CONTENTS
İçindekiler

REVIEWERS of the 76th ISSUE
76. Sayı Hakemleri

Relationships between Political Behaviors of School Principals and Perceived Coworkers' Social Loafing among Teachers*

Beyza HIMMETOGLU[1] , Damla AYDUG[2] , Cetin TERZI[3]

A R T I C L E I N F O	A B S T R A C T
	Purpose: This study was carried out to examine the relationships between teachers' opinions on the political behaviors used by school principals and their perceptions of their coworkers' social loafing practices. **Research Methods:** A correlational survey model was used in the study. The population of the study consists of 1948 primary school teachers. The sample of the study consists of 652 primary school teachers. **Findings:** Teachers' opinions on school principles' use of idealized effect, praise and legitimization behaviors is reported to be at the level of "I agree", but for exchange and pressure behaviors their opinions were at the level of "I neither agree nor disagree".

It is found that teachers' perceptions coworker social loafing are low. Legitimization and pressure behaviors respectively are statistically significant predictors of teachers' perceptions of their coworkers' social loafing. **Implications for Research and Practice**: Political behaviors used by school principals can be examined through qualitative or mixed methods to obtain detailed information. In order to reduce the social loafing in a given educational organization to minimum, studies can be conducted to determine which precautions should be taken and how motivation levels of teachers can be increased in collaborative work In order to decrease negative results arising from social loafing in educational organizations, it can be suggested that school principals should use legitimization behavior more and pressure behavior less.

---

[1] Corresponding Author: Anadolu University, TURKEY, e-mail: beyzahimmetoglu@anadolu.edu.tr, ORCID: https://orcid.org/0000-0003-0757-232X
[2] Anadolu University, TURKEY, e-mail: damlaaydug@anadolu.edu.tr, ORCID: https://orcid.org/0000-0001-8348-5098
[3] Anadolu University, TURKEY, e-mail: cterzi@anadolu.edu.tr, ORCID: https://orcid.org/0000-0002-1413-8233

Introduction

The main qualification that differentiates an organization from any community is the gathering of individuals to achieve a common purpose (Parsons, 1956, p. 64). Members of the organization coordinate their activity towards achieving the common, organizational goal (Bittner, 1965, p.175). Increasing the goal achievement level of an organization is the main priority to improve operational effectiveness. However, ignoring individual goals is an important factor that decreases the level of organizational goals achievement (Argrys, 1964; Ouchi, 1980). In this regard, aligning individual goals with organizational goals is a crucial issue for organizational effectiveness. However, this harmony is not always easily achieved (Hall, Schneider & Nygren, 1970). Individual goals may be prioritized ahead of organizational goals in 21st-century organizations in which individual tendencies towards competitive, global and postmodern perspectives become more prevalent. This situation often transforms organizations into political arenas in which members display behaviors serving their individual goals. Another factor transforming organizations into political arenas is the power concept in organizations. Power is defined as the ability to direct and affect other individuals' behaviors by controlling the resources others need (Beycioglu & Sincar, 2013, p. 247). In this sense, it can be said that people with more power want to direct and affect others' behaviors. Thus, it can be concluded that individual goals and the factors related to power in organizations increase the politicization levels of organizations by increasing the use of political behaviors.

Political behaviors are defined as behaviors that serve to achieve individual goals rather than organizational goals and operate outside of the formal task and role definition of members (Farrel & Peterson, 1982, p.405; Islamoglu & Boru, 2007, p. 136). Based on these definitions, an important aspect of political behavior is that the motives behind the behavior are shaped around individual drives and needs rather than organizational ones. However, if the individual goals coincide with the organizational goals, political behaviors can serve to achieve organizational goals. There are three main factors triggering political behaviors in organizations: ambiguity, lack of trust and resource shortage (Poon, 2003, p. 138). Ambiguity prevents individuals from internalizing organizational goals and roles. It also causes information pollution, which directs people to protect their own interests. Lack of trust in organizations make members think that others behave on behalf of their own interests, which results in showing a tendency to display political behaviors. Resource shortages cause competitions and struggles among individuals and groups for the same resources at the same time. This can lead people to behave outside their formal task definitions in order to get these resources.

Since the various reasons causing political behaviors and the awareness of organizational goals among individuals vary nature, political behaviors vary, too. In this study, the political behaviors used by managers are examined by considering the classification of political behaviors according to horizontal and vertical behaviors. The examples of political behaviors used by managers in organizations can be rational persuasion, consultation, inspiration, legitimization, coalition, pressure and gaining the support of superiors. Political behaviors are generally seen as organizational

behaviors that produce negative behaviors (Yukl, Falbe, & Youn, 1993, p. 6). However, the consequences of political behaviors differ as positive and negative according to both how they are perceived (Parker, Dipboye, & Jackson, 1995, p. 892) and whether or not they serve organizational goals (Landells & Albrecht, 2013, p. 363). Although there are perspectives that see political behaviors as positive or negative according to their consequences, it is generally accepted that political behaviors are an indispensable part of all organizations (Curtis, 2003, p.293). On the other hand, examining the consequences of political behaviors can be essential for taking precautions against negative ones.

There are theoretical and applied studies in the literature that demonstrate which organizational and personal behaviors and variables are affected by organizational politics. These studies show that organizational politics generally have negative results such as increasing stress, burnout, and tension (Ferris & Kacmar, 1992); turn-over intensions and psychological withdrawal (Cropanzano, Howes, Grandey, & Toth, 1997; Randall, Cropanzano, Bormann, & Birjulin, 1999) and decreasing commitment, trust, and job satisfaction (Kumar & Ghaidally, 1989) of members. Thus, based on the related literature, it is possible to say that political behaviors, which decrease trust, cause the necessary projects and tasks to go wrong by affecting members' commitment and job satisfaction negatively and making them reluctant to do their jobs by affecting their motivation negatively. In this regard, it is possible for political behaviors to make organizational members avoid contributing to group tasks.

One of the concepts that describe how working in a group can decrease the productivity of individual is social loafing. Social loafing implies that individuals put forth less effort than expected in group projects (Ilgin, 2013, p.239). Karau and Williams (1995, p. 135) define social loafing as a decrease in the motivations and efforts of individuals when they work together with others as compared to working alone. The first person who mentioned this decrease of individual efforts in collective environments is Ringelmann. This decrease in individuals' efforts when they are in a group was initially called the "Ringelmann Effect," but then Latane, Williams and Harkins (1979, p. 823) labelled that situation "social loafing" and defined it as a decrease in individual effort due to the social pressure of others. Because people generally work as hard as the rate of the pressure they feel, as the group grows larger, the efforts and contributions of individuals decrease.

Social loafing may be the result of various sources. Some reasons for social loafing are mentioned in the literature as a lack of motivation, a lack of control, ambiguity about individuals' contributions to the group and ambiguity of purpose (Roy, Gauvin & Limayem, 1996). However, the reasons behind social loafing can be grouped into two categories: individual-based and group-based reasons. Individual-based reasons include the interrelatedness of tasks, task visibility, distributive justice and procedural justice perceptions, and group-based reasons include group size, group commitment and the perceived social loafing of coworkers (Liden, Wayne, Jaworski, & Bennett, 2004, pp.287-291). Perceived social loafing of coworkers indicates the perception levels of group members about one or more coworkers displaying social loafing (Comer, 1995, pp.647-677). Whatever the actual behavior, the emphasis is on perception since

group members can witness what their coworkers do and do not do better and closer than managers (Liden et al., 2004, p.291). Thus, it is believed that defining the perceptions of employees related to their coworkers' social loafing can be an effective way of evaluating social loafing behaviors at the organizational level.

Social loafing is evaluated as a negative situation especially in effectiveness- and productivity-based organizations. Thus, organizations need to be aware of the precautions against social loafing in addition to the reasons behind social loafing (Dogan, Bozkurt, & Demir, 2012, p. 56). Reasons for social loafing in organizations can be summarized as organizational and administrative implementations causing negative attitudes and behaviors among members. In particular behaviors that decrease trust in coworkers' and administrators' justice perception and motivation can cause social loafing. In this sense, it is possible to assume that political behaviors causing negative attitudes and behaviors of members is one of the reasons behind social loafing.

