergroup Behavior, *Theory, Research, and Practice*, 19 (1), 45–61.
- Wu, Y. T. (2013). Research trends in technological pedagogical content knowledge (TPACK) research: A review of empirical studies published in selected journals from 2002 to 2011. *British Journal of Educational Technology*, 44 (3), 73-76.
- Yalçın, H. & Yayla, K. (2016). Teknolojik Pedagojik Alan Bilgisi Konusunda Yapılan Araştırmaların Bibliyometrik Analizi ve Bilimsel İletişim, *Eğitim ve Bilim*, 41 (188), 291-307.
- Zhang, H., Huang, M., Quing, X., Li, G. & Tian, C. (2017) Bibliometric Analysis of Global Remote Sensing Research during 2010–2015, *International Journal of Geo information*, 6 (332), 1-19.