Article History
Received: 01.01.2021
Received in revised form: 12.07.2021
Accepted: 18.08.2021
Article Type: Research Article

International Journal of Contemporary Educational Research (IJCER)

www.ijcer.net

How Central is the PISA Outcomes on Human Development?

Akif Avcu[1]
[1]Marmara University, 0000-0003-1977-7592

To cite this article:

Avcu, A. (2021). How Central is the PISA Outcomes on Human Development?. *International Journal of Contemporary Educational Research, 8*(4), 16-26. https://doi.org/10.33200/ijcer.851852

International Journal of Contemporary Educational Research
Volume 8, Number 4, December 2021, Page 16-26 ISSN: 2148-3868

How Central is the PISA Outcomes on Human Development?

Akif Avcu[1*]
[1]Marmara University

Abstract

In contrast to traditional statistical approaches, which assume the existence of a latent common cause leading to the emergence and covariance of indicators, network modeling assumes that latent features emerge because of interactions between indicators. Clearly, such a way of treating the results of the Program for International Student Assessment and other development indicators better reflects the mutual interactions among indicators. With this aim, the network pattern of development indicators was uncovered and graphically represented, the most important and least important indicators were identified. In addition, the indicators that are more closely related to the results of the Program for International Student Assessment were also identified. The United Nations Development Program data were used for the analyzes. Data from 2015 for sixty-six countries were used, consisting of thirteen development indicators. The data were analyzed in the statistical program R using the package "qgraph". The results showed that Program for International Student Assessment scores were not central to other development indicators, while they were closely associated with gender inequality, secondary school completion rate, and unequal life expectancies. These findings were discussed based on the existing literature and some recommendations were made for policy makers and for future research.

Keywords: PISA, network modeling, development indicators

Introduction

PISA (Program for International Student Assessment) was developed in 1997 by the member countries of OECD (Organization for Economic Cooperation and Development) to evaluate students' performance in reading, mathematics, and science (OECD, 2001). The PISA mastery scores are expected to assess how much 15-year-old students have the knowledge and skills necessary for their participation in the labor market and society (OECD, 2006). The main aim of PISA is directly related to the measurement of the knowledge and skills of students, to associate these with the data collected from students, teachers, schools and educational systems to understand the differences between performances and then improve those educational systems (OECD, 2019). In other words, the worldwide data obtained by PISA are used to determine the factors associated with student achievement and to establish standards for increasing the quality of education systems (OECD, 2017).

Although it was originally developed for the OECD countries, in the following years, PISA has been turned into a global standard and is currently used in over 70 countries and economies. Currently, no other educational work has received such a substantive media coverage (Grek, 2009) and such a high interest from general population (Pongratz, 2006). Importantly, since the first results were published, reforms of all school systems have been shaped by PISA data to some extend (see Takayama, 2009). This general interest is understandable because a well-educated population is crucial to a country's economic and social development. After world war II, they are improving educational quality and expanding mass education have become a more important component for development at the global level (Chabbott & Ramirez, 2000, p. 163). More and better education has come to be seen as a prelude to rapid economic development world widely. According to this perspicacity, education promotes economic growth and foster development in people's lives through many channels (Barro, 1997). Therefore, societies have an irrepressible interest in ensuring that children and adults have access to better educational opportunities (OECD, 2010). In line with this reality, the PISA results is currently used by the United Nations (UN) system as an important indicator to track progress towards the sustainable development goals established by the international community as a plan to achieve a better and more sustainable development for all (OECD, 2018, p. 2). Toward achieving a sustainable development, the United Nations track the development of

* Corresponding Author: *Akif Avcu, avcuakif@gmail.com*

countries with some indicators. Those indicators were grouped into six sub-titles (demography, education, environmental sustainability, gender, human security, income/composition of resources, inequality, mobility and communication, poverty, socio-economic sustainability, trade and financial flows, work/employment, and vulnerability). the PISA results are also seen as a development indicator under the subtitle of education.

Although many empirical studies show a positive relationship between development and education, the effects of education on development are uncertain (Chabbott & Ramirez, 2000, p. 163). A positive effect of education on economic, political and cultural development is widely assumed in most of the modern and modernizing world. Accordingly, many studies consider education as a lever for development (Bloom, Canning, Chan, & Luca, 2014; Clark & Royer; 2013; Cinnirella & Streb, 2017), but contrary to this general assumption, it is clear that this relationship is reciprocal rather than causal (Afşar, 2009), as developments in other areas and spending on education also promote educational outcomes (Baldacci, Guin-Sui, & de Mello, 2004) and the relationship between development indicators is dynamic rather than causal. For this reason, instead of modeling the relationship between education and developmental indicators with causal models, a different approach is needed to realize better how the PISA results is associated with other development indicators.

In the last few decades, the common latent cause model has been one of the most widely used approaches to identify latent traits in social sciences (i.e. van Borkulo et al., 2015; Beard et al., 2016). Based on this approach, it is possible to make inferences about the existence of a hypothetical causes and effects. This approach helps in examining causal relationships between latent traits (Kelava & Brandt, 2014). It also made important contributions to our understanding of the educational and psychological phenomena under study. In this framework, variables are generally considered passive indicators of some common causes. When indicators are considered passive, it means that they are grouped together based on a common cause (Fried, 2015). Consistent with this understanding, indicators have generally been considered passive indicators of a latent trait (development) and have been modeled in several studies using a traditional factor analytic approach (e.g., Ganegodage, Rambaldi, Rao, & Tang, 2006; Qiu, Sung, Davis, & Tchernis, 2018).

On the other hand, it may be better to regard indicators as dynamic systems. One branch of this new understanding consider observable variables as causal systems or networks (Cramer, Waldorp, van der Maas, & Borsboom, 2010). The network modeling has already been proposed as an alternative method to solve psychometric problems (Borsboom & Cramer, 2013) because, handling of data in this manner enable to see the dynamic nature of the relationships among indicators and to detect the most, the least important and the abundant indicators in a network.

In contrast to the traditional psychometric approach, which assumes the existence of a latent trait that causes the emergence and covariance of indicators, network psychometrics assumes that indicators exist because of interactions between them (Borsboom, 2017). In this context, variables are considered to be properties that interact directly with each other and do not arise because of a common latent characteristic (i.e., development). Similarly, it is possible to consider development indicators as network units that interact with each other. The existence of these units depends on their interaction with each other and is reinforced by the existence of each other.
It was stated that network analysis will transform our understanding of social constructs to a certain extent (Mcnally, 2016). This analysis, which used a quite new approach, has started to take place in the psychology literature in the last 10 years. There is even more opportunity in the field of education to fully understand the network structure of related concepts. With the proliferation of network analysis, a number of important findings such as understanding practically more important variables, better understanding of underlying mechanisms, and examining protective factors could be achieved. In addition, this new model will help to better understand social phenomena because it provides different information than the ones provided by the common latent model. This new technique can also be used with longitudinal (Snijders, 2009) or multi-group (Kim & Leskovic, 2012) data. This flexibility will also facilitate the widespread use of this approach.

Regarding the novelty of network modeling in educational research and the potential of network modeling to provide new insights on phenomena under investigation, this study aims to understand the network structure of development indicators of United Nations and see how central the PISA results in this network. By this aim, network pattern of development indicators was revealed and graphically represented, the most and the least central indicators were detected.

Method

Participants and Data

The data of the current study were obtained at the country level. Data consisted of development indicators of countries released by UNDP (United Nations Development Program). The data are available on the program's website (http://hdr.undp.org/en/data) and consist of 150 global human development indicators for over 190 different countries. Those indicators cover different aspects of human development under different dimensions: income, inequality, education, health etc. The key dimensions (longevity, education and income) of human development were measured with combined score, named as human development index (HDI). In order to prevent overlapping of indicators, HDI was not included in the current study. The educational quality of countries is investigated with different indicators including the PISA results. Instead of using science and mathematics reading literacy scores separately for the current study, the combined PISA score was calculated and used. This score was simply obtained by calculating arithmetic average of reading, science and mathematics literacy scores. Because these scores are highly correlated (the correlation between science and reading scores is .96; the correlation between science and math scores is .98; the correlation between math and reading scores is .94) and they are standardized scores with a mean of 500 and a standard deviation of 100, the unweighted arithmetic mean was used as an indicator of success from PISA.

For 2015 assessment, there were only 72 countries with available PISA scores. On the other hand, the data was available for 67 countries in the UNDP database. In addition, Qatar was removed from the database because the data was missing for some key development indicators. Hence, development indicators of 66 countries were included in the final dataset. These variables were selected based on two criteria: (1) availability of data for the 66 countries included; (2) the indicators selected must come from different development dimensions (health, education, gender inequality, power, labour, employment and vulnerability, socio-economic sustainability, demography and environmental sustainability) and be as diverse as possible. On the other hand, as with other model prediction techniques, as the number of nodes to be estimated increases, more parameters need to be estimated, which requires a larger sample size (Epskamp, Borsboom, & Fried, 2018). Given the number of countries included in the dataset, only the representative indicators for each dimension and with complete data were included. For the current study, thirteen different development indicator was selected to estimate the network structure. They are as follows: combined PISA score (PIS); gross domestic product (GDP) per capita; unemployment percentage of the total labor force (Unm); net migration rate (Mgr); percentage of total population living in the urban area (r)); percentage of the total population access to the internet(Int); mortality rate attributed to unsafe water, sanitation and hygiene services (Mrt); life expectancy at birth (Lfx); percentage of GDP for current health expenditure (Hlt); inequality in life expectancy (Inq); gender inequality index (GII); gender development index (GDI); the percentage of the population with at least some secondary education (ScE).

Network Analysis

Networks are abstract models consisting of a set of nodes, a set of edges connecting nodes, and information about the nature of the nodes and the edges (De Nooy, Mrvar, & Batagelj, 2011). In the psychometric network approach, all indicator variables are represented as nodes and the association of these variables as edges in a graphical network model. For example, how often a person had sleep problems in the past week is represented as a node in the network model. For the current study, each developmental indicator was considered a node in the network. These nodes are connected by edges and show the relationship between the nodes. The main advantage of the network approach is that it also allows us to measure the overall importance of the indicators.

The edges in the networks could be weighted or unweighted. In weighted networks, a value or a coefficient is indicating the magnitude of connection representing the association of nodes while in unweighted networks, nodes are connected with edges which don't specify such a magnitude. For weighted edges, the magnitude of association is graphically represented by the thickness of the edge and numerically can range from - / + 1. The closer the value is to +1 or -1, the greater the edge strength and the stronger the relationship between the nodes. In the graphical representation, a negative relationship is usually represented by a red line and a positive relationship by a green line. The value zero means that there is no edge connecting the nodes.

Furthermore, the edges of the networks can be directed or undirected. Undirected networks consist of edges or simple lines connecting the pair of nodes where the direction of this relationship is not specified. In this case, arrows are not present at the end points of the colored lines. On the other hand, in directed networks, there are arrows on one or both end of the edge specifying the direction of predictions and causal relationships.

The analysis is based on network psychometrics and consists of two main steps: (1) the statistical model is estimated; (2) using graph theory, the undirected and weighted network structure is plotted. (Newman, 2010). In more concrete terms, in the first step, the edge weights were estimated using one of three different correlation coefficients: (a) simple correlation; (b) partial correlations; and (c) regularized partial correlations. Perhaps the simplest way to calculate a network of a psychological construct is to draw an edge between any two related nodes and represent the association with Pearson correlation. In this way, an undirected, weighted and signed (sign indicates the direction of the relationship) network could obtained (Cramer et al., 2012). Networks based on simple correlations may be useful for visualizing a complex relationship patterns. Still, the main limitation of such networks is that it is not possible to see whether the observed relationship is due to the real interaction between the two variables or the confusing effect of other variables in the network.

Due to this limitation in using the correlation coefficients, using partial correlation coefficients to estimate a network became common in constructing networks in data assumed to have multivariate normality (McNally et al., 2015). Such networks are also known as Gaussian Graphic Models (GGM; Lauritzen, 1996). On the other hand, like simple correlations, use of partial correlations carry an inherent limitation. Due to the sampling variability, when the partial correlation network is estimated, zero value can almost never be estimated between the two nodes. Even though the two variables are conditionally independent, estimations most likely yield relatively small partial correlation values and these small values will be represented as weak edges in the model. These negligible links are called spurious (Costantini, et al., 2015). Karl Pearson (1897) was the first person who pointed to the concept of spurious correlation and explained it as a significant correlation between the two variables that don't actually exist. It occurs due to a third variable omitted during the data collection process (also termed as confounding factors). In addition, sampling error and biased estimated can also inflate the correlation values (Abelson, 2012). The reason why they are called spurious is that they represent relationships that do not exist. Controlling these pseudo-connections is desirable because their existence in partial correlation networks cause biased positive relationships. In addition, controlling also reduce the number of parameters estimated and the risk of overfitting.

For this reason, partial correlation networks are generally estimated using regularization techniques commonly used in machine learning. In this way, possibly fake edges are removed in the model and sparser connections can be obtained. In this way, the interpretability of the network prediction is increased. Regularized estimation is achieved using LASSO (the Least Absolute Shrinkage and Selection Operator) or Graphical-LASSO (GLASSO), a variant of LASSO (Epskamp et al., 2018).

In practice, LASSO shrinks the partial correlation coefficients when predicting the network model. LASSO uses a penalization approach to compute the values by adding absolute parameter values. In more concrete terms, small coefficients are estimated as zero. In this way, fewer connections retained in an estimated network which is called a sparse network. LASSO controls the degree of regularization using an adjustment parameter. This regularization parameter is selected for minimizing the Extended Bayesian Information Criterion (EBIC; Chen & Chen 2008). The adjustment parameter can be set between 0 to 1. If it is set lower, fewer connections are removed, resulting in less sparse connections in the network model. On the other hand, if the parameter is set high, too many connections are eliminated and cause some of the real connections to be eliminated and spurious connections. A well accepted recommendation is to set tuning parameter as .5 (Foygel & Drton, 2015). The penalization approach has been reported to perform well in predicting the actual network structure (Foygel & Drton, 2010; Friedman, Hastie, & Tibshirani, 2008; Meinshausen, & Bühlmann, 2006).

The main advantage of the network analysis approach is that they also allow us to specify important nodes/symptoms/indicators. Centrality indices determine the importance of the nodes in a network. They allow us to evaluate the relative importance of nodes based on where the node is located in the network and on the pattern of connections. Generally speaking, a node is central if it has many connections while it is regarded as peripheral if it has a few connections with other nodes.

In simple terms, the centrality of a node can be defined by how strong its connections to neighboring nodes are. This definition corresponds to the strength of the node (Barrat, Barthélemy, Pastor-Satorras, & Vespignani., 2004). The strength centrality indices refer to the extent of a node's connections with other nodes. If the strength parameter of a node is high, it means that it influences many other nodes. The strength of a node is obtained by simply adding all the edges connected to a node. If the network is formed with partial correlations, the node strength is equal to the sum of the partial correlation coefficients between the node and all other nodes.

Moreover, the centrality of a node can be defined by the shortness of the direct or indirect edges connecting it to other nodes. This type of operational definition of centrality corresponds to closeness (Freeman, 1978). The

Closeness Centrality Index is defined as the inverse of the sum of the distances between a node and all other nodes in the network. If the Closeness Centrality Index of a node is high, it is a node that can predict other nodes well. Node strength indicates how strongly a node is connected to other nodes in the network, while closeness indicates how strongly a node is indirectly connected to other nodes in the network.

Finally, centrality is also defined according to how important it is to the interactions of other nodes with each other. Defining centrality in this way is known as betweenness (Brandes, 2001). The value of betweenness index is determined by calculating how often a particular node is the shortest path between two other nodes. It quantitatively indicates how many of the shortest paths between the two nodes lead to that node. The higher the Betweenness value, the more important the node becomes for connecting other nodes.Further estimations were made to calculate the clustering coefficients of each indicator. Formally, the clustering coefficient indicates to what degree nodes in an estimated network cluster together. This coefficient provide evidence for the abundance of given a node in a trio of nodes. In more straightforward terms, if two nodes do notneed a third node to be connected, this third node is called abundant because it doesn't give any function to connect different nodes. Thereby, the clustering coefficient value of this node become higher (Masuda, Sakaki, Ezaki, & Watanabe, 2018).

Statistical Analysis

Before starting analysis part, it was investigated whether or not the assumption of multivariate normal distribution was met. To this aim, the Mahalanobis distances were calculated and assessed at p<0.001 significance level. As Epskamp, Maris, Waldorp and Borsboom (2016) pointed out, network modelling was the only assumption. After checking the assumption, the analyzes of the study were carried out in two steps. In the first step, the network structure of 21 BDI items was estimated by using the "*qgraph*" package (Epskamp et al., 2012) available in R program (R Core Team, 2019). The estimated model was undirected and weighted. Even the unweighted networks only concern whether or not two nodes are connected, the weighted networks also take the strengths of the connection between each node pair into account. The strength of the connection is determined with the centrality indices and clustering coefficient. These concepts will be introduced at the end of this part. GLASSO regularization was applied to obtain a sparse network where spurious connections were estimated as exactly zero. In this way of analyzing the data, false positive relationships can be reduced. In the GLASSO approach, the edges correspond to the partial correlation coefficients between the nodes. In this way the relationship between two symptoms can be obtained after controlling for the effects of other indicators. A shrinkage parameter was used to minimize the EBIC. Doing so aimed to increase the accuracy of the estimated network structure (van Borkulo et al., 2014). In addition, using the Fruchterman-Reingold algorithm (Fruchterman & Reingold, 1991), stronger nodes were placed at the center of the network. In the second step, centrality indices (strength, closeness, betweenness) and clustering coefficients were estimated. There are four different clustering coefficients available in qgraph package: clustering coefficient for unweighted networks (Watts & Strogatz, 1998), Zhang's clustering coefficient (Zhang & Horvath, 2005), Onnela's clustering coefficient (Onnela, Saramaki, Kertesz, & Kaski, 2005), and Barrat's weighted clustering Coefficient (Barrat et al., 2004). Zhang's clustering coefficient was selected to report and discuss because of the similarity of findings of other coefficients. Finally, Small Worldedness Index was calculated to see whether or not development indicators have a small world topography.

Results

The assumption of multivariate normality was investigated by calculating the Mahalanobis distances. The results suggested that there was no violation to this assumption at $p<0.001$ significance level. As mentioned in previous session, the network model of development indicators was estimated as weighted and undirected using the qgraph package in R environment. 13 different development indicators, selected from the UNDP database, were represented by one node in this network. The GLASSO algorithm was used for estimation. For this reason, the edges connecting nodes represent partial correlation coefficients between symptoms. These partial correlation coefficients representing the relationship between the two nodes can be regarded as the edge-weight. When performing GLASSO regularization, a tuning parameter that controls the model's sparsity by minimizing EBIC was utilized. In this way, the Type I error was controlled by narrowing down all the parameters with near-zero values. In addition, using the Fruchterman-Reingold algorithm, more central parameters were placed in the middle of the graph.

The Pearson correlations and regularized partial correlations obtained from the estimated network model were given in Table 1. The results showed that the combined PISA score is highly correlated with "GII", "HDI" and "Ineqlife" indexes at $p<0.01$ *level*. On the others hand, it has no significant relationships with "Mobile", "Migration" and "Urb%" indexes. Regarding the regularized partial correlations, some of the values were estimated as zero because they were spurious. After controlling the effect of all other indexes, the highest partial correlations of the combined PISA score were observed for “Sec. Ed” and “GDI”. This result implies that the conventional Pearson correlations can be misleading to see how the PISA scores are related to other development indexes because the indexes that are highly related to the PISA scores are totally different when investigated with a network approach using regularized partial correlations.

Table 1. Pearson Correlations and Regularized Partial Correlations Between Development Indexes

		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
1	PIS	-			.07	-.22		.08	.03	-.18	.06			.23
2	HDI	.77**	-					.14	.09		.11	.16		
3	Urb%	.15	.47**	-			-.04			.10	-.20		-.01	-.12
4	Mort	-.37**	-.49	-.24*	-									.10
5	GDI	.34**	.24	-.12	-.25*	-	-.04	-.13		.30	-.03	.04		-.06
6	GII	-.79**	-.80**	-.12	.47**	-.19	-	.15			.06			.03
7	Ineqlife	-.74**	-.84**	-.30*	.55**	-.25*	.85**	-	.16	-.25	.08			
8	Internet	.63**	.84**	.51**	-.60**	.25*	-.68**	-.79**	-	-.02	.31	.20	-.09	
9	Mobile	.05	.10	.06	-.05	.20	.00	-.06	.05	-	-.24			-.10
10	Migr	-.08	.08	.43**	-.08	-.18	.20	-.16	.30*	.03	-			.13
11	Agr	-.42**	-.71**	-.64**	.42**	-.05	.46**	.54**	-.69**	-.14	-.29*	-	-.04	
12	Unemp	-.23	-.11	-.17	-.20	-.20	-.12	-.07	-.16	-.14	-.27*	.08	-	
13	Sec. Ed	.67**	.62**	.02	-.51**	.36**	-.60**	-.65**	.63**	.06	-.04	-.28*	-.13	-

Note: the lower diagonal values are the Pearson correlations; the upper diagonal values are regularized partial correlations; the missing parts in upper diagonal represent zero correlations estimated as such after regularization process; *p<0.05; **p<0.01.

The results for the estimated network are shown graphically in Figure 1 below. As figure was examined, it could be seen that some of the edges were represented by green lines while red lines represented some others. That is, regularized partial correlation coefficients calculated for some connections had a negative value. In this prediction, 37 of the 78 edges (47.4%) between the nodes were estimated differently from zero. It implies that, all the indicators are tightly connected directly or indirectly.

Furthermore, the edges are not of equal thickness. Some edges are thicker, while others are thinner. This is an indication that some of the nodes are relatively more connected with each other. In the figure, thicker edges represent stronger connections. The strongest link is between "*the Inequality in life expectancy*" - "*the life expectancy at birth*" and "*the Inequality in life expectancy*" – “*The percentage of the total population access to the internet*" indicator pairs. However, the association between those node pairs were negative. On the other hand, the highest positive association was observed between “*GDP*” and “*the percentage of total population access to the internet*” node pair.

If we look at the development indicators associated with the PISA results, the results showed seven edges connecting the PISA results to other indicators in the network. These indicators were “*the Gender Inequality Index*”, “*the percentage of population with at least some secondary education*”, "*the Inequality in life expectancy*", “*the Gender Development Index*”, “*the life expectancy at birth*”, “*the percentage of total population access to the internet*” and (GDP). Especially, the association of the PISA results with gender inequality, the percentage of the population with at least some secondary education and inequality in life expectancy was found as higher. Interestingly, the association of PISA outcomes with GDP is relatively lower when all other development indices in the network were controlled. On the other hand, no association was found between the results of PISA and "unemployment as a percentage of total labor force", "net migration rate", "total population living in urban areas", "unsafe water mortality rate", "sanitation and hygiene services", "current health expenditure as a percentage of GDP" and "life expectancy inequality". This result implies that when we control for other variables in the network, these indicators are completely independent of the results from PISA.

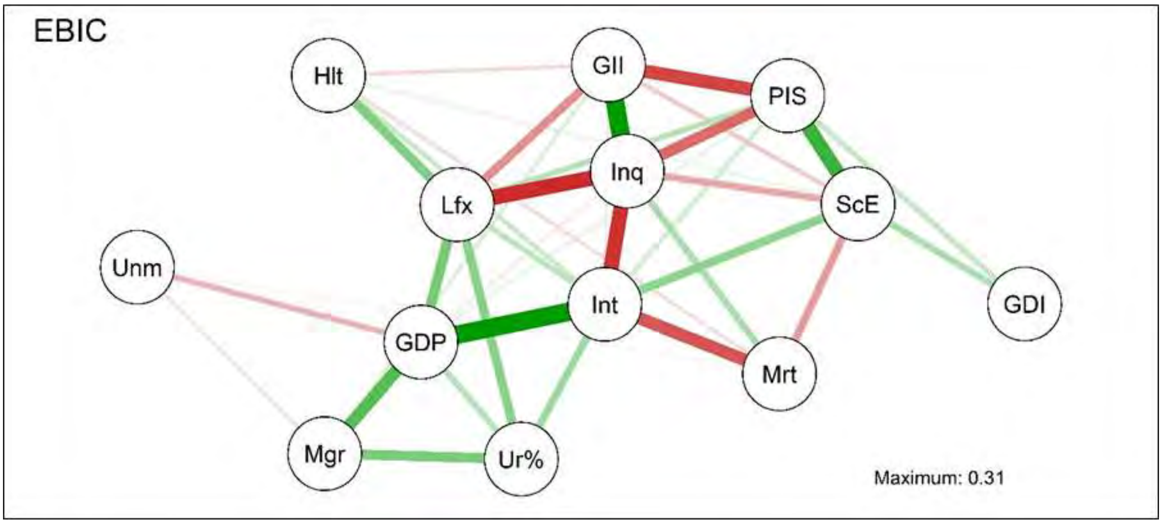

Figure 1. Centre Estimated network for human development indicators

One of the most important features of network analysis is that it identifies the most central and peripheral nodes using centrality indices. As mentioned earlier, the more connections a node has, the more centrally it is located and the centrality is determined mainly by strength, closeness and betweenness indices. The centrality criteria for the model predicted for development indicators were given in Figure 2. Each centrality value is standardized so they can be compared and they are on the same scale. Accordingly, "*percentage of total population access to the internet*", "*the inequality in life expectancy*" and "*the life expectancy at birth*" are the indicators with highest centrality indices. On the other hand, the PISA results was classified the fifth indicator in terms of the magnitude of strength centrality indice. In indice value of the PISA results is slightly less than the one for GDP indicator. This result implies that the PISA results are not at the top four development indicators in terms of the importance in the network.

When the index of closeness was analyzed, slightly similar results are observed. Accordingly, the highest values were estimated for the "*the percentage of total population access to the internet*", "*the inequality in life expectancy*" and "*GDP*" indicators, while The PISA results took only the seventh position in terms of the magnitude of closeness index. The result implies that the PISA results is relatively more distant than some other indicators in the network, less indirect connection and less close to them. In addition, it has relatively less predictive power in comparison to the indicators with higher closeness index values.

Lastly, the results for betweenness index values revealed similar findings. The highest values were observed for "*the percentage of total population access to the internet*", "*the inequality in life expectancy*" and "*GDP*" indicators. On the other hand, the PISA results took sixth position in terms of the magnitude of the betweenness indice value. The result suggests that PISA results are relatively less important for the interactions of other indicators compared to the indicators with higher betweenness values. In other word, the PISA results provide a bridge for the connections of other indicators less than some of the other indicator.

In addition, the Small-worldness index has been calculated as a global indicator of centrality. This value was estimated to be 1.38. Humphries & Gurney (2008) stated that if the small World index calculated for a network is greater than 1, the network has small-world property. Accordingly, this finding is evidence that the network of development indicators shows small world topography. The main reason for this situation is that development indicators form a high-density network.

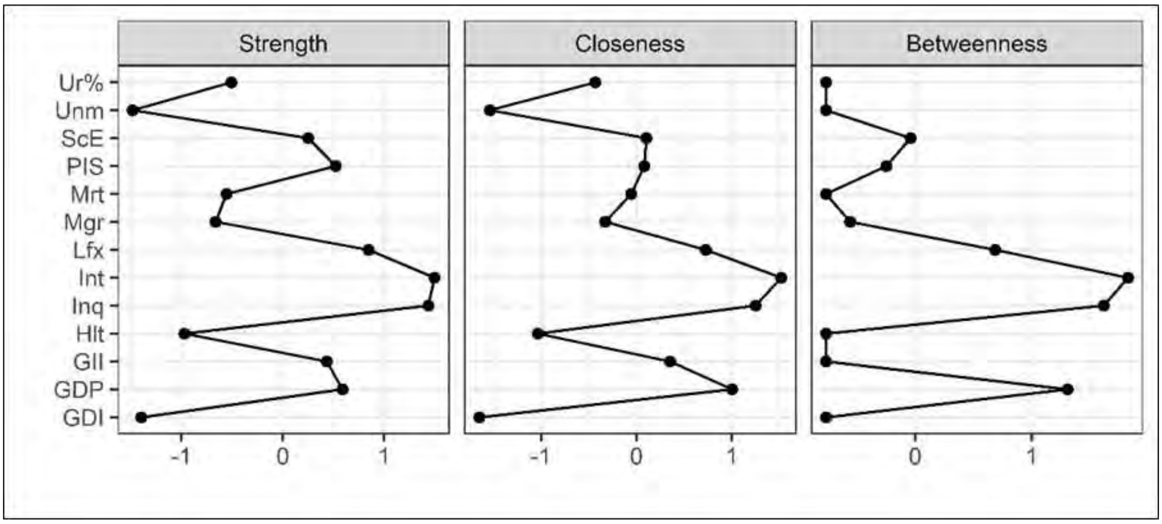

Figure 1. Standardized centrality indices for human development indicators

One The results for Zhang's clustering coefficients are shown in Figure 3. According to this, by far the highest value was observed for the GDP indicator. It can be said that this indicator is the most locally abundant. The local frequency of these nodes depends on the connections of the neighboring nodes. For example, "net migration rate" and "share of total population living in cities" are neighboring nodes of the GDP indicator, but they are already connected and do not need GDP to stay connected. In other words, for a country, the net migration rate and the percentage of the total population living in urban areas are conceptually linked indicators by themselves. This relationship exists independently of the GDP of that country. On the other hand, the indicator PISA has been classified as the fifth rich indicator and plays a relatively larger role in the relationship with other indicators.

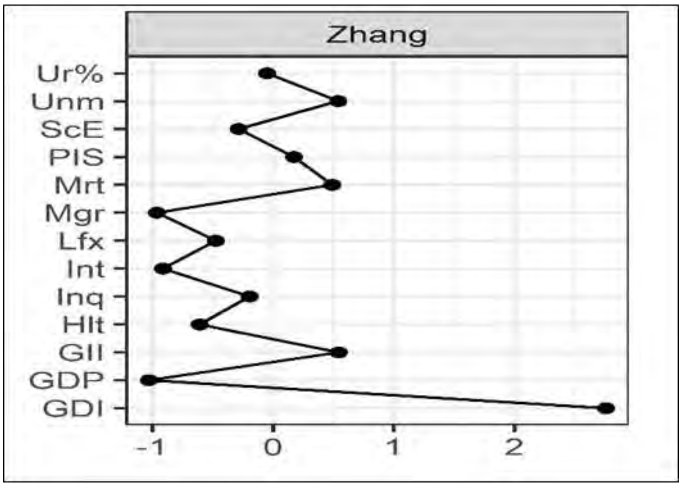

Figure 1. Clustering coefficient plots of human development indicators

Conclusion

As stated in the introduction part, this study aims to understand the network structure of the development indicators of United Nations and see how central the PISA results in this network. The results showed that all of the development indicators in the estimated network were connected directly or indirectly. Accordingly, they form a “small world” of a network. In small networks, the nodes need fewer intermediary nodes to connect indirectly (Watts & Strogatz, 1998). The practical implication of this finding is that, all the development indicators are directly or indirectly related but need less mediator nodes to be connected. They form a tightly interconnected network. In addition, one more implication is that, none of the indicators are abundant and contribute to the formation of the development indicators network.

On the other hand, some of these observed connections were positive while some others were negative. This result was clearly expected one because, some of the development indicators focus on negative outcomes. In contrast, some others focus on the negative outcomes and their regularized partial correlations become positive or negative accordingly.

Centrality analysis has shown that some indicators had more influence on the network than the others. For example, the percentage of total population access to the internet, the inequality in life expectancy and the life expectancy at birth were the most central indicators on the network and, therefore, had a major impact on the persistence of other indicators. In addition, they have a potential to activate other indicators when they are activated. When this result was evaluated in terms of the PISA scores, if a country wants to enhance its PISA success, they need to invest more to improve internet access, equality in life expectancy, and average life expectancy. On the other hand, the GDI and unemployment rates were the least central indicators and have little effect on other developmental indicators, including the PISA scores. This result can also be deduced intuitively because the GDI (which is related to the gap in living standards by gender) and the unemployment rates are the result rather than the effect compared to other developmental indicators.

On the other hand, the PISA scores seemed to locate at middle position in terms of its centrality among other development indicators. Accordingly, the PISA scores are not the cause or the effect of development in absolute way. It effects some indicators while it is affected by some other indicators. For instance, internet access (as the most central node in the network) affect the PISA scores which further affects the GDP and the unemployment rates. Hence, as Chabbott & Ramirez (2000) pointed out, education is seen as a means to foster development. On the other hand, the current study results showed that the causal function between education and other development indicators is not totally unidirectional.

Looking at the practical implications, it is clear that policy makers should focus on the indicators with the higher centrality indices. After all, if one changes the most central nodes in a network model, it is expected that others will be changed as well, without having to make additional efforts to change the least central nodes.. The interventions that aim to improve those central indicators would later improve other indicators, including the PISA results. Today, the PISA exam results, their effects, and how to improve these results are generally discussed mainly by educators. They focus on the relationship between the PISA results and other education-related variables (i.e. Meroni, Vera-Toscano, & Costa, 2015; Mikk, Krips, Säälik, & Kalk, 2016).

On the other hand, the experts in other fields generally discussed the educational issues in economic approaches (Crespo, 2002). The findings from this current study showed that the PISA results are not the most central member of the development indicators in the network and that PISA results will not improve until more centrally located PISA indicators improve. For this reason, it should not be forgotten that PISA results are a part of a tightly intertwined development network, and it would be right for experts from different fields to work in cooperation to improve the PISA results. By combining the information from experts of different areas (economy, social services, etc.), policymakers could take more efficient actions in improving educational outcomes.

Recommendations

With this study, a better understanding of the structural relationship between different development indicators in the participating countries of PISA can be gained. Moreover, a clear picture of how the results of PISA are related to other development indicators can be obtained. However, in future studies, it would be valuable to compare the development indicators of OECD countries and other countries that later participated in the program. In this way, the impact of the inclusion of new countries in PISA on the network structure can be studied. Such a study can show how the situation of PISA results changed with the inclusion of other countries. Finally, as mentioned earlier, network modeling can be studied longitudinally. The network structure of development indicators can be studied in this way. By estimating network structure over time, researchers can understand how the centrality of PISA outcomes (and their influence on other development indicators) in the network changes over time.

References

Abelson, R. P. (2012). *Statistics as principled argument*. Ney York, NY: Psychology Press.

Afşar, M. (2009). Türkiye'de eğitim yatırımları ve ekonomik büyüme ilişkisi. *Anadolu Üniversitesi Sosyal Bilimler Dergisi*, *9*(1), 85-98.

Baldacci, E., Guin-Sui, M. T., & de Mello, L. (2004). More on the Effectiveness of Public Spending on Healthcare and Education: A Covariance Structure Model. *Journal of International Development*, *15*(6), 709-725.

Barrat, A., Barthélemy, M., Pastor-Satorras, R., & Vespignani, A. (2004). The architecture of complex weighted networks. *Proceedings of the National Academy of Sciences*, *101*(11), 3747-3752.

Barro, R.J., (1997). *The Determinants of Economic Growth*. Cambridge, MA: MIT Press.

Beard, C., Millner, A. J., Forgeard, M. J., Fried, E. I., Hsu, K. J., Treadway, M. T., ... & Björgvinsson, T. (2016). Network analysis of depression and anxiety symptom relationships in a psychiatric sample. *Psychological medicine*, *46*(16), 3359-3369.

Bijou, M., & Liouaeddine, M. (2018). Gender and Students' Achievements: Evidence from PISA 2015. *World Journal of Education*, *8*(4), 24-35.

Bloom, D. E., Canning, D., Chan, K. J., & Luca, D. L. (2014). Higher education and economic growth in Africa. *International Journal of African Higher Education*, *1*(1), 22-57.

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry*, *16*, 5-13.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology*, *9*, 91-121.

Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, *25*(2), 163-177.

Chabbott, C. & Ramirez, F. O. (2000). Development and education. In *Handbook of the Sociology of Education* (pp. 163-187). Boston, MA: Springer.

Chen, J. & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*. *95*(3), 759-771.

Cinnirella, F., & Streb, J. (2017). The role of human capital and innovation in economic development: evidence from post-Malthusian Prussia. *Journal of economic growth*, *22*(2), 193-227.

Clark, D., & Royer, H. (2013). The effect of education on adult mortality and health: Evidence from Britain. *American Economic Review*, *103*(6), 2087-212.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. J. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13-29.

Cramer, A. O. J., Waldorp, L., van der Maas, H., & Borsboom, D. (2010). Comorbidity: A Network Perspective. *Behavioral and Brain Sciences*, *33*(2-3), 137-15.

Cramer, A.O., Sluis, S.V., Noordhof, A., Wichers, M., Geschwind, N., Aggen, S.H., Kendler, K.S., & Borsboom, D. (2012). Dimensions of Normal Personality as Networks in Search of Equilibrium: You Can't Like Parties if You Don't Like People: Dimensions of normal personality as networks. *European Journal of Personality*, *26*(4), 414-431.

Crespo, P. F. (2002). *An analysis of the relationship between higher education and development by applying Sen's human capabilities approach: The case of three technological universities in Mexico* (Doctoral dissertation, University of York).

De Nooy, W., Mrvar, A., & Batagelj, V. (2011). *Exploratory social network analysis with Pajek* (2nd ed.). Cambridge: Cambridge University Press.

Epskamp, S., Maris, G. K., Waldorp, L. J., & Borsboom, D. (2016). Network psychometrics. *arXiv preprint arXiv:1609.02818*.

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*, 195-212.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J. ,Schmittmann V. D. & Borsboom, D. (2012). qgraph: Network Visualizations of Relationships in Psychometric Data. *Journal of Statistical Software*, *48*(4), 1-18.

Foygel, B.R. & Drton, M. (2010). Extended Bayesian information criteria for Gaussian graphical models. *Advances in Neural Information Processing Systems*, *23*, 2020-2028.

Foygel B.R., & Drton, M. (2015). High-dimensional Ising model selection with bayesian information criteria. *Electronic Journal of Statistics*, *9*(1), 567-607.

Freeman, L. C. (1978). Centrality in social networks conceptual clarification. *Social Networks*, *1*(3), 215-239.

Fried, E. I. (2015). Problematic assumptions have slowed down depression research: why symptoms, not syndromes are the way forward. *Frontiers in Psychology*, 6, 309.

Friedman, J. H., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, *9*(3), 432-441.

Fruchterman, T. M., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Software: Practice and experience*, *21*(11), 1129-1164.

Ganegodage, K. R., Rambaldi, A. N., Rao, D. P., & Tang, K. K. (2006). *Analysis of Development Performance Using a Development Index Based on Factor Analysis (Old Title-Constructing Multideminsional Indexes of Economic Structure and Development)* (No. WP052006). School of Economics, University of Queensland, Australia.

Grek, S. (2009). Governing by numbers: The PISA "effect" in Europe. *Journal of Education Policy*, *24*(1), 23-37.

Humphries, M. D., & Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PloS one*, *3*(4), e0002051.

Kelava, A., & Brandt, H. (2009). Estimation of nonlinear latent structural equation models using the extended unconstrained approach. *Review of Psychology*, *16*, 123-131.

Kim, M., & Leskovec, J. (2012). Latent Multi-group Membership Graph Model. *arXiv preprint arXiv:1205.4546*.

Lauritzen, S. L. (1996). *Graphical Models.* Oxford Statistical Science Series. volume 17. Oxford: Clarendon Press.

Masuda, N., Sakaki, M., Ezaki, T., & Watanabe, T. (2018). Clustering Coefficients for Correlation Networks. *Frontiers in Neuroinformatics*. 12, 7.

McNally, R. J., Robinaugh, D. J., Wu, G. W. Y., Wang, L., Deserno, M. K., & Borsboom, D. (2015). Mental disorders as causal systems: A network approach to posttraumatic stress disorder. *Clinical Psychological Science*, *3*(6), 836–849.

Mcnally, R.J. (2016). Can network analysis transform psychopathology? *Behaviour research and therapy*, *86*, 95-104.

Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *The annals of statistics*, *34*(3), 1436-1462.

Meroni, E. C., Vera-Toscano, E., & Costa, P. (2015). Can low skill teachers make good students? Empirical evidence from PIAAC and PISA. *Journal of Policy Modeling*, *37*(2), 308-323.

Mikk, J., Krips, H., Säälik, Ü., & Kalk, K. (2016). Relationships between student perception of teacher-student relations and PISA results in mathematics and science. *International journal of science and mathematics education*, *14*(8), 1437-1454.

Newman, M. E. J. (2010). *Networks: An introduction*. New York, NY: Oxford University Press.

OECD (2001). *Knowledge and Skills for Life: First Results from PISA 200.* Paris: OECD Publishing.

OECD (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. Paris: OECD Publishing.

OECD. (2010). *Education at a glance 2010: OECD indicators*. Paris: OECD publishing.

OECD. (2017). *PISA 2015 assessment and analytical framework: science, reading, mathematic, financial literacy and collaborative problem solving (Revised Edition)*. Paris: OECD publishing.

OECD. (2018). *PISA for development: results in focus*. Paris: OECD publishing.

OECD. (2019). PISA 2018: *Insights and interpretations*. Paris: OECD publishing.

Onnela, J. P., Saramaki, J., Kertesz, J., & Kaski, K. (2005). Intensity and coherence of motifs in weighted complex networks. *Physical Review E*, *71*(6), 065103.

Pearson, K. (1897). Mathematical contributions to the theory of evolution.—on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the royal society of london*, *60*(359-367), 489-498.

Pongratz, L. A. (2006). Voluntary Self-Control: Education reform as a governmental strategy. *Educational Philosophy and Theory*, *38*(4), 471-482.

R Core Team (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Snijders T.A.B. (2009) Network Analysis, Longitudinal Methods of. In: Meyers R. (Eds.) *Encyclopedia of Complexity and Systems Science*. New York, NY: Springer.

Takayama, K. (2009). Politics of externalization in reflexive times: Reinventing Japanese education reform discourses through "Finnish PISA success". *Comparative Education Review*, *54*(1), 51-75.

Qiu, Q., Sung, J., Davis, W., & Tchernis, R. (2018). Using spatial factor analysis to measure human development. *Journal of Development Economics*, *132*, 130-149.

van Borkulo, C. D., Borsboom, D., Epskamp, S., Blanken, T. F., Boschloo, L., Schoevers, R. A., & Waldorp, L. J. (2014). A new method for constructing networks from binary data. *Scientific reports*, *4*(1), 1-10.

van Borkulo, C., Boschloo, L., Borsboom, D., Penninx, B. W., Waldorp, L. J., & Schoevers, R. A. (2015). Association of symptom network structure with the course of depression. *JAMA psychiatry*, *72*(12), 1219-1226.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of "small-world" networks. *Nature*, *393*(6684), 440-442.

Zhang, B., & Horvath, S. (2005). A general framework for weighted gene co-expression network analysis. *Statistical Applications in Genetics and Molecular Biology*, 4.