The number of studies in the literature examining the relationships between political tactics used by principals in educational organizations and the perceived social loafing of coworkers among teachers is limited. By the means of this study it is believed that needed information related to both the effects of school principals' political behaviors on the faculty and the possible results of these behaviors will be provided. Also, examining the possible causes of social behavior among teachers will contribute to increasing the cooperation level in educational organizations. Thus, this study can contribute to filling the aforementioned gap in the literature. In this regard, it will be possible to make suggestions for the necessary precautions to prevent or minimize social loafing behavior by identifying the kinds of political behaviors that increase or decrease social loafing among teachers via this study. This study was carried out to examine the relationships between teachers' opinions on organizational political behaviors used by school principals and their perceptions of their coworkers' social loafing. To this end, following research questions were posed for the study:

1. What is the degree of school principals' using political influence behaviors?

2. What is the level of perceived coworker social loafing among primary school teachers?

3. Are the political influence behaviors used by school principals statistically significant predictors of perceived coworker social loafing among teachers?

Method

Research Design

A correlational survey model was used in this study. Correlational models are often used to determine whether two or more concepts vary consistently and the consequent degree of relationship between these concepts (Cresswell, 2012, p. 338). It was aimed to define the current relationships between the examined concepts in this study as well.

Research Sample

The population of the study consisted of 1948 teachers who work at primary schools affiliated with the Directorates of National Education in Tepebasi and Odunpazari during the 2015–2016 academic year. There are forty-three public primary schools affiliated with the Odunpazari Directorates of National Education, and there are forty-four public primary schools affiliated with the Tepebasi Directorates of National Education. The cluster sampling method was used to determine the sample group. The cluster sampling method employs random selection to constitute the sample instead of selecting individuals one by one. It was important for this sampling method that all groups should have individuals with similar features (Ozen & Gul, 2007, p. 406–407). Since the population is very large, and all the members' features are similar, cluster sampling was preferred in this study. In this study, the Tepebasi and Odunpazari districts were accepted as clusters, so, first data collection tools were sent to twenty-five of forty-four primary schools affiliated with the Directorate of National Education of Tepebasi and to twenty of forty-three primary school affiliated with the Directorate of National Education of Odunpazari. Data could be collected from thirty-eight of fifty primary schools. After answered data collection tools were examined, it was detected that seventeen primary school teachers had answered the questions inadequately, and these tools were excluded. Finally, the sample group constituted of 652 primary school teachers which is 33.47% percent of these two districts. The necessary sample size for the population consisting of 1948 teachers was calculated as 321 teachers with a 95% confidence level. In this regard, the sample size of the study consisting of 652 primary school teachers was evaluated as adequate.

Research Instruments and Procedures

Data of the study were collected via "Political Influence Tactics Scale," "Perceived Coworker Social Loafing Scale," and "Personal Information Form," the last of which including questions to determine participants' demographic features such as gender, age and educational degree. To determine the political behaviors of school principals the "Political Behaviors Scale," which was developed by Yukl and Falbe (1990) and then revised by Berson and Sosik (2007), was used. The original form of this scale consists of forty-four items to evaluate eleven influence tactics. However, the revised form, including eight dimensions and thirty-two items, was employed in this study since this form excludes influence tactics used upwardly. This five-point Likert scale was translated into Turkish by Mehtap (2011), who also perform the reliability and validity analysis of the scale. However, the reliability and validity of the scale were re-examined for this study. In to determine the validity of the scale, a factor analysis was performed with data collected from the 652 primary school teachers constituting the sample of the study. Before exploratory factor analysis, KMO and Barttlett sphericity tests were performed (KMO = 948, Barttlett sphericity = 18101.644, df = 465, p = 0.000). Then the exploratory factor analysis was performed using the Principal Components-Varimax Rotated Solution technique. The exploratory factor analysis was performed twice. After the first rotation, items fifteen and sixteen were excluded because the differences of their factor loads for different factors was under 0.10. After the second

rotation, a scale with five dimensions (idealized effect, exchange, praise, legitimization and pressure) and thirty items was created. The total variance of all five dimensions in the scale was calculated to be 72.75 %. The first sub-dimension of the scale includes fourteen items, and each of the other four sub-dimensions include four items. The factor loads of the items vary between 0.666 and 0.900. The Cronbach-Alpha reliability co-efficient values were calculated as 0.960, 0.948, 0.930, 0.859 and 0.832 respectively.

The "Perceived Coworker Social Loafing Scale" was developed by Liden et al. (2004) who took the 10-item social loafing scale developed by George (1992) into consideration. Ulke (2006) transformed the scale developed by Liden et al. (2004) into a five-point Likert scale with thirteen items. The reliability and validity of the scale were re-examined for this study. In order to determine the suitability of the scale for exploratory factor analysis, KMO and Barttlett sphericity tests were performed (KMO = 0.915, Barttlett sphericity = 4217.790, df = 78, p = 0.000). Then exploratory factor analysis was performed using the Principal Components-Varimax Rotated Solution technique. Exploratory factor analysis was performed twice. After the first rotation, items two and thirteen were excluded since they didn't work as reverse items. Item five also was excluded from the scale since its factor load was under 0.50. After the second rotation and the exclusion of these three items, a scale with one dimension and ten items was formed. The total variance of the scale was calculated to be 56.87%. The factor loads of the items vary between 0.850 and 0.562. The Cronbach-Alpha reliability co-efficient value of the scale was calculated to be 0.912.

Data Analysis

Before analyzing the data, it was necessary to determine whether or not the data distribution was normal. In order to see the distribution of data, the skewness and kurtosis coefficients and stem-leaf and histogram graphics were examined. Since it was determined that the data distribution was normal, parametric tests were used. In order to evaluate the opinions of participants on political behaviors used by school principals and the perceived levels of coworkers' social loafing, the arithmetic mean, standard deviation, and minimum and maximum values were calculated. A stepwise regression analysis was used to determine whether political behaviors used by school principals are statistically significant predictors of perceptions of social loafing among teachers. Before applying regression analysis, its assumptions were tested. First, in order to perform this analysis appropriately, extreme values were controlled. In this regard, Mahalanobis distance values were considered. The data of five participants were identified as extreme values higher than the recommended Mahalanobis distance value (13.82 in Pallant, 2011, p. 159), and they were removed before the regression analysis. Another assumption of multiple regression analysis is singularity. The relationships between the variables of this study were examined, and it was found that there was not an instance of singularity since correlation values between the variables were under 0.70. Also, in order to examine collinearity between predictor variables according to the last model of stepwise regression analysis, VIF and tolerance values were examined (VIF = 1.04 and tolerance value = 0.960). Last, the assumption indicating no OTO correlation between independent variables was tested by calculating the Durbin-Watson value. According to results, the Durbin-Watson value

is 1.890, which is between the recommended values of one and three (Field, 2009, p. 224).

Results

First, an analysis was performed to determine the political behaviors used by school principals according to the opinions of primary school teachers and to determine the teachers; perceptions of their coworkers' social loafing. Based on these analyses, descriptive statistics values related to the "Political Influence Tactics Scale" and the "Perceived Coworker Social Loafing Scale" of primary school teachers are displayed in Table 1.

Table 1

Descriptive Statistical Values of Political Influence Behaviors Used by School Principals and Primary School Teachers' Perceptions of their Coworkers' Social Loafing Levels

Variables	n	No of Items	Min.	Max.	$\overline{X}$	sd	$\overline{X}$ /number of items	sd/number of items
Idealized effect	652	14	28.00	70.00	56.17	7.74	4.01	0.55
Exchange	652	4	4.00	20.00	12.27	4.42	3.07	1.11
Praise	652	4	4.00	20.00	14.74	3.23	3.69	0.81
Legitimization	652	4	4.00	20.00	15.06	2.73	3.77	0.68
Pressure	652	4	4.00	20.00	10.84	3.76	2.71	0.94
Perceived Coworker Social Loafing	652	10	10.00	50.00	24.86	7.50	2.49	0.75

As seen in Table 1, the values of the items in the sub-dimensions of "Political Influence Tactics Scale" vary. In order to make a comparison between the sub-dimensions of the scale, first the means of each sub-dimension were divided into item numbers and transformed means varying between one and five. According to the findings, the opinions of teachers for the idealized effect, legitimization, and praise sub-dimensions are at the level of "Agree" ($\overline{X}$ = 4.01, $\overline{X}$ = 3.77, $\overline{X}$ = 3.69 respectively), and the exchange and pressure sub-dimensions are at the level of "Neither Agree Nor Disagree" ($\overline{X}$ = 3.07 and 2.71 respectively). When the findings related to perceived coworkers' social loafing levels were examined, it was seen that teachers' opinions are at the level of "Disagree" ($\overline{X}$ = 2.49). Based on this finding, it can be said that primary school teachers' perceptions of coworker social loafing is low. To test whether political behaviors used by school principals are significant predictors of teachers' perceived social loafing levels, a stepwise regression analysis was used. Before the stepwise regression analysis was performed, the Pearson Product-Moment Correlation technique was applied to examine the relationships between variables.

Correlation coefficients indicating the relationships between variables and descriptive statistics are displayed at Table 2.

Table 2

Results of Pearson Product-Moment Correlation Analysis Applied to Examine Relationships between Political Influence Behaviors and Social Loafing Levels

Variables	1	2	3	4	5	6
1- Idealized effect	-					
2- Exchange	0.172**	-				
3- Praise	0.611**	0.483*	-			
4- Legitimization	0.450**	0.377**	0.475**	-		
5- Pressure	-0.134**	0.386**	0.030	0.211**	-	
6- Social Loafing	-0.189**	0.029	-0.114**	-0.189**	0.144**	-

*$p<.05$, **$p<.01$

As seen Table 2, there are positive and statistically significant relationships between the perceived social loafing levels of teachers and political behavior ($r = 0.144$, $p < 0.01$) behavior. There are negative and statistically significant relationships between the perceived social loafing levels of teachers and the idealized effect ($r = -0.189$, $p < 0.01$), praise ($r = -0.114$, $p < 0.01$) and legitimization ($r = -0.189$, $p < 0.01$) behaviors. On the other hand, there are not any statistically significant relationships between the perceived social loafing levels of teachers and exchange behavior. Therefore, it is concluded that including exchange behavior in regression analysis as a predictor variable is not necessary. The results of a stepwise regression analysis, which were calculated for the remaining four predictor variables, indicate that two (legitimization and pressure) of these four variables are significant predictors of the perceived social loafing levels of teachers.

Table 3

Results of Stepwise Regression Analysis Related to Political Influence Behaviors of School Principals as Predictors of Perceived Social Loafing Levels

Model	Variables	B	SH_B	β	t	p	R^2	ΔR^2	F	p
1	Legitimization	-.0534	0.109	-0.189	-4.894	0.000	0.036	0.036	23.96	0.000
2	Legitimization	-0.638	0.110	-0.226	-5.826	0.000	0.068	0.033	23.65	0.000
	Pressure	0.369	0.078	0.184	4.748	0.000				

According to the results of the stepwise regression analysis, which are displayed in Table 3, the first model involves legitimization behavior. According to first model, the legitimization behavior of school principals explains 3.6% of the perceived social loafing levels of teachers ($R^2 = 0.036$). In the model, it is seen that legitimization

behavior predicts perceived social loafing levels of teachers negatively. In this regard, when the usage of legitimization behavior increases, the perceived social loafing levels of teachers decrease, or when the use of legitimization behavior decreases, the social loafing level of teachers increases. In the second step, pressure behavior takes part as a predictor in the model. According to second model, pressure behavior explains 3.3 % of the social loafing behaviors of teachers ($\Delta R^2 = 0.033$). It is also seen that pressure behavior predicts perceived social loafing behavior positively. According to the last model of stepwise regression analysis, legitimization and pressure behaviors together explain 6.8% of the perceived social loafing levels of teachers ($R^2 = 0.068$). According to the last model, legitimization ($\beta = 0.23$, $p < 0.05$) and pressure ($\beta = 0.18$, $p < 0.05$) behaviors are statistically significant predictors of perceived social loafing. While legitimization, one of the variables in the last model, has a more powerful prediction level, pressure is still a positive predictor of social loafing. When school principals use increased pressure behavior to make teachers obey their requests, the perceived social loafing levels of teachers also increases, and when principals' use of pressure behavior decreases, perceived social loafing levels of teacher decrease as well.

Discussion, Conclusion and Recommendations

The findings indicate that, according to teachers' opinions, school principals use idealized effect and legitimization behaviors from political behaviors most. The findings of many researches in the literature also supported the idea that school principals prefer moderate political behaviors, which refer to idealized effect behavior like rational persuasion and consultation (Barbuto, Fritz, & Marx, 2002; Berson & Sosik, 2007; Ozgul-Katlav, 2016; Knippenberg & Steensma, 2003). Although the findings of these studies vary, it was seen that they generally supported findings such as: the most often used political behaviors are rational persuasion and consultation behaviors and moderate political behaviors are preferred over rigid ones. One of the political behaviors, called "idealized effect" in this study, includes such political behaviors as rational persuasion, inspiring, consultation and collaboration mentioned in the literature. Idealized effect behavior comprises behaviors like persuading others by explaining the reason for the one's request and offering reasonable statements and factual information, generating eagerness by appealing to others' desires and values, supporting them in their tasks and providing adequate resources (Kipnis & Schmidt, 1980, p. 447; Yukl et al., 1993, p. 7). According to teachers' opinions in this study, it can be inferred that school principals usually prefer the idealized effect behavior of political behavior. Thus, it can be said that there are not rigid hierarchical relationships between teachers and school principals. Additionally, maintaining informal relationships in schools instead of using rigid enforcements can be evaluated as the necessity of a loose structure within educational systems because of teachers' classroom autonomy.

The results of this study indicated that the secondarily preferred political behavior used by school principals is legitimization. Legitimization behavior is related to basing

the school principals' requests on their authority and emphasizing how their requests are reinforced by the schools' rules and procedures (Yukl et al., 1993, p. 7). Legitimization behavior can be described as compatibility with rules. Even though some studies (Daglı, 2015; Ozgul-Katlav, 2016) support that legitimatization behavior is one of the frequently preferred political behaviors, some other studies show that it is one of the underutilized political behaviors (Barbuto & Moss, 2006; Falbe & Yukl; 1992; Yukl et al., 1993). This situation indicates that the results related to legitimization behavior vary from study to study. However, these differences between the results of these studies can be derived from the differences between the institutions where the research was conducted. The reasons explaining why the school principals frequently prefer legitimization behavior can usually be traced to basing the school principals' authority on rules and procedures and the general perceptions of schools as bureaucratic institutions.

The results of the study show that the least preferred political behavior used by school principals is pressure, according to the opinions of primary school teachers. Pressure behavior is associated with behaviors including threats, insistent demands and enforcement (Barbuto & Moss, 2006, p. 32). Similar to this study, several studies found that pressure behavior is one of the least preferred political behaviors (Barbuto et al., 2002; Daglı, 2015; Falbe & Yukl, 1992; Ozgul-Katlav, 2016). However, there are some studies that asserted that pressure is one of the most preferred political behaviors or moderately preferred political behaviors (Berson & Sosik, 2007; Duyar, Aydin, & Pehlivan, 2009; Kuru-Cetin, 2013; Yukl & Falbe, 1993). The reasons for the differences between the findings of these studies can be attributed to the institutions where the research was conducted, the data collection tools used in the studies and the population and sample from whom the data of the studies was collected. However, the school principals' not using pressure behavior frequently in the schools, which are organizations having warm and intimate relationships between individuals and being human-oriented organizations, can be evaluated as a desirable situation because it is known that using pressure-based power and enforcements like punishment may alienate members of the organizations (Hoy & Miskel, 2010, p. 212). As a result, it can be said that according to the teachers' opinions in this study, school principals prefer moderate political behaviors instead of punishment and enforcement.

It teachers' perceptions about the level of coworker social loafing were also investigated in this study. Findings indicated that teachers' perceptions about the level of coworker social loafing is low. Similarly, it was found that participant's perceptions about social loafing is also low in Dogan, Bozkurt and Demir's (2012) study conducted on workers in the service and manufacturing sectors. However, it was determined that participant's perceptions about social loafing is on a moderate level in many studies in the local literature (Ilgin, 2010; Kesen, 2015; Tolukan, Bayrak, & Karacan-Dogan, 2017). In the foreign literature, there are some findings showing that perceptions of people about the level of social loafing are low (Chang, 2008; Murphy et al., 2003; Mulvey & Klein, 1998). However, even low-level perception of social loafing should be seen as a problem that needs to be addressed (Piezon & Ferree, 2008). When teachers' perceptions about the level of coworker social loafing were analyzed, it can

be inferred that social loafing should be seen as a problem for educational organizations and precautions should be taken to change the teachers' perceptions about social loafing.

The reasons behind teachers' perceptions about coworker social loafing being low may be attributed to the unique characteristics of a given culture. There are lots of studies that mention that social loafing is influenced by cultural differences. For example, in Karau and Williams' (1995, p. 139) meta-analysis study, people in Eastern cultures, which are more collectivistic, give more importance and attribute a meaning to collective works. Therefore, social loafing is less common in these cultures. Earley (1989, p. 577) as compared with the effects of individualism and collectivism on social loafing. Furthermore, it is observed that American managers who have individualism-based beliefs display social loafing behavior, but Chinese managers who have collectivism-based beliefs do not display social loafing behavior. Therefore, Turkish culture, which is collectivist rather than individualistic, can be the reason for the finding that the perceptions about coworker social loafing of the teachers who participated in this study is low.

Finally, the predictive role of school principals' political behavior in teachers' perceptions about coworkers' social loafing was examined. It was determined that legitimization and pressure behaviors are statistically significant predictors of teachers' perceptions about coworkers' social loafing. According to these results, principals' legitimization behavior negatively predicted teachers' perceptions about social loafing. On the other hand, principals' pressure behavior positively predicted teachers' perceptions about social loafing. In other words, an increase in legitimization behavior leads to a decrease in social loafing behavior, and an increase in pressure behavior leads to an increase in social loafing behavior. When the fact that pressure is one of the rigid political behaviors and is associated with negative organizational outputs is considered, this result can be viewed as an expected result. However, like pressure behavior, legitimization behavior is also one of the rigid political behaviors (Berson & Sosik, 2007, p. 679). The findings of this study showed that legitimization behavior leads to the opposite consequences of pressure behavior. The important reason of this finding is that because the Turkish education system employs a bureaucratic structure, legal, authority-based power is commonly used by school principals. Additionally, legitimization behavior focuses on applying rules for everyone equally, which is identified with a fair management approach.

Studies examining the organizational and behavioral results of the political behaviors used by managers are found in the literature. However, there have not been any studies examining the relationships between political behaviors and social loafing. For this reason, the findings of this study can be evaluated within the scope of the negative and positive results of pressure and legitimization behaviors. Studies in Turkey indicated that legitimization behavior is generally related to positive organizational outcomes. For instance, a study designed by Mehtap (2011) showed that legitimization behavior is positively correlated with organizational identification, affective commitment and organizational citizenship behavior. Another similar example was the study conducted by Daglı (2015). Daglı (2015) pointed out that the

influencing behavior named as impartiality is positively correlated with organizational citizenship behavior and mindfulness. Pressure behavior has similar results to those of this study, as well. Studies on pressure behavior indicated that pressure behavior is correlated with negative results like demoralizing, damaging relationships between teachers, seeing teachers avoid making an effort and/or expending energy (Blase, 1990), decreasing awareness (Daglı, 2015), and limiting performance (Higgins, Judge, & Ferris, 2003). From these negative results, especially avoiding making contributions, can be associated with social loafing behavior. Furthermore, decreasing performance can be specified as potential result of social loafing.

Based on the results of this study, in order to make a contribution to the theory and practice related to behavior politics and social loafing, school principals could be informed about the functional results of political behavior when they employ it as compatible with organizational goals. In addition, political behaviors used by school principals can be examined through qualitative or mixed methods to obtain detailed information. In order to reduce the social loafing in educational organizations to minimum, research can be conducted to determine which precautions should be taken and how motivation levels can be increased in collective works. Finally, to decrease negative results arising from social loafing in educational organization, principals should use legitimization behavior more and pressure behavior less.

References

Argyris, C. (1964). *Integrating the individual and the organization.* New York: John Wiley & Sons, Inc.

Barbuto Jr, J. E., Fritz, S. M., & Marx, D. (2002). A field examşnation of two measures of work motivation as predictors of leaders' influence tactics. *The Journal of Social Psychology*, 142(5), 601-616.

Barbuto Jr, J. E., & Moss, J. A. (2006). Dispositional effects in intra-organizational influence tactics: A meta-analytic review. *Journal of Leadership & Organizational Studies*, 12 (3), 30-48.

Berson, Y., & Sosik, J. J. (2007). The relationship between self-other rating agreement and influence tactics and organizational processes. *Group and Organization Management,* 32(6), 675-698.

Beycioglu, K., & Sincar, M. (2013). Örgütlerde güç ve politika [Power and politics at organizations]. S. Özdemir (Ed.), *Eğitim yönetiminde kuram ve uygulama* içinde (s. 245-261). Ankara: Pegem Akademi.

Bittner, E. (1965). The concept of organization. *Social Research*, 32(3), 239-255.

Blase, J. J. (1990). Some negative effects of principals' control-oriented and protective political behavior. *American Educational Research Journal,* 27(4), 727-753.

Chang, Y. (2008). *Social loafing and moral emotions: The janus-headed aspect of moral identity.* Yayınlanmamış Yüksek Lisans Tezi, University of Delaware, Delaware, USA.

Comer, D.R. (1995). A model of social loafing in real work groups. *Human Relations*, 48, 647-677.

Creswell, J.W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research.* (Fourth Edition). Boston, MA: Pearson.

Cropanzano, R., Howes, J.C., Grandey, A.A. & Toth, P. (1997). The relationship of organizational politics and support to work behaviors, attitudes, and stress. *Journal of Organizational Behavior*, 18(2), 159-180.

Curtis, S. (2003). Lies, damned lies and organisational politics. *Industrial and Commercial Training*, 35(7), 293-297.

Daglı, E. (2015). *İlköğretim okullarında müdürlerin kullandıkları etkileme taktiklerinin öğretmenlerin örgütsel vatandaşlık davranışları ve okul farkındalığı ile ilişkisi* [*The relationship between principals' influence tactics and teachers' organizational citizenship behaviors and school mindfulness in primary schools].* Yayınlanmamış Doktora Tezi, Gazi Üniversitesi, Ankara

Dogan, A., Bozkurt, S., & Demir, R. (2012). Sosyal kaytarma davranışı ile algılanan görev görünürlüğü arasındaki ilişkinin incelenmesine yönelik bir araştırma. [A research about investigating the relationship between social loafing behavior and the perception of task visibility] *Sosyal ve Ekonomik Araştırmalar Dergisi,* 24, 53-79.

Duyar, İ., Aydin, İ., & Pehlivan, Z. (2009). Analyzing principal influence tactics from a cross-cultural perspective: Do preferred influence tactics and targeted goals differ by national culture? *Educational Leadership: Global Contexts and International Comparisons,* 11, 191-220.

Earley, P. C. (1989). Social loafing and collectivism: A comparison of the United States and the people's Republic of China. *Administrative Science Quarterly*, 34, 565-581.

Falbe, M., & Yukl, G. (1992). Consequences for managers of using single influence tactics and combinations of tactics. *The Academy of Management Journal,* 35(3), 638-652.

Farrel, D., & Peterson, J. C. (1982). Patterns of political behvaior in organizations. *The Academy of Management Review,* 7(3), 403-412.

Ferris, G. R., & Kacmar, K. M. (1992). Perceptions of organizational politics. *Journal of Management,* 18(1), 93-116.

Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll).* (Third Edition). London: SAGE.

George, J. M. (1992). Extrinsic and instrinsic origins of perceived social loafing in organizations. *Academy of Management Journal,* 35, 191-202.

Hall, D.T., Schneider, B., & Nygren, H.T. (1970). Personal factors in organizational identification. *Administrative Science Quarterly,* 15(2), 176-190.

Higgins, C., A. Judge, T. A., & Ferris, G.R. (2003). Influence tactics and work outcomes: A meta-analysis. *Journal of Organizational Behavior,* 24, 89-106.

Hoy, W. K., & Miskel, C.G. (2010). *Eğitim yönetimi* (S Turan, çev.). (7. Baskı). Ankara: Nobel Yayıncılık.

Ilgin, B. (2010). *Örgütsel vatandaşlık davranışlarının oluşumunda ve sosyal kaytarma ile ilişkisinde, duygusal zeka ve lider üye etkileşiminin rolü* [*The roles of the emotional intelligence and leader-member exchange in the occurance of the organizational citizenship behaviors and its relationship with the social loafing].* Yayınlanmamış Doktora Tezi, Hacettepe Üniversitesi, Ankara.

Ilgin, B. (2013). Toplumsal bir hastalık: Sosyal kaytarma. [A social disease: Social loafing] *Uşak Üniversitesi Sosyal Bilimler Dergisi,* 6(3), 238-270.

Islamoglu, G., & Boru, D. (2007). Politik davranış boyutları: Bir ölçek geliştirme çalışması. [Scale development for the dimensions of political behavior] *Akdeniz İ.İ.B.F. Dergisi,* 14, 135-153.

Karau, S.J., & Williams, K.D. (1995). Social loafing: Research findings, implications and future directions. *Current Directions in Psychological Science,* 4(5), 134-140.

Kesen, M. (2015). Psikolojik güçlendirme çalışanların sosyal kaytarma davranışlarını azaltır mı? [Does psychological empowerment decrease social loafing behaviors of employees?] *Journal of Yasar University,* 10(38), 6478-6554.

Kipnis, D., & Schmidt, S. M. (1988). Upward influence styles: relationship with performance evalutaions, salary, and stress. *Administrative Science Quarterly,* 33(4), 528-542.

Kumar, P., & Ghadially, R. (1989). Organizational politics and its effects on members of organizations. *Human Relations,* 42(4), 305-314.

Kuru-Cetin, S. (2013). *Okul yöneticileri ve öğretmenlerin birbirlerini etkileme taktiklerinin örgütsel adalet ile ilişkisi* [*Relationship between school managers and teachers*

downward and upward influence tactics and organizational justice]. Yayınlanmamış Doktora Tezi, Ankara Üniversitesi, Ankara.

Landells, E., & Albrecht, S. L. (2013). Organizational political climate: Shared perceptions about the building and use of power bases. *Human Resource Management Review*, 23(4), 357-365.

Latane, B., Williams, K. & Harkins, S. (1979). Many hands make light the work: The causes and consequences of social loafing. *Journal of Personality and Social Psychology*, 37 (6), 822-832.

Liden, R.C., Wayne, S.J., Jaworski, R.A., & Bennett, N. (2004). Social loafing: A field investigation. *Journal of Management,* 30, 285-304.

Knippenberg, B., & Steensma, H. (2003). Future interactiın expectation and the use of soft and hard influence tactics. *Applied Psychology: An International Review*, 52(1), 55-67.

Mehtap, Ö. (2011). *Örgüt içi politik davranışların örgütsel vatandaşlık davranışı üzerindeki etkisi [The effect of political behaviors on the organizational citizenship bejavior]*. Yayınlanmamış Doktora Tezi, İstanbul Üniversitesi, İstanbul

Mulvey, P. W., & Klein, H. J. (1998). The impact of perceived loafing and collective efficacy on group goal processes and group performance. *Organizational Behavior and Human Decision Processes,* 74(1), 62-87.

Murphy, S. M., Wayne, S. J., Liden, R. C., & Erdogan, B. (2003). Understanding social loafing: The role of justice perceptions and exchange relationships. *Human Relations,* 56(1), 61-84.

Ouchi, W. G. (1980). Markets, bureaucracies, and clans. *Administrative Science Quarterly*, 25(1), 129-141.

Ozgul-Katlav, E. (2016). *Algılanan etkileme taktiklerinin işgören performansı üzerindeki etkisini belirlemede kontrol odağının düzenleyici rolü: Otel işletmelerinde bir uygulama [Moderating role of locus of control in determining the effect of employee performance on perceived influence].* Yayınlanmamış Doktora Tezi, Nevşehir Hacı Bektaş Veli Üniversitesi, Nevşehir.

Parker, C. P., Dipboye, R. L., & Jackson, S. L. (1995). Perceptions of organizational politics: an investigation of antecedents and consequences. *Journal of Management*, 21(5), 891-912.

Parsons, T. (1956). Suggestions for a Sociological Approach to the Theory of Organizations-I. *Administrative Science Quarterly,* 1(1), 63-85.

Piezon, S. L., & Ferree, W. D. (2008). Perceptions of social loafing in online learning groups: A study of public university and U.S. naval war college students. *The International Review of Research in Open and Distance Learning,* 9(2), 1-17.

Pallant, J. (2011). *SPSS survival manual: A step by step guide to data analysis using SPSS*. (4th Edition). Australia: Allen & Unwin.

Poon, M.L. (2003). Situational antecedents and outcomes of organizational politics perceptions. *Journal of Managerial Psychology*, 18(2), 138-155.

Randall, M. L., Cropanzano, R., Bormann, C. A., & Birjulin, A. (1999). Organizational politics and organizational support as predictors o work attitudes, job perfromance, and organizational citizenship behavior. *Journal of Organizational Behavior,* 20, 159-174.

Roy, M. C., Gauvin, S., & Limayem, M. (1996). Electronic group brainstorming: The role of feedback on productivity. *Small Group Research*, 27(2), 215-247.

Tolukan, E., Bayrak, M., & Karacan-Doğan, P. (2017). Antrenörlerin sosyal kaytarma algılarının incelenmesi. [The examination of the level of social loafing perceptions of coaches] *Gazi Üniversitesi Spor Bilimleri Dergisi,* 2(1), 1-13.

Ulke, H. E. (2006). *Investigating the role of personality and justice perceptions on social loafing.* Yayınlanmamış Yüksek Lisans Tezi, Orta Doğu Teknik Üniversitesi, Ankara.

Vigoda, E. (2002). Stress-related aftermaths to workplace politics: the relationships among politics, job distress, and aggressive behavior in organizations. *Journal of Organizational Behavior*, 23(5), 571-591.

Yukl, G., & Falbe, C. M. (1990). Influence tactics and objectives in upward, downward and lateral influence attempts. *Journal of Applied Psychology,* 75(2), 132-140.

Yukl, G., Falbe, C. M., & Youn, J. Y. (1993). Patterns of influence behavior for managers. *Group & Organization Management*, 18(1), 5-28.

Okul Yöneticilerinin Sergiledikleri Politik Davranışlar ile Öğretmenlerin Algıladıkları İş Arkadaşlarının Sosyal Kaytarma Düzeyleri Arasındaki İlişkiler

Atıf:

Özet

Problem Durumu: Bireyin, grup içinde olmasının, verimliliğini azalttığını vurgulayan kavramlardan biri sosyal kaytarmadır. Sosyal kaytarma, grup halinde yapılan çalışmalarda bireylerin beklenenin altında çaba gösterdiklerini ifade etmektedir. Sosyal kaytarmayı ortaya çıkarabilecek nedenleri, üyelerin örgüte karşı olumsuz tutum ve davranışlar geliştirmelerine neden olabilecek örgütsel ve yönetsel uygulamalar ile iş arkadaşları tarafından sergilenen olumsuz tutum ve davranışlar şeklinde özetlemek olanaklıdır. Özellikle üyelerin iş arkadaşlarına ve yöneticilerine duydukları güvenin, örgütsel adalet algılarının ve motivasyonlarının azalmasına neden olabilecek davranışların sosyal kaytarma davranışına yol açabileceği görülmektedir. Bu bağlamda örgüt üyelerini rahatsız eden ve örgüte yönelik olumsuz tutum ve davranışlar sergilemelerine neden olabilecek politik davranışların, sosyal kaytarmanın nedenlerinden biri olarak nitelendirilmesi olasıdır.

Alanyazında eğitim örgütlerinde yöneticilerin kullandıkları politik etkileme davranışları ile bu davranışların öğretmenlerin sosyal kaytarma düzeyleri ile arasındaki ilişkiyi belirlemeye yönelik çalışmaların sınırlı sayıda olduğu belirlenmiştir. Bu çalışma aracılığıyla yöneticilerin politik davranışlarının, hem okullarda üyelerin davranışlarını etkileme süreçlerine hem de bunların olası sonuçlarına ilişkin bilgi sağlanabileceği düşünülmektedir. Böylelikle alanyazındaki bu eksikliğin giderilmesi amaçlanmaktadır. Bununla birlikte okul yöneticilerinin kullandığı politik etkileme davranışlarını belirlemenin, bu davranışların örgütsel sonuçlar üzerindeki etkilerini belirlemeye temel oluşturarak, uygulamaya yönelik faydalar da sağlayabileceği düşünülmektedir. Bu bağlamda söz konusu çalışma aracılığıyla incelenen politik davranışlardan hangilerinin öğretmenlerde sosyal kaytarmayı arttırdığı hangilerinin azalttığı ortaya konularak, okul için olumsuz sonuçlar üretebilecek sosyal kaytarma davranışını yok edebilmek veya kısmen de olsa azaltabilmek için alınabilecek önlemlere yönelik öneriler geliştirilebilecektir.

Araştırmanın Amacı: Bu araştırmada, ilkokul öğretmenlerinin görüşlerine göre okul yöneticilerinin sergiledikleri politik davranışlar ile öğretmenlerin algıladıkları iş arkadaşlarının sosyal kaytarma düzeyleri arasındaki ilişkiler incelenmiştir. Ayrıca

okul yöneticilerinin politik davranışlarının öğretmenlerin iş arkadaşlarının sosyal kaytarma düzeylerinin bir yordayıcısı olup olmadığı da araştırılmıştır.

Araştırmanın Yöntemi: Okul yöneticilerinin sergiledikleri politik davranışların öğretmenlerin sosyal kaytarma düzeyini etkileyip etkilemediği, etkiliyorsa ne yönde etkilediği ilişkisel tarama modeliyle belirlenerek, var olan durum saptanmaya çalışılmıştır. Araştırmanın evrenini 2015-2016 eğitim-öğretim yılında Eskişehir Tepebaşı ve Odunpazarı İlçe Milli Eğitim Müdürlüğü'ne bağlı ilkokullarda görev yapan 1948 ilkokul öğretmeni oluşturmaktadır. Araştırmada küme örnekleme yönteminden yararlanılmıştır. Sonuç itibariyle araştırmanın çalışma grubu 652 öğretmenden oluşmuştur. Araştırmanın verileri "Politik Etkileme Davranışları Ölçeği", "İş Arkadaşlarının Algılanan Sosyal Kaytarma Düzeyleri Ölçeği" ve "Kişisel Bilgi Formu" aracılığı ile elde edilmiştir.

Okul yöneticilerinin politik davranışlarını belirlemede; Yukl ve Falbe (1990)'nin çalışmaların ardından Berson ve Sosik (2007) tarafından yeniden düzenlenen "Politik Etkileme Davranışları Ölçeği" kullanılmıştır. Özgün ölçekte 11 taktik toplam 44 madde ile ölçülmüştür. Ancak "aşağıdan yukarıya" uygulanan taktiklerin ölçek dışında bırakıldığı son halinin bu çalışma için daha uygun olduğu düşünülmüş ve 8 boyut ve 32 maddeden oluşan yeniden düzenlenmiş ölçek formunun kullanılması kararlaştırılmıştır. 5'li likert türünde olan ölçeğin Türkçeye uyarlama çalışmaları ile geçerlik ve güvenirlik analizleri Mehtap (2011) tarafından yapılmıştır. Bu çalışmada Mehtap (2011) tarafından Türkçe'ye uyarlanan 32 maddelik ölçek formunun geçerlik ve güvenirlik değerleri yeniden hesaplanmıştır. Ölçeğin geçerliğini belirlemek üzere, çalışmanın örneklemini oluşturan 652 öğretmenden elde edilen veriler üzerinde açımlayıcı faktör analizi gerçekleştirilmiştir. Sonuç olarak 5 boyut (ideal etki, karşılıklı değişim, övgü, meşrulaştırma ve baskı) ve 30 maddeden oluşan bir veri toplama aracı elde edilmiştir. Veri toplama aracını oluşturan 5 faktörün toplam varyansı açıklama oranı % 72.75'tir.

İş arkadaşlarının algılanan sosyal kaytarma ölçeği George (1992) tarafından geliştirilen 10 maddelik sosyal kaytarma ölçeğinin maddelerinden yararlanılarak Liden vd. (2004) tarafından geliştirilmiştir. Ülke (2006) ise çalışmasında iş arkadaşlarının algılanan sosyal kaytarmalarını belirlemek amacıyla Liden ve arkadaşları tarafından geliştirilen ölçeği 13 maddelik 5'li likert türünde bir ölçeğe dönüştürmüştür. Bu çalışmada ölçeğin geçerlik ve güvenirlik özellikleri, yeniden hesaplanmıştır. Ölçeğin yapı geçerliliğini test etmek için öncelikle örneklem grubundan elde edilen verilerle açımlayıcı faktör analizi yapılmıştır. Analiz sonucunda ölçeğin tek boyuttan ve 10 maddeden oluştuğu belirlenmiştir. Elde edilen ölçeğin toplam varyansı açıklama oranı % 56.87'dir.

Araştırmanın Bulguları: Araştırmadan elde edilen sonuçlara göre, öğretmenlerin okul yöneticilerinin kullandıkları politik davranışlardan ideal etki, meşrulaştırma, övgü boyutlarına "katılıyorum", karşılıklı değişim ve baskı boyutlarına "ne katılıyorum ne katılmıyorum" şeklinde görüş bildirdikleri belirlenmiştir. Öğretmenlerin iş arkadaşlarının sosyal kaytarma düzeylerine ilişkin görüşleri incelendiğinde, öğretmenlerin "katılmıyorum" düzeyinde görüş bildirdikleri ve dolayısıyla öğretmenlerin iş arkadaşlarının sosyal kaytarma düzeylerini düşük olarak

algıladıkları saptanmıştır. Öğretmenlerin sosyal kaytarma düzeylerinin okul yöneticilerinin politik davranışlarından baskı (r= .144, p< .01) ile olumlu yönde, ideal etki (r= -.189, p< .01), övgü (r= .-114, p< .01) ve meşrulaştırma davranışları (r= -.189, p< .01) ile olumsuz yönde ilişkili olduğu saptanmıştır. Diğer taraftan öğretmenlerin sosyal kaytarma düzeyleri ile karşılıklı değişim boyutu arasında istatistiksel olarak anlamlı bir ilişkiye rastlanmamıştır. Adımsal çoklu regresyon analizinin sonuçlarına göre meşrulaştırma ve baskı davranışları birlikte, öğretmenlerin sosyal kaytarma düzeylerinin %6.8'ini açıklamaktadır (R^2=.068). Son modele göre meşrulaştırma (β= .23, p<.05), ve baskı (β= .18, p<.05), davranışlarının sosyal kaytarmayı istatistiksel olarak anlamlı bir şekilde yordadığı görülmektedir.

Sonuç ve Öneriler: Araştırma sonucunda, öğretmenlerin sosyal kaytarma algısının okul yöneticilerinin kullandıkları politik etkileme davranışlarından önemli yordayıcılarının sırasıyla meşrulaştırma ve baskı davranışları olduğu ortaya çıkmıştır. Bu ilişki çerçevesinde; okul yöneticilerinin politik etkileme davranışlarından meşrulaştırma davranışını kullanma düzeyleri arttıkça, öğretmenlerin sosyal kaytarmaya ilişkin algı düzeylerinin azaldığı anlaşılmıştır. Diğer taraftan, okul yöneticilerinin politik etkileme davranışlarından baskı davranışını kullanma düzeyleri arttıkça, öğretmenlerin sosyal kaytarmaya ilişkin algı düzeylerinin de arttığı saptanmıştır. Baskı davranışının, genellikle olumsuz davranışsal ve örgütsel çıktılarla ilişkilendirilen katı bir etkileme davranışı olarak nitelendirildiği düşünüldüğünde, elde edilen sonucun beklendik bir sonuç olduğu söylenebilir. Eğitim örgütlerinde sosyal kaytarmanın neden olduğu olumsuz sonuçları en aza indirgeyebilmek adına okul yöneticilerinin politik davranışlardan meşrulaştırmaya daha fazla, baskı davranışına ise mümkün olduğunca az başvurmaları önerilebilir.

Anahtar Kelimeler: Politik davranış, sosyal kaytarma, okul müdürü, öğretmen.

An Investigation of the Factors That Affect High School Students' Attitudes Towards Social Media by CHAID Analysis[*]

Ismail KARAKAYA[1], Mehmet SATA[2], Ergun Cihat CORBACI[3], Bayram CETIN[4]

A R T I C L E I N F O	A B S T R A C T
	Purpose: The purpose of this study is to determine the level of attitude of high school students towards social media and the variables that have an effect on these attitudes. **Research Methods:** In this study, the relational screening model was the descriptive method that was used. The data were collected by the "Social Media Attitude Scale." In the data analysis, the two-step cluster analysis and CHAID analysis (Chi-square Automatic Interaction Detector) were

used. **Findings:** According to the results of the CHAID analysis, the most effective variable on high school students' attitudes towards social media was "students' dream job." The other variables that affected attitudes towards social media were as follows: student's grade level, having a computer, the purpose of computer use, the mother's educational level, the father's educational level and the father's job. In addition, it was found that students had a moderate attitude towards social media. The effect size and relational degree for each significant variable was calculated. **Implications for Research and Practice:** As a result of the analyses, it was found that the level of high school students' attitudes towards social media was a moderate/threshold level and that their dream jobs had an effect on these attitudes. In order to determine whether these variables had significance in practice, the effect sizes and relational degrees for each significant variable were calculated. According to the calculations, effect sizes were large and medium, and their relational degrees were also high and moderate.

---

[1] Faculty of Gazi Education, Gazi University, TURKEY, e-mail: ikarakaya2002@gmail.com, ORCID: 0000-0003-4308-6919.

[2] Corresponding Author: Faculty of Education, Agri Ibrahim Cecen University, TURKEY, e-mail: mehmetwsata@gmail.com, ORCID: 0000-0003-2683-4997.

[3] Faculty of Gazi Education, Gazi University, TURKEY, e-mail: e.cihat.corbaci@gmail.com, ORCID: 0000-0002-7874-956X.

[4] Ministry of National Education, General Directorate of Measurement, Evaluation and Examination Services, TURKEY, e-mail: bcetin27@gmail.com, ORCID: 0000-0001-5321-8028.

Introduction

Thanks to the development of communication technologies, the internet has become the most popular communication environment where our daily lives have been affected from a very early age (Altunay, 2010). Online content from every area--such as education, healthcare, military, entertainment and science (Kenanoglu and Kahyaoglu, 2012)—has reshaped social, economic, and cultural life as well as interpersonal relations (Ozmen, Akuzum, Sunkur and Baysal, 2012). Within this context, it is seen that the internet and its usage is an important factor.

The internet technology that developed at the end of 21st century was, in the beginning, a one-way communication tool that traditional communication methods were supposed to use; thereafter, the development called Web 2.0, which enabled a user to communicate with other users and to interact with the network he/she was in redirected the internet technology and has enabled large masses of people to adopt the internet's usage area (Buyuksener, 2009). New applications have arrived thanks to Web 2.0. One of these applications is defined using terms such as "social networks," "social communication networks," or "social networking sites" (Otrar and Argin, 2014). Social networks serve as a public area for the internet as well as an environment where personal communication takes place (Solmaz ve Gorkemli, 2012). Social media stands out as the most ideal environment in terms of being easily accessible and allowing persistent updates (Ozturk and Talas, 2015). Social media that has become widespread with Web 2.0, which turns the internet into a networking environment, provides a potential power in terms of maximizing communication and interaction.

The use of social networking sites not only in daily life but also in education has provided this kind of content to often be handled in literature (Karademir and Alper, 2012). To give an example from the literature, Aijan and Harsthone (2008) stated that the use of social networks would be beneficial in educational activities. In another research, it was expressed that educational systems should be updated by benefiting from young individuals' interests and tendencies towards communication technologies (Gulsecen, Gursul, Bayrakdar, Cilengir and Canim, 2010). The research involving students' habits surrounding using social networking sites (students aged 9-16) was investigated and performed in partnership with Middle East Technical University and Information Technologies and Communications Authority (2011). It was determined that 66% of the individuals in the research group used social networking sites at least one time a day and they spent at least 72 minutes a day on these social networking sites (ODTU and TIB, 2011). In view of this research, it is true that individuals spend a lot of time on the internet. Another result of this research is that daily lives and the time spent on study times of the individuals who used social networking sites was adversely affected.

The foundations of the social network applications on education were based on different learning approaches because of both their nature and their functions and the factors they had. Ferdig (2007) stated that social network applications were closely related to many pedagogic points in constructivism and asserted that they supported pedagogic approaches such as active learning, social learning, societies of practice and

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

23

learning societies. Selwyn (2007) and Albion (2008) asserted that social networks provided students an opportunity to join new networks in terms of cooperative learning and these networks were appropriate for informal learning in their nature.

In the research done by Togay, Akdur, Yetisken and Bilici (2013) where prominent examples of social media applications were evaluated, it was aimed to determine whether the use of social media in university education was effective and helpful. The researchers provided 60 students with educational support based on social media. As a result, it was determined that the support of educational processes via social media was effective in students' learning and it facilitated learning processes and developed educational processes.

The purpose of this study is to determine the variables that are effective in high school students' attitudes towards social media. When literature was reviewed, it was seen that teenagers used social media actively and perceived it as an indispensable part of their lives (Ari, Yilmaz and Bektes, 2016; Aydin, 2016; Sahin, Kaynakci and Aytop, 2016). Within this context, the attitudes of individuals towards social media are essential. Within the scope of this study, the variables that have not been observed before were taken into consideration in terms of the attitudes towards social media. In addition, this study will contribute to the literature of interest in terms of effect sizes of the variables that have an effect on attitude.

In this regard, it is important to determine the attitude level of high school students who spend a considerable amount of time on the internet involved with social media and the factors that affect these attitudes. In this sense, the following question will be answered:

1. What are the attitude levels of high school students towards social media and which variable(s) is/are effective in the attitude of high school students?

Method

Research Design

In this study, the relational screening model that is one of the descriptive methods was used because this study tries to determine the attitude level of high school students towards social media and the factors that affect these attitudes (Buyukozturk, Cakmak, Akgun, Karadeniz and Demirel, 2013). Relations identified with this model cannot be interpreted as a real cause-and-effect relationship.

Research Sample

The population of the study that aims to determine high school students' attitudes towards social media and the factors affecting these attitudes consists of students who studied at high schools bound to Batman provincial directorate for national education during the 2015-2016 educational years. The sample of the study consists of 640 female and 954 male students (1594 students in total) who were chosen by the maximum variation sampling method, one of the non-random sampling methods. Different types

of high schools were chosen not only to incorporate both successful and unsuccessful students, but also to increase the representation of the population. In addition, the incomes of the students in different high schools were chosen to maximize the variation. The reason the study sample was chosen from Batman province was accessibility to the sample. The demographic information for the sample is given at Tables 1 and 2.

Table 1

Frequencies and Percentages for Independent Variable

Variables	Variable Categories	Frequency	Percentage
Student's Gender	Male	954	59.8
	Female	640	40.2
	15	512	32.1
	16	148	9.3
Student's Age	17	382	24.0
	18	427	26.8
	19	125	7.8
	Teacher	202	12.7
	Doctor	472	29.6
	Security Staff	77	4.8
Dream Job	Engineer	294	18.5
	Lawyer	172	10.8
	Architect	74	4.6
	Other	303	19.0
Monthly Income	1500 TL and below	513	32.2
	1501 – 3000 TL	838	52.6
	3001 – 5000 TL	177	11.1
	5001 TL and above	66	4.1
	Low (two hours or less a day)	324	20.3
The Frequency of Mobile Phone Use	Moderate (2 – 5 hours a day)	229	14.4
	High (more than 5 hours a day)	1041	65.3
The State of Computer Use	Yes	389	24.4
	No	1205	75.6
	To play games	190	11.9
	To do homework	441	27.7
The Purpose of Computer Use	Social networking sites	470	29.5
	To study	235	14.7
	Others (film, music etc.)	258	16.2
	9th grade	544	34.1
Grade	10th grade	376	23.6
	11th grade	281	17.6
	12th grade	393	24.7
Total		1594	100

A simple look at Table 1 reveals that 40.2% of the students are female and 58.9% of them are male, that their ages ranges from 15 to 19, that they have different dream jobs, and that their income level is moderate. Table 1 also shows that most students have no

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

25

personal computer, they use computers for different purposes, and they are in different grades.

Table 2

Frequencies and Percentages for Independent Variable

Variables	Variable Categories	Frequency	Percentage
Mother's level of education	Illiterate	591	37.1
	Literate	166	10.4
	Primary School Graduate	549	34.4
	Secondary School Graduate	154	9.7
	High School Graduate	92	5.8
	University Graduate	36	2.3
	Post Graduate	6	0.4
Father's level of education	Illiterate	96	6.0
	Literate	96	6.0
	Primary School Graduate	658	41.3
	Secondary School Graduate	289	18.1
	High School Graduate	252	15.8
	University Graduate	186	11.7
	Post Graduate	17	1.1
Mother's job	House wife	1538	96.5
	Officer	8	0.5
	Teacher	7	0.4
	Nurse	14	0.9
	Retired	3	0.2
	Other	24	1.5
Father's Job	Farmer	433	27.2
	Officer	135	8.5
	Policeman	13	0.8
	Teacher	26	1.6
	Doctor	13	0.8
	Soldier	31	1.9
	Worker	272	17.1
	Retired	38	2.4
	Craftsman	221	13.9
	Driver	39	2.4
	Other	373	23.4
Total		1594	100

From Table 2, we see that the educational level of the participants' parents is not high and most of the participants' mothers are housewives and their fathers have different jobs.

Research Instruments and Procedures

"Social Media Attitude Scale (SMAS)," which was developed by Otrar and Argin (2013), was used to determine high school students' attitudes towards social media. Socio-demographic information on students was collected via a "Personal Information Form" organized by the researcher.

The psychometric properties of the scale are as follows. The scale consists of six negative items and 17 positive items (23 items in total) and it has four factors (sharing necessity, social competence, social isolation and relation with teachers). It is a self-report scale. Factors account for 52.65% of the variance. For internal consistency, the Cronbach Alpha coefficient was estimated as .85 for the whole scale. Test-retest correlation coefficient was found to be significant (r=.83; *p<.001*) (Otrar ve Argin, 2014). As a result, it is stated that the measurements obtained using the measurement tool are valid and reliable.

The reliability and validity analyses were repeated for the present sample. For validity, a confirmatory factor analysis (CFA) was used. The number of students who participated in the confirmatory factor analysis was 1594 in total. Before conducting CFA, in order to test the assumption of multivariate normal distribution assumption for the data multivariate skewness (Z_s) and kurtosis (Z_k), multivariate χ^2 and relative multivariate kurtosis (RMK) were estimated. In the present study, the assumption of the multivariate normal distribution was violated (Z_s= between -11.207 and 14.654, Z_k = between -10.577 and 0.625, χ^2 = between 51.443 and 215.139, RMK = 1.135). Because the assumption of the multivariate normal distribution was violated, the Satorra-Bentler χ^2 (S-B χ^2) value was estimated by Robust Maximum Likelihood (MLR) as the parameter estimation method (Brown, 2015). To handle missing data, an EM (Expectation Maximization) algorithm was used. In addition, the outliers were excluded from the data set. One of the assumptions of CFA for unbiased estimations is that indicators/variables do not have a multicollinearity problem. In order to check this assumption, the bivariate correlations between variables were examined and these correlations ranged from -.427 to .534. As these correlation values were not higher than .85, multicollinearity did not exist (Cokluk, Sekercioglu & Buyukozturk, 2014, p. 277). Another assumption for unbiased estimation of CFA models is to have adequate sample size. This measurement tool was applied to 1594 individuals; therefore, this study has an adequate sample size. After these assumptions were tested and met, CFA analysis was conducted.

In order to test model-data fit, the Goodness of Fit Index (GFI≥0.90), the Adjusted Goodness of Fit Index (AGFI≥0.90), and the Root Mean Square Error of Approximation (RMSEA≤0.08) were taken into consideration (Brown, 2015). According to the confirmatory factor analysis, GFI, AGFI, RMSEA were calculated as 0.77, 0.72, and 0.054, respectively. GFI and AGFI values were below the nominal criterion of 0.90 and the main reason for these low values was because of the violation of the multivariate normality assumption. Comparative Fit Index (CFI) and Non-Normed Fit Index (NNFI) values are taken into consideration when multivariate normality assumption is violated. As a result, NNFI was estimated as 0.97 and CFI was estimated as 0.97;

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

27

these values indicated model-data fit. When these analyses were taken into consideration, model-data fit was considered to be reasonable and a four-factor measurement model was validated.

Table 3

Model Fit Indices for the Model (Kline, 2011, p.208)

Fit Indices	Good Fit Range	Acceptable Fit Range	Estimates	Model Fit Value of This Study
χ^2/df	$0 \leq \chi^2/df < 2$	$2 \leq \chi^2/df \leq 5$	709.96/203=3.49	Acceptable Fit
RMSEA	$0 \leq RMSEA < 0.05$	$0.05 \leq RMSEA \leq 0.10$	0.05	Good fit
NNFI	$0.95 \leq NNFI \leq 1.00$	$0.90 \leq NNFI < 0.95$	0.97	Good fit
NFI	$0.95 \leq NFI \leq 1.00$	$0.90 \leq NFI < 0.95$	0.97	Good fit
CFI	$0.95 \leq CFI \leq 1.00$	$0.90 \leq CFI < 0.95$	0.97	Good fit
SRMR	$0 \leq SRMR < 0.05$	$0.05 \leq SRMR \leq 0.10$	0.08	Acceptable Fit

Cronbach α ve McDonald ω coefficients were calculated in order to prove evidence of reliability in terms of internal consistency for the estimates obtained from 23 items. While Cronbach α coefficient is calculated for a one-factor scale, Stratified Cronbach α coefficient is calculated for the scales that have more than one factor (Tan, 2009).

The Cronbach's alpha can be interpreted as a lower bound to real reliability because it will be equal to the real reliability of the test only if the items are parallel or tau-equivalent or essentially tau-equivalent (Traub, 1994). Therefore, McDonald developed an appropriate reliability coefficient for congeneric measures. This coefficient (McDonald ω) is also called construct reliability and obtained from CFA. McDonald ω coefficient is suggested instead of Cronbach alpha coefficient in order to have proof of the reliability of the measures where congeneric measures exist (Osburn, 2000). In this context, the McDonald ω coefficient is used in the study.

Table 4

Reliability Values Obtained from the Present Study and the Original Form for the Scale

Subscales	Stratified Cronbach α This study	Cronbach α This Study	Cronbach α Original Study	McDonald ω This Study	McDonald ω Original Study
Sharing Necessity	--	.710	.805	.800	--
Social Competence	--	.772	.814	.807	--
Social Isolation	--	.654	.792	.842	--
Relation with Teachers	--	.750	.814	.696	--
Whole Scale	.751	--	.852	--	--

When reliability values were investigated, it could be derived that acceptable results were obtained for both subscales and the whole scale. As seen in Table 4, McDonald ω coefficient is higher than Cronbach α coefficient because McDonald ω coefficient is more suitable for congeneric measurements. This is also valid for this

study. The reason for using Cronbach α coefficient is to compare the estimate obtained in this study with the original scale. As a result, it can be inferred from these statistics that adequate proofs for validity and reliability were obtained.

Data Analysis

CHAID analysis and two-step cluster analysis were used in data analysis. Two-step cluster analysis was conducted to convert continuous data collected from the scale into nominal data so as to determine the variables that affected high school students' attitudes towards social media. The reason CHAID analysis was used in data analysis is as follows (Kayri, 2007): a) CHAID analysis is effective in classifying categorical data, b) every variable in the data is categorical, and c) it is possible that data sets are not from the same population. Two-step cluster analysis divides data sets into two homogeneous subgroups. Dividing heterogeneous data sets into homogeneous subgroups or clusters ensures statistical studies have better results (Kayri, 2007).

What is intended here is not to fail homogeneity and normality assumptions that classical regression should meet. When analyzing data consisting of the large sample size, because of the violation of homogeneity assumption, an estimate obtained from a regression model may not reflect reality. Therefore, dividing data into subgroups and investigating possible classified subgroups can provide better results and the assumption of homogeneity is satisfied. In this respect, a regression equation obtained by CHAID analysis is exempt from familiar assumptions (normality, linearity, additivity and homogeneity) because the whole population can be divided into nodes via an iteration algorithm (Kayri and Boysan, 2007; Kim, 2009). CHAID analysis uses Chi-square ($\chi2$) statistics as the relationship and interactions between independent variables are handled. Chi-square statistics handles the dependency between variables (Kirk, 2008). For the independent variables that had a significant chi-square statistic, their relation degree with outcome variable and effect sizes were calculated. Contingency coefficient (C) is used in Chi-square statistics in order to estimate the relationship between variables. Contingency coefficient is interpreted as Pearson correlation coefficient and defined as follows:

$$C = \sqrt{\frac{\chi^2}{\chi^2+N}} \tag{1}$$

Effect size was defined by Cohen (1988). This statistic helