A Multi Criteria Approach For Statistical Software Selection in Education

Eğitimde İstatistiksel Yazılım Seçimine Çok Kriterli Bir Yaklaşım

Zeliha KAYGISIZ ERTUĞ[*], Nuray GİRGİNER[**]

ABSTRACT: Statistical software is commonly used in the statistical lessons at universities. The developments and enhancement in statistical software in recent years has considerably eased statistics education in these institutions. The purpose of this study is to develop an evaluation model considering the quantitative and qualitative criteria for statistical software selection in an outsourcing user of these programs variety fields, especially in education. An integrated model is proposed by combining Analytic Hierarchy Process (AHP) and Grey Relation Analysis (GRA) into a single evaluation model. The model is illustrated with a case study of a team of five people including academics and software developers well versed in the use and development of such software to demonstrate the effectiveness of this integrated method. AHP has been applied to determining weight of criteria and GRA has been performed for determining the most appropriate statistical software. The results indicate that when analysis characteristics are the main criteria with the highest priority, financial and vendor firm characteristics are the main criteria with the lowest priorities. Also according to GRA results, the most appropriate statistical software is SPSS and Statgraph is in last rank with a low level of significance.
Keywords: Statistics education, statistical software, analytic hierarchy process, grey relation analysis.

ÖZ İstatistiksel yazılımlar, üniversitelerdeki istatistik derslerinde sıklıkla kullanılmaktadır. Son yıllarda istatistik yazılımlarının iyileştirilerek geliştirilmesi, bu kurumlarda verilen istatistik eğitimini de büyük ölçüde kolaylaştırmıştır. Bu çalışmanın amacı, eğitim başta olmak üzere çok çeşitli alanlarda kullanılan istatistiksel yazılımların seçimi için nicel ve nitel kriterlerin bir arada değerlendirildiği bir model geliştirmektir. Çalışmada Analitik Hiyerarşi Süreci (AHS) ve Gri İlişkisel Analizin (GİA) birlikte kullanımıyla bütünleşik bir model önerisinde bulunulmuştur. Söz konusu modelin etkililiğini ortaya koyabilmek için, gerek kullanıcı gerekse programcı olarak uzman olan akademisyenler ve programcılardan oluşan beş kişilik bir ekiple bir örnek olay uygulaması yapılmıştır. Kriterlerin ağırlıklarını belirlemek için AHS, en uygun istatistiksel yazılım seçimini gerçekleştirmek için ise GİA kullanılmıştır. Elde edilen sonuçlar, en yüksek önceliğe sahip olan ana kriterin analiz özellikleri, en düşük önceliklere sahip olan ana kriterlerin ise finansal özellikler ile satıcı firma özellikleri olduğunu ortaya koymuştur. Aynı zamanda GİA sonuçlarına göre SPSS en uygun istatistiksel yazılım olarak belirlenirken, Statgraph en düşük önem derecesine sahip yazılım olmuştur.
Anahtar sözcükler: İstatistik eğitimi, istatistiksel yazılım, analitik hiyerarşi süreci, gri ilişkisel analiz.

1. INTRODUCTION

Various statistical methods are commonly used in the analysis of obtained data and the interpretation of the findings in today's scientific studies due to the wide spreading of the statistical methods and techniques in all areas starting especially from the last quarter of the 20th century along with the changes in the priorities of researchers in this field. However, since the manual solution of the required statistical analyses of complex and multi-variable problems containing mass data takes too long and has a high probability of error, the use of software developed for this purpose has become obligatory. The determination of software most suited to the data during the analysis and evaluation stage is important for the reliability of the obtained results. Thus, the selection of the proper statistical software is more important than the carrying out of the statistical operations.

[*] Assist. Proff. Dr, Eskisehir Osmangazi University Faculty of Economic and Administrative Sciences, Eskişehir, zelihak@ogu.edu.tr

[**] Proffessor Dr, Eskisehir Osmangazi University Faculty of Economic and Administrative Sciences, Eskişehir, ngirginer@gmail.com

Statistical software differs from each other in many different areas (the technical support of the vendor, user friendliness of the software, ease of database access, ability to generate output-reports, included methods etc.). Thus, the satisfaction of the demands of the user depends on which software is chosen. For example, one of the places where such software is commonly used is the statistical lessons at universities. The developments and enhancement in statistical software in recent years has considerably eased statistics education in these institutions. Such software enables the students to gain information and develop skills regarding the statistical analysis and interpretation of results obtained during the study. This in turn brings forth the necessity to incorporate computer applications as well as theory during statistics education. However, the selection is not correct if the selected software does not coincide with the solution of the problems that the students will face during their professional lives following graduation.

On the other hand, selection of the proper software is important to ensure the quality of the scientific studies carried out by universities apart from education. Significant problems arise when the user does not clearly know the algorithms and solution formulations included in the software. The reliability of the analyses may be problematic when the users are people without any education in statistics. Since the proper and effective use of statistical software will provide benefits; it is important that researchers show the attention and sensitivity they employ during the planning and execution of their studies in the selection of the proper software. However, since the required statistical software necessary for both statistics education and scientific studies cannot be obtained as part of their budgets, it is an obligation that such software is purchased via the faculty budget. The limited budgets of faculties along with the fact that it is not possible to purchase all the software demanded by academics pose a problem of selecting software. In this study, the necessity of statistical software both for statistics education and scientific studies along with the problem of statistical software selection due to limited budgets has been examined.

Technical features, after sales services provided by the selling company, included analysis properties, ease of use, speed, graphics support, and interactivity between databases are considered in addition to the cost when making the software decision. In other words, there are many qualitative and/or quantitative criteria that affect the decision of the researcher. Therefore, a selection method should be used where priorities of the factors with respect to each other can be evaluated for all criteria concurrently. In practice, there is no standard method for this selection process. In addition to methods such as cost-benefit analyses, gradation, risk analyses and scoring; multi-criteria methods such as Analytic Hierarchy Process (AHP), Analytic Network Process (ANP) are used. On the other hand, Grey Relational Analysis (GRA) is another method that can be used for software selection enabling the selection of the best alternative for all criteria based on the closeness with the best alternative among all alternatives for problems where more than one alternative are considered in the existence of more than one criteria. GRA successfully selects the best alternative in total by calculating the closeness of an alternative in all criteria with the best and worst values for each respective criterion.

In this study, the weights of the criteria and the sub-criteria that are thought to be effective in the selection of statistical software obtained via AHP have been used as data in GRA. Therefore, a degree of significance has been put forth for each alternative (software) via GRA aiming to select the most suitable statistical software with the highest degree of significance..

1.1. Literature Review

The large number of criteria that should be evaluated together has resulted in the frequent use of AHP in software selection literature. This complex and multi-criteria decision problem has attracted the attention of many researchers thereby increasing the number of relevant studies in literature. Başlıgil (2005) has studied the selection of software for highest level of customer satisfaction via fuzzy AHP; Koçak (2003) and Çörekçioğlu and Güngör (2005) have examined the selection of the most suitable enterprise resource planning (ERP) software via AHP. Various international study examples can be given in which AHP has been used for software selection. Santhanam and Kyparisis (1996) have developed a non-linear model for software selection and have compared the relevant software based on the relationship between criteria whereas Jalal and Ray (1999) have examined the evaluation methods that can be used to select software for production simulation. Lai et.al (1999, 2002) have used AHP to carry out comparisons for software alternatives in multi-environment control systems based on group decisions according to two main criteria which they have named as technical features and executive features. Wei, Chien and Wang (2004) have set out from AHP method to develop a model enabling ERP selection according to the needs of an enterprise and during their studies have observed that this method is quite accessible for executives and yields productive results. Lee, Shen and Chih (2004) have developed a multi-criteria method for software selection. In a similar study, Mulabeke and Zheng (2006) have examined the selection of software used in product development via analytic network process. Ayağ and Özdemir (2007) have examined ERP software selection using analytic network process whereas Perera and Costa (2008) have used AHP and Yazgan et.al. (2009) have used artificial neural networks and analytic network process methods.

The fact that most errors in statistical calculations arise due to user or statistical software errors has led researchers to carry out studies regarding the reliability of statistical software. Studies in literature regarding statistical software focus mostly on the comparison of the reliabilities of statistical software with the exception of the study carried out by Girginer and Kaygısız (2009). For instance; Altman and McDonald (2001) suggest a guide regarding the selection of reliable statistical software. In addition, there are also studies which compare widely used statistical software. For instance; Dielman (2002) has compared four statistical software widely used in business administration departments whereas Kitchen et.al. (2003) have compared six of statistical softwares two of which are web based. The development of statistical software among estimation software has been included in the study carried out by Küsters et.al. (2006). Last study about this topic has been carried out by Keeling and Pavur (2007) in which they have compared the reliability of nine statistical software. Different from these studies, Girginer and Kaygısız (2009) have evaluated the selection of the best statistical software that will be used for academic studies as well as during education in universities from among three statistical software using AHP and 0-1 Goal Programming (GP) methods according to five basic criteria each of which include sub-criteria.

As seen in the literature review, no studies were found on the evaluation of statistical softwares using GRA; meanwhile, no study in which the statistical softwares were ranked by their priorities by obtaining the weights of criteria and sub-criteria by applying AHP was encountered excluding the study carried out by Girginer and Kaygısız (2009). It has been stated by examples above that apart from the study carried out by Girginer and Kaygısız (2009), a single method is used for the comparison of software used in fields outside of statistics instead of combined methods. Also AHP determine the most appropriate alternative instead of sorting them. However using GRA, alternatives can be sorted. All these reasons in this study, GRA had been used together with AHP which is another multi-criteria decision making technique. The objective of the study is to determine the criteria considered during the selection of statistical software which should be evaluated together along with their priorities and to make the most suitable statistical software decision using GRA. It is thought that this study and its results will be helpful to all parties faced with such a decision.

2. METHOD

The selection of the most appropriate statistical program alternative involves multiple objectives or/and criteria and hierarchy process. In this study used analytical hierarchy process (AHP) proposed by Saaty, (1980) and grey relation analysis (GRA) to select the statistical program alternatives. AHP provides an optimal solution considering both qualitative and quantitative aspects of a decision. Another advantage of AHP is that it reduces the level of comparison from large number of factors to few. GRA provide the optimal decision making using relational coefficient matrices which is used to compute the weights to the criteria. The objective of this study is to apply the AHP and GRA for optimal selection of statistical software alternative.

2.1. Determining the Weights of Criteria by Using AHP

As a decision method that decomposes a complex multi-criteria decision problem into a hierarchy (Saaty 1980, 1994; Tung and Tang, 1998; Lee et al., 1999; Macharis et al., 2004), AHP is also a measurement theory that prioritizes by group of decision makers. Assuming there is no interdependence among sub-criteria should be more emphasized in determining the perspective upper level criteria. AHP incorporates the evaluations of all decision makers into a final decision (Javalgi et al., 1989; Forman and Peniwati, 1998; Chou et al., 2004; Chang et al., 2007a; Wu et al., 2008), without having to elicit their utility functions on subjective and objective criteria, by pair-wise comparisons of the alternatives (Saaty, 1990; Lipovetsky, 1996; Saaty, 2000; Altuzarra et al., 2007).

Prior to the AHP application of the statistical software selection problem, a team of five people including academics and software developers well versed in the use and development of such software was assembled. These experts consisted of a computer software developer, a computer teacher, an academic giving statistics lessons and two other academicians who use statistical software in other numeric fields (marketing, medicine, etc.). In line with the interviews carried out with the team and literature survey studies regarding software selection, criteria and sub-criteria that should be considered in the selection of statistical software the hierarchical structure were determined and the hierarchical structure was prepared as shown in Figure 1 for four widely used software (SPSS, Statistica, Minitab and Statgraph). Five basic criteria (financial, technical, usage, analysis and vendor firm properties) all of which include other sub-criteria were considered during the formation of the hierarchical structure.

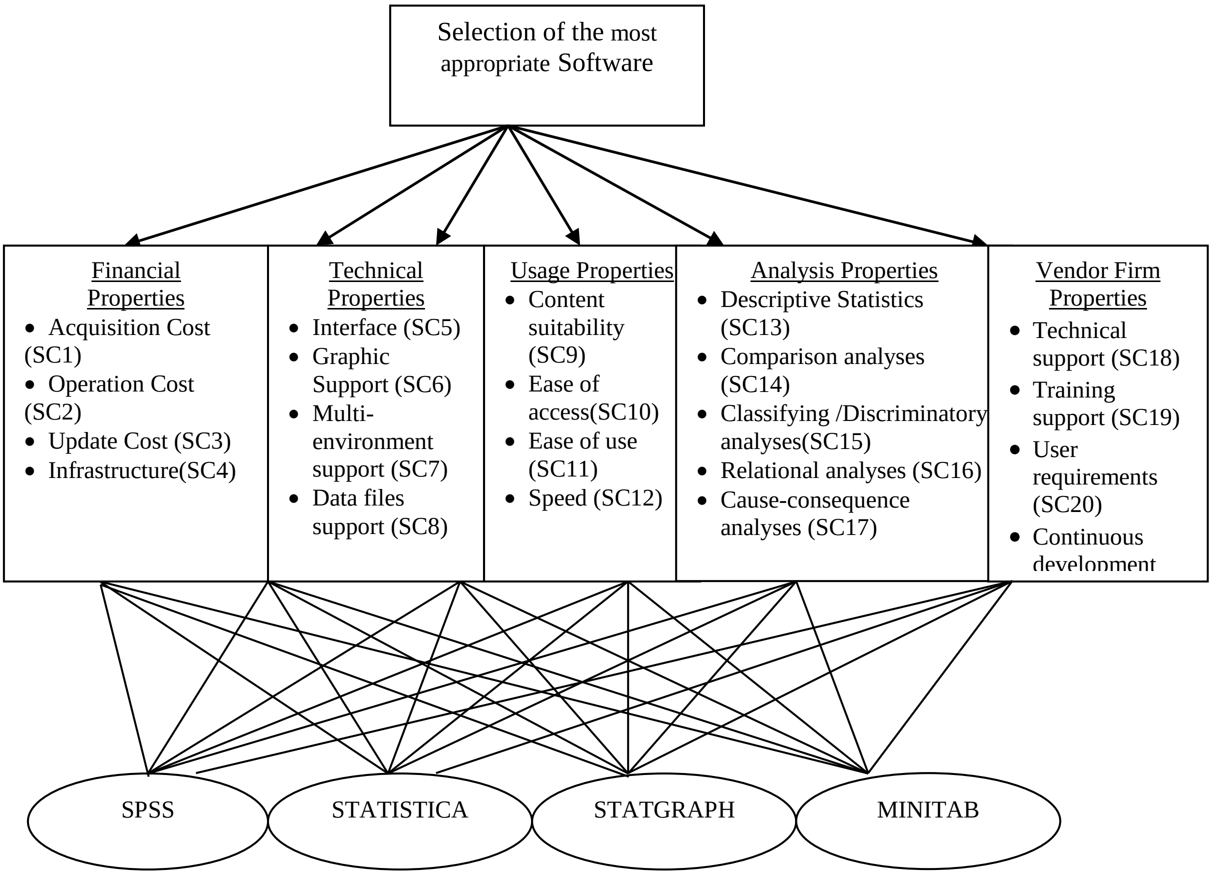

Figure 1: Hierarchical Structure of Problem

The next important stage of the AHP decision process following the formation of the hierarchical structure of the problem is obtaining paired comparison values. Survey form has been prepared based on the paired comparison matrices formed for each level of the hierarchy given in Figure 1. Using the scale suggested by Saaty, participants have judged the paired comparisons between the alternatives listed in the survey form, main criteria and sub-criteria according to their levels of significance. Thinking that the judgments of people who have knowledge about statistical software through their work as academic or through trainings will differ, the geometric mean was calculated and AHP group decision making was applied.

Following the paired comparisons for all main and sub-criteria, the values obtained for each main and sub-criteria based on respective significance levels and alternatives were obtained via group decision making and listed in Table 1.

Table 1: AHP Result Matrix for the Statistical Software Selection Problem

Main-Criteria	Sub-Criteria \ Alternatives	SPSS	Statistica	Minitab	Statgraph
Financial Properties (0.051)	SC1 (0.011)	0.388	0.262	0.230	0.120
	SC2 (0.011)	0.407	0.231	0.230	0.133
	SC3 (0.018)	0.390	0.212	0.271	0.127
	SC4 (0.010)	0.389	0.247	0.246	0.117
Technical Properties (0.166)	SC5 (0.025)	0.426	0.319	0.183	0.072
	SC6 (0.028)	0.178	0.365	0.210	0.247
	SC7 (0.046)	0.371	0.288	0.212	0.130
	SC8 (0.066)	0.426	0.248	0.199	0.128
Usage Properties (0.237)	SC9 (0.045)	0.346	0.277	0.280	0.097
	SC10 (0.038)	0.494	0.166	0.255	0.085
	SC11 (0.112)	0.416	0.215	0.204	0.165
	SC12 (0.042)	0.304	0.304	0.196	0.196
Analysis Properties (0.498)	SC13 (0.032)	0.295	0.311	0.264	0.129
	SC14 (0.114)	0.323	0.295	0.262	0.120
	SC15 (0.104)	0.399	0.298	0.219	0.084
	SC16 (0.107)	0.282	0.415	0.199	0.104
	SC17 (0.140)	0.366	0.375	0.161	0.098
Vendor Firm Properties (0.049)	SC18 (0.012)	0.413	0.223	0.240	0.124
	SC19 (0.014)	0.405	0.236	0.231	0.129
	SC20 (0.010)	0.324	0.298	0.249	0.129
	SC21 (0.013)	0.480	0.151	0.227	0.142

When Table 1 is examined, it is observed that the criteria with the highest importance in the selection of statistical software is analysis properties (%49.8) followed by the main criteria of usage properties (%23.7) and technical properties (%16.6). Financial properties (%5.1) and vendor firm properties (%4.9) were ranked last based on their levels of significance.

When financial properties are considered as a main criteria, the sub-criteria with the highest relative significance level among all sub-criteria came out to be update cost (%1.8) followed by acquisition cost and operation cost (%1.1) whereas infrastructure (%1.0) was ranked last. When technical properties are considered, it is observed that the sub-criterion with the highest level of relative significance is data files support (%6.6). This sub-criterion was followed by multi-environment support with a significance level of %4.6 and graphic support of the software with a significance level of %2.8. The sub-criterion with the lowest level of significance was the software interface with a value of %2.5. The ease of use for the software was the criteria with the highest significance level for the usage properties main criteria (%11.2) whereas the sub-criteria with the lowest significance level was ease of access with a value of %3.8. The sub-criteria with the highest significance level among analysis features main criteria were cause-effect analyses (%14). This was followed respectively by comparison analyses, relational analyses and classifying/discriminatory analyses whereas sub-criterion named descriptive statistics was ranked last with a significance level value of %3.2. For vendor firm features which is the last main criteria, training support (%1.4) was determined to be the most important sub-criterion and user requirements (%1) was determined to be the least important sub-criteria.

2.2. Determining the Most Appropriate Statistical Software by Using GRA

Deng (1982) introduced "The Grey System Theory" to supplement the limitations of using traditional statistical methods. Grey System Analysis (GRA) is useful for capturing the correlations between the reference factor and other factors which can be compared within a system (Deng, 1989). One of the features of GRA is that both qualitative and quantitative relationship can be identified among complex factors with insufficient information relative to conventional statistical methods. Under such conditions, the results generated by conventional statistical techniques may not be acceptable without sufficient data to achieve desired confidence levels. In contrast, grey system theory can be used to identify major correlations among factors of a system with a relatively small amount of data. Because of these features GRA has been extensively applied in many fields, such as financial institutions, hospitals, banks, airlines firms, etc.

The procedure for calculating the GRA is as follows (Wu et. al., 2010):

1. Calculate the Grey Relation Grade

Let X_0 be the referential series with k entities (or criteria) of X_1, X_2,, …X_i, ….X_N (or N measurement criteria). Then

X_0={x_0 (1), x_0 (2), …, x_0 (k)},

X_1={x_1 (1), x_1 (2), …, x_1(k)},

.

.

X_i={x_i (1), x_i (2), …, x_i(k)},

.

.

X_N={x_N (1), x_N (2), …, x_N(k)}

The grey relation coefficient between the compared series X_i and the referential series of X_0 at the j-th entity is defined as:

$$\gamma_{0i}(j) = \frac{\Delta \min + \Delta \max}{\Delta_{0i}(j) + \Delta \max}$$

(1)

Where $\Delta_{0i}(j)$ is the absolute value of difference between X_0 and X_i at the j-th entity, that is $\Delta_{0i}(j) = |x_0(j) - x_i(j)|$ and $\Delta_{\max} = Max_i \max_j \Delta_{0i}(j)$, $\Delta_{\min} = Min_i \min_j \Delta_{0i}(j)$

The grey relational grade (GRG) for series of X_i is given as:

$$\Gamma_{0i} = \sum_{j=1}^{K} w_j \gamma_{0i}(j)$$

(2)

Where, w_j is the weight of j-th entity. If it is not necessary to apply the weight, take $\omega_j = \frac{1}{K}$ as an average.

2. Data Normalization (or Data Dimensionless)

Before calculating the grey relation coefficients, the data series can be treated, based on the following three kinds of situation and the linearity of data normalization, to avoid distorting the normalized data (Hsia and Wu, 1997).

These are:

Benefit target: Upper-bound effectiveness measuring (i.e. larger-the-better)

$$x_i^*(j) = \frac{x_i(j) - \min_j x_i(j)}{\max_j x_i(j) - \min_j x_i(j)}$$

(3)

Cost Target: Lower bound effectiveness measuring (i.e. smaller-the-better)

$$x_i^*(j) = \frac{\max_j x_i(j) - x_i(j)}{\max_j x_i(j) - \min_j x_i(j)}$$

(4)

c) Medium Target: Moderate effectiveness measuring (i.e. nominal-the- best)

If $\min_j x_i(j) \le x_{ob}(j) \le \max_j x_i(j)$, then

$$x_i^*(j) = \frac{\left| x_i(j) - x_{ob}(j) \right|}{\max_j x_i(j) - \min_j x_i(j)}$$

(5)

If $\max_j x_i(j) \le x_{ob}(j)$, then

$$x_i^*(j) = \frac{x_i(j) - \min_j x_i(j)}{x_{ob}(j) - \min_j x_i(j)}$$

(6)

If $x_{ob}(j) \le \min_j x_i(j)$, then

$$x_i^*(j) = \frac{\max_j x_i(j) - x_i(j)}{\max_j x_i(j) - x_{ob}(j)}$$

(7)

Where $x_{ob}(j)$ is the objective value of entity j.

The GRA calculation process explained above has been applied as shown below in steps in line with the purpose of the study.

Step 1. Establishing decision making matrix: The weight is estimated for five experts with each respondent using Saaty's relative importance scale and averaging their scale to assess candidates, then establishing a decision making matrix as shown Table 2.

Table 2: The Decision-Making Matrix

Sub-Criteria	**Statistical Software**				
	Reference	**SPSS**	**STATISTICA**	**MINITAB**	**STATGRAPH**
SC1	0.388	0.388	0.262	0.230	0.120
SC2	0.407	0.407	0.231	0.230	0.133
SC3	0.390	0.390	0.247	0.246	0.117
SC4	0.389	0.389	0.247	0.246	0.117
SC5	0.426	0.426	0.319	0.183	0.072
SC6	0.365	0.178	0.365	0.210	0.247
SC7	0.371	0.371	0.288	0.212	0.130
SC8	0.426	0.426	0.248	0.199	0.128
SC9	0.346	0.346	0.277	0.280	0.097
SC10	0.494	0.494	0.166	0.255	0.085
SC11	0.416	0.416	0.215	0.204	0.165
SC12	0.304	0.304	0.304	0.196	0.196
SC13	0.311	0.295	0.311	0.264	0.129
SC14	0.323	0.323	0.295	0.262	0.120
SC15	0.399	0.399	0.298	0.219	0.084
SC16	0.415	0.282	0.415	0.199	0.104
SC17	0.375	0.366	0.375	0.161	0.098
SC18	0.413	0.413	0.223	0.240	0.124
SC19	0.405	0.405	0.236	0.231	0.129
SC20	0.324	0.324	0.298	0.249	0.129
SC21	0.480	0.480	0.151	0.227	0.142

The twenty one sub-criteria are "the-larger-the-better" because of each of them are compared according to which one more important. Accordingly the referential series can be X_0 = (0.388, …, 0.480). The statistical softwares (alternatives) are X_1 (SPSS), X_2 (Statistica), X_3 (Minitab) and X_4 (Statgraph).

Step 2. Normalizing Data: After establishing a decision making matrix (Table 2), it is established referential series can be X0 = {1.00, 1.00, 1.00, …, 1.00}. The statistical softwares are X1, X2, X3 and X4 . Data are normalized for 21 sub-criteria by using equations (6). Table 3 summarizes normalization data.

Table 3: Summary of Normalization Data

Sub-Criteria	Statistical Softwares				
	Reference (X_0)	SPSS (X_1)	Statistica (X_2)	Minitab (X_3)	Statgraph (X_4)
SC1	1.00	1.00	0.529	0.410	0.00
SC2	1.00	1.00	0.358	0.354	0.00
SC3	1.00	1.00	0.476	0.472	0.00
SC4	1.00	1.00	0.478	0.474	0.00
SC5	1.00	1.00	0.698	0.314	0.618
SC6	1.00	0.00	1.00	0.068	0.00
SC7	1.00	1.00	0.655	0.340	0.00
SC8	1.00	1.00	0.403	0.238	0.00
SC9	1.00	1.00	0.723	0.735	0.00
SC10	1.00	1.00	0.198	0.416	0.00
SC11	1.00	1.00	0.199	0.155	0.00
SC12	1.00	1.00	1.00	0.00	0.00
SC13	1.00	0.912	1.00	0.742	0.00
SC14	1.00	1.00	0.862	0.699	0.00
SC15	1.00	1.00	0.679	0.428	0.00
SC16	1.00	0.572	1.00	0.305	0.00
SC17	1.00	0.967	1.00	0.227	0.00
SC18	1.00	1.00	0.342	0.401	0.00
SC19	1.00	1.00	0.387	0.369	0.00
SC20	1.00	1.00	0.866	0.615	0.00
SC21	1.00	1.00	0.026	0.251	0.00

Step 3. Computing absolute values [$\Delta_{0i}(j)$ *]: A*bsolute value is the difference X0 (differential series) and Xi at the j-th sub-criteria. Computed absolute values are displayed in Table 4.

Step 4. Computing Grey Relation Coefficients [$\gamma_{0i}(j)$ *]:* The relational coefficients, $\gamma_{0i}(j)$ of the compared series are computed using equation 4. Table 5 presents the results.

Table 4: Absolute Values

Sub-Criteria	Statistical Software			
	SPSS (X_1)	Statistica (X_2)	Minitab (X_3)	Statgraph (X_4)
SC1	0.00	0.471	0.590	1.00
SC2	0.00	0.642	0.646	1.00
SC3	0.00	0.524	0.528	1.00
SC4	0.00	0.522	0.526	1.00
SC5	0.00	0.302	0.686	1.00
SC6	1.00	0.00	0.932	0.382
SC7	0.00	0.345	0.660	1.00
SC8	0.00	0.597	0.762	1.00
SC9	0.00	0.277	0.265	1.00
SC10	0.00	0.802	0.584	1.00
SC11	0.00	0.801	0.845	1.00
SC12	0.00	0.00	1.00	1.00
SC13	0.088	0.00	0.258	1.00
SC14	0.00	0.138	0.301	1.00
SC15	0.00	0.321	0.572	1.00
SC16	0.428	0.00	0.695	1.00
SC17	0.033	0.00	0.773	1.00
SC18	0.00	0.658	0.599	1.00
SC19	0.00	0.613	0.631	1.00
SC20	0.00	0.134	0.385	1.00
SC21	0.00	0.974	0.749	1.00

Table 5: Grey Relation Coefficients $\gamma_{0i}(j) = \dfrac{\Delta \min + P\Delta \max}{\Delta_{0i}(j) + P\Delta \max}$

Main Criteria	Sub-Criteria (w_j)	Statistical Softwares			
		SPSS (X_1)	Statistica (X_2)	Minitab (X_3)	Statgraph (X_4)
(0.5662)	**SC1 (0.5767)**	1.00	0.515	0.459	0.333
	SC2 (0.550)	1.00	0.437	0.436	0.333
	SC3 (0.561)	1.00	0.481	0.463	0.333
	SC4 (0.5772)	1.00	0.489	0.487	0.333
(0.5727)	**SC5 (0.595)**	1.00	0.623	0.424	0.333
	SC6 (0.5622)	0.333	1.00	0.349	0.567
	SC7 (0.589)	1.00	0.592	0.431	0.333
	SC8 (0.5447)	1.00	0.456	0.396	0.333
(0.5976)	**SC9 (0.6577)**	1.00	0.644	0.654	0.333
	SC10 (0.5445)	1.00	0.384	0.461	0.333
	SC11 (0.5222)	1.00	0.384	0.372	0.333
	SC12 (0.6665)	1.00	1.00	0.333	0.333
(0.6472)	**SC13 (0.7105)**	0.850	1.00	0.659	0.333
	SC14 (0.6852)	1.00	0.784	0.624	0.333
	SC15 (0.602)	1.00	0.609	0.466	0.333
	SC16 (0.5722)	0.538	1.00	0.418	0.333
	SC17 (0.666)	0.938	1.00	0.393	0.333
(0.5747)	**SC18 (0.553)**	1.00	0.424	0.455	0.333
	SC19 (0.5562)	1.00	0.450	0.442	0.333
	SC20 (0.6717)	1.00	0.789	0.565	0.333
	SC21 (0.518)	1.00	0.339	0.400	0.333

Step 5: Computing Grey Relation Grade: The sub-criteria weights from AHP and GRA are derived by equation (5). Sub-criteria weights (wj) obtained from AHP are shown second column in parentheses in Table 5. Equation 5 has been applied for each alternative (statistical software) as the ratio of the sum of grey relation coefficients multiply relative weights by the sum of the relative weights. Table 6 summarizes these results.

Table 6: Grey Relation Grades $\Gamma_{0i} = \sum_{j=1}^{K} w_j \gamma_{0i}(j)$

Statistical Software	Γ_{0i}	Rank
SPSS	%93.8	1
Statistica	%63.8	2
Minitab	%46.8	3
Statgraph	%34.46	4

As shown in Table 6, SPSS (%93.8) is the most appropriate statistical software. Other software follow such as Statistica (%63.8), Minitab (%46.8) and in last rank Statgraph with %34.46.

According to AHP results (Table 1), the main criterion with the highest importance in the selection of statistical software is analysis properties (0.498) followed by the main criteria of usage properties (0.237) and technical properties (0.166). Financial properties and vendor firm properties were ranked last based on their levels of significance. On the other hand according to GRA results (Table 6), first software is SPSS as the most appropriate statistical software. SPSS is followed by Statistica and Minitab, respectively. Statgraph is last in the ranking.

4. DISCUSSION and CONCLUSIONS

This study proposes an integrated approach for evaluating and selecting statistical software of AHP and GRA in a preference measurement model for users of these programs. The main advantage of this research is that it can be used for both qualitative and quantitative criteria. This study indicates that the AHP and GRA are powerful tools which can be multi criteria decision making problem as selecting statistical software problem. The proposed method comprises two parts. The first part employs the AHP to determine the weights of criteria. The second part applies the GRA to rank alternatives and select the best statistical software. In this model if new criteria added, they can be included in the proposed model to select the best software. Also any new potential software can be included in the evaluation process. Hence in comparison with other models proposed model is more flexible, applicable and effective.

According to the results of the AHP carried out to determine the main and sub-criteria affecting the most suitable software for researchers using statistical software; the main criteria were ranked according to their levels of significance as follows: analysis properties, usage properties, financial properties and vendor firm properties. Whereas according to the results of the GRA carried out to determine the most suitable statistical software; SPSS package software was determined as the most suitable software followed respectively by Statistica, Minitab and Statgraph. On the other hand, when the main criteria weights obtained according to the GRA results are examined, it is striking to see that similar to AHP results, the two main criteria with the highest level of importance were analysis properties and usage properties.

Since academicians purchase statistical software using the budget allocated by their universities, financial properties main criteria were not determined to be significant in this study according to both the AHP and the GRA results. In addition, when vendor firm properties criteria are considered, since users generally meet their own technical and training support requirements following the purchase and installation of the software, this criteria was also not specified as significant in the selection of statistical software. On the other hand, descriptive statistics property was determined to be a more significant analysis feature in comparison with the others whereas relational analysis was determined to be the least significant. With respect to the usage properties, the sub-criteria with the highest level of significance was determined to be the speed of the software whereas ease of use was determined to be the property with the lowest significance level (See Table 5).

Depending upon the importance of the statistical software mentioned in this study, the expectation among universities regarding software that researchers or academicians cannot purchase personally makes it necessary to increase relevant costs. This necessity has increased in recent years and has comprised the basic competition points among universities. It is thought that another factor providing the motivation to increase these costs is the shaping of the competition between universities around international relations and the international publications of the academicians.

Even though the most suitable statistical software was determined within the scope of data obtained by appealing to the opinions of five experts within the scope of the hierarchical model, the same study can be repeated with a different group of experts and updated data in addition to considering different criteria and sub-criteria. Weights of criteria weights can also be obtained by using the Analytic Network Process (ANP) allowing the analysis of the statistical software selection problem in a network structure, GRA application can be repeated and results can be compared.

5. REFERENCES

Altman, M. & Mcdonald, M. (2001). Choosing reliable statistical software, *PS: Political Sci. Politics*, 43 (3), 681–687.

Altuzarra, A., Maria, J., Jimenez, M. & Salvador, M. (2007). A Bayesian priorization procedure for AHP-group decision making, *European Journal of Operational Research*, 182 (1), 367-382.

Ayağ, Z. ve Özdemir, R.G. (2007). An intelligent approach to ERP software selection through fuzzy ANP, *International Journal of Production Research,* 45 (10), 2169-2194.

Başlıgil, H. (2005). Bulanık AHP ile yazılım seçimi, *Sigma Mühendislik ve Fen Bilimleri Dergisi*, 3, 27–33.

Chang, C.W., Wu, C.R., Lin, C.T. & Chen, H.C. (2007). An application of AHP and sensitivity analysis for selecting the best slicing machines, *Computers and Industrial Engineering*, 52 (3), 296–307.

Chou, T.Y., Lin, W.T., Lin, C.Y., Chou, W.C. & Huang, P.H. (2004). Application of the PROMETHEE technique to determine depression outlet location and flow direction in DEM, *Journal of Hydrology*, 287, 49–61.

Çörekçioğlu, M. ve Güngör, A. (2005). ERP yazılımı seçiminde Analitik Hiyerarşi Sürecinin kullanımı, 4. Üretim Araştırmaları Sempozyumu, 8–10 Ekim, Konya, 2-4.

Deng, J. (1989). Introduction to Grey System Theory, *The Journal of Grey System*, 1 (1), 1-24.

Dielman, T. (2002). The numerical accuracy of statistical calculations in Excel 2000 and Minitab version, *Journal of Business and Management*, 8 (2), 199–207.

Forman, E. & Peniwati, K. (1998). Aggregating individual judgments and priorities with the Analytic Hierarchy Process, *European Journal of Operational Research*, 108 (1), 165–169.

Girginer, N. ve Kaygısız, Z. (2009). İstatistiksel yazılım seçiminde Analitik Hiyerarşi Süreci ve 0-1 Hedef Programlama yöntemlerinin birlikte kullanımı, *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 10 (1), 211-233.

Hsia, K.H. & Wu, J.H. (1997). A study on the data preprocessing in Grey Relational Analysis, *Journal of Chinese Grey System*, 1, 47–53.

Jalal, N. & Ray, J.P. (1999). Software selection for simulation in manufacturing: A review, *Simulation Practice and Theory*, 7 (1), 1-14.

Javalgi, R.G, Armacost, R.L. & Hosseini, J.C. (1989). Using the Analytical Hierarchy Process for bank management: analysis of customer bank selection decisions, *Journal of Business Research*, 19 (1), 33-49.

Keeling, K.B. & Pavur, R.J. (2007). A comparative study of the reliability of nine statistical software packages, *Computational Statistics & Data Analysis*, 51 (8), 3811-3831.

Kitchen, A. M., Dranchenberg, R. & Symnazik, J. (2003). Assessing the reliability of web-based statistical software, *Computational Statistics*, 18 (1), 107–122.

Koçak, A. (2003). Yazılım seçiminde Analitik Hiyerarşi Yöntemi yaklaşımı ve bir uygulama, *Ege Akademik Bakış Dergisi*, 3 (1), 67–77.

Kusters, U., Mccullough, B.D. & Bell, M. (2006). Forecasting software: Past, present and future, *International Journal of Forecasting*, 22, 599–615.

Lai, S.V, Trueblood, P.R. & Wong, K. B. (1999). Software selection: A case study of the AHP to the selection of a multimedia authoring system, *Information and Management*, 36, 221-232.

Lai, S.V, Wong, K. B. & Cheung, W. (2002). Group decision making in a multiple criteria environment: A case using the AHP in software selection, *European Journal of Operational Research*, 137 (1), 134-144.

Lee, M., Pham, H. & Zhang, X. (1999). A methodology for priority setting with application to software development process, *European Journal of Operational Research*, 118 (2), 375-389.

Lee, H. S., Shen, P. D. and Chih, W. L. (2004), A fuzzy multiple criteria decision making model for software selection, 2004 IEEE International Conference (1098-7584), 25-29 July, 3, 1709-1713.

Lipovetsky, S. (1996). The Synthetic Hierarchy Method: An optimizing approach to obtaining priorities in the AHP, *European Journal of Operational Research*, 93 (3), 550-564.

Macharis C., Springael J., De Brucker K. & Verbeke A. (2004). PROMETHEE and AHP: The design of operational synergies in multicriteria analysis. Strengthening PROMETHEE with ideas of AHP, *European Journal of Operational Research*, 153 (2), 307–317.

Mulabeke, J.A.W. & Zheng, L. (2006). Analytic network process for software selection in product development: A case study, *Journal of Engineering Technology Management*, 23, 337-352.

Perera, H.S.C. & Costa, W.K.R. (2008). Analytic Hierarchy Process for selection of ERP software for manufacturing companies, *The Journal of Business Perspective*, 12 (4), 1-11.

Saaty, T.L. (1980). The Analytic Hierarchy Process, McGraw Hill, New York.

Saaty, T.L. (1990). Eigenvector and logarithmic least squares, *European Journal of Operational Research*, 48, 156–160.

Saaty, T.L. (1994). Fundamentals of decision making and priority theory with the AHP. RWS Publications, Pittsburgh, PA, U.S.A.

Saaty, T.L. (2000). Fundamentals of the Analytic Hierarchy Process, RWS Publications, Pittsburgh, PA.

Santhanam, R. & Kyparisis, G.J. (1996). A decision model for interdependent information system project selection, *European Journal of Operational Research*, 89 (2), 380-399.

Tung, S.L. & Tang, S.L. (1998). A comparison of the Saaty's AHP and modified AHP for Right and left eigenvector inconsistency, *European Journal of Operational Research*, 106 (1), 123–128.

Wu, W.Y., Shih, H.A. & Chan, H.C. (2008). The Analytic Network Process for partner selection criteria in strategic alliances, *Expert Systems with Applications*, 36 (3), 4646–4653.

Wu, C.R., Lin, C.T. & Tsai, P.H. (2010). Evaluating business performance of wealth management banks, *European Journal of Operational Research,* 207 (2), 971-979.

Wei, C.C., Chien, C.F. & Wang, M.J. (2005). An AHP-based approach to ERP system selection, *International Journal of Production Economics*, 96, 47-62.

Yazgan, H.R., Boran, S. ve Göztepe, K. (2009). An ERP software selection process with using artificial neural network based on Analytic Network Process approach, *Expert Systems with Applications,* 36 (5), 9214–9222

Uzun Özet

Özellikle yirminci yüzyılın son çeyreğinden itibaren bir yandan istatistik yöntem ve tekniklerinin her alanda yaygınlaşması, diğer yandan bu alanda çalışma yapanların önceliklerinin değişmesi gibi nedenlerle günümüzde bilimsel araştırmalarda, elde edilen verilerin çözümlenmesinde ve bulguların yorumlanmasında çeşitli istatistiksel yöntemler yoğun olarak kullanılmaktadır. Ancak yığın veri içeren, karmaşık ve çok değişkenli problemlerin çözümlenmesinde gereken istatistiksel analizlerin elle çözümünün çok zaman alması ve hesaplama işlemlerinde hata yapma olasılığının artması, bu yönde geliştirilen yazılımların kullanımını da zorunlu hale getirmiştir. Verilerin analiz ve değerlendirilmesi aşamasında hangi yazılımların çalışma verilerine en uygun olduğunun saptanması, sonuçların güvenirliliği açısından önem taşımaktadır. Dolayısıyla istatistiksel işlemlerin yapılmasından ziyade uygun istatistiksel tekniğin ve bu teknik için kullanılacak olan istatistiksel yazılımın seçimi ön plana çıkmaktadır. Diğer taraftan üniversitelerin eğitim-öğretim dışında gerçekleştirdikleri bilimsel araştırmaların kalitesinde, amaca uygun, doğru yazılımın seçilmesi etkilidir. Her fakültenin sınırlı bir bütçeye sahip olması ve akademisyenler tarafından talep edilen yazılımların tamamının karşılanamaması da araştırmacıları çoğu zaman bir yazılım seçimi problemi ile karşı karşıya bırakmaktadır. Bu çalışmada, istatistiksel yazılımların hem istatistik eğitimindeki hem de bilimsel araştırmalardaki gerekliliği ve fakültelerin sınırlı bütçelerinden dolayı istatistiksel yazılım seçimi problemi ele alınmıştır.

Çalışmanın amacı, istatistiksel yazılımların seçiminde dikkate alınan ve birlikte değerlendirilmesi gereken kriterleri ve önceliklerini Analitik Hiyerarşi Süreci (AHS) ile belirleyerek, Gri İlişkisel Analiz (GİA) ile en uygun istatistiksel yazılım seçimini gerçekleştirmektir. İstatistiksel yazılım seçiminde etkili olduğu düşünülen kriter ve alt kriterlerin AHP ile elde edilen ağırlıkları, GİA'de veri olarak kullanılmıştır. Böylelikle GİA ile her bir alternatif yazılım için bir önem derecesi ortaya konularak, en yüksek önem derecesine sahip istatistiksel yazılımın en uygun istatistiksel yazılım olarak belirlenmesi amaçlanmıştır.

İstatistiksel yazılımlar konusunda gerek kullanıcı gerekse programcı olarak uzman olan akademisyenler ve programcılardan oluşan beş kişilik bir ekip oluşturulmuştur. Söz konusu uzmanlar bir bilgisayar programcısı, bir bilgisayar öğretmeni, istatistik eğitimi veren bir akademisyen ve diğer sayısal alanlarda (pazarlama, tıp, vb.) istatistiksel yazılımları kullanan iki akademisyenden oluşmaktadır. Ekip ile gerçekleştirilen görüşmeler ve literatürde yazılım seçimiyle ilgili yapılan çalışmalar doğrultusunda, istatistiksel yazılım seçimi probleminde göz önüne alınması gereken kriterler ve alt kriterler belirlenmiş ve problemin yaygın kullanımı olan dört yazılım seçeneği (SPSS, Statistica, Minitab ve Statgraph) için hiyerarşik yapısı oluşturulmuştur. Hiyerarşik yapı oluşturulurken, her birisi alt kriterler içeren beş temel kriter (finansal, teknik, kullanım, analiz ve satıcı özellikleri) dikkate alınmıştır. Problemin hiyerarşik yapısının oluşturulmasından sonra istatistiksel yazılımlar hakkında bilgisi olan ve gerek akademik gerek öğretim amaçlı kullanıcı durumundaki kişilerin, ikili karşılaştırmalardaki yargılarının farklılaşacağı düşüncesiyle katılımcıların yargılarının geometrik ortalamaları alınarak grup karar vermeli AHP uygulanmıştır. AHP ile elde edilen kriter ve alt kriter ağırlıkları, GİA'de veri olarak kullanılarak her bir alternatif yazılım için bir önem derecesi ortaya konulmuş ve en yüksek önem derecesine sahip istatistiksel yazılım en uygun istatistiksel yazılım olarak belirlenmiştir.

Bulgulara göre istatistiksel yazılım seçiminde en fazla önem verilen ana kriterin analiz özellikleri (%49,8) olduğu; bu kriteri kullanım özellikleri (%23,7) ve teknik özellikler (%16,6) ana kriterlerinin izlediği görülmektedir. İstatistiksel yazılım seçiminde finansal özellikler (%5,1) ve satıcı firmaya ait özellikler (%4,9) önem derecesi bakımından son sıralarda yer almaktadır. Finansal özellikler ana kriteri göz önüne alındığında, alt kriterler arasında göreli önem değeri en yüksek olan alt kriter güncelleme fiyatı (0,018) olurken, bunu edinme ve işletim maliyeti (0,011) izlemektedir. Alt yapı (0,010) ise son sırada yer almaktadır. Teknik özellikler ana kriteri dikkate alındığında, göreli önem değeri en yüksek olan alt kriterin veri dosyası desteği (0,066) olduğu görülmektedir. Bu alt kriteri 0,046 önem düzeyi ile çoklu ortam desteği izlerken, yazılımın grafik desteğinin önem derecesi 0,028'dir. En az önem derecesine sahip olan alt kriter ise 0,025 önem derecesiyle yazılımın ara yüzüdür. Kullanım özellikleri ana kriteri için; en yüksek önem derecesine sahip olan yazılımın kullanımının kolaylığıdır (0,112). En düşük önem düzeyine sahip olan alt kriter ise 0,038 önem düzeyi ile erişim kolaylığıdır. Analiz özellikleri ana kriterinde ise en yüksek öneme sahip olan alt kriter neden-sonuç analizleri (0,140) olarak ortaya çıkmıştır. Bunu sırasıyla karşılaştırma analizleri, ilişkisel analizler ve sınıflayıcı/ayırma analizleri izlemektedir. Tanımlayıcı istatistikler ise 0,032'lik önem derecesiyle en son sırada ortaya çıkmıştır. Son ana kriter olan satıcı özelliklerinde ise eğitim desteği (0,014) en önemli alt kriter, kullanıcı gereksinimleri (0,010) ise en az önemli alt kriter olarak belirlenmiştir. Diğer taraftan, GRA sonucunda elde edilen ana kriter ağırlıklarına bakıldığında da AHP sonuçlarına benzer şekilde en fazla önem verilen ilk iki ana kriterin sırasıyla analiz özellikleri ve kullanım özellikleri olduğu dikkat çekmektedir. Ayrıca GRA bulgularına göre, en uygun istatistiksel yazılım %93.8 ile SPSS olmuştur. SPSS'i %63.8 ile Statistica, %46.8 ile Minitab ve %34.46 ile Statgraph izlemiştir. Akademisyenler istatistiksel yazılımlara bağlı bulundukları üniversite bütçesi içerisinden yapılan alımlarla sahip olduklarından, finansal özellikler kriteri bu çalışmada hem AHP hem de GRA sonuçlarına göre önemli bir kriter olarak ortaya çıkmamıştır. Ayrıca satıcı özellikleri kriteri açısından düşünüldüğünde ise, yazılımlar satın alınıp kullanılmaya başlandıktan sonra gerekli olan teknik ve eğitim desteği gibi bir takım gereksinimleri, kullanıcılar genellikle kendileri karşıladıklarından yine bu kriter de istatistiksel yazılım seçiminde önemli bir kriter olarak belirlenmemiştir.

İstatistiksel yazılımların bu çalışmada da sözü edilen önemine bağlı olarak araştırmacıların ve akademisyenlerin kişisel olarak edinemeyecekleri ya da zorlukla edinebilecekleri yazılımları sağlama konusunda üniversiteler içindeki beklenti, bu yöndeki harcamaların artırılmasını zorunlu kılmaktadır. Bu zorunluluk son yıllarda daha yüksek düzeyde kendini göstermektedir ve bir ölçüde üniversiteler arasındaki rekabetin temel noktalarından birini oluşturmaktadır. Üniversiteler arasındaki rekabetin ve nitelik ölçütünün uluslararası ilişkiler ve akademisyenlerin uluslararası yayınları etrafında şekillenmesinin de bu harcamaların artışında motive edici bir unsur olacağı düşünülmektedir. Her ne kadar kurulan hiyerarşik model kapsamında beş uzman görüşüne başvurularak elde edilen veriler kapsamında en uygun istatistiksel yazılım belirlenmiş olsa da, ele alınabilecek farklı kriter ve alt kriterlerin yanı sıra sadece akademisyenlerden oluşan bir uzman grup ile de güncellenen verilerle çalışma tekrarlanabilir. Aynı zamanda kriterler, alt kriterler ve alternatifler arasındaki bağımlılığın da dikkate alınacağı ve böylelikle istatistiksel yazılım seçimi probleminin bir ağ yapısında incelenmesini sağlayacak Analitik Network Süreci kullanılarak kriter ağırlıkları tekrar elde edilip, GRA uygulaması yinelenebilir ve sonuçlar karşılaştırılabilir.

Citation Information

Kaygısız-Ertuğ, Z., & Girginer, N. (2014). A multi criteria approach for statistical software selection in education. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi [Hacettepe University Journal of Education],* 29(2), 129-143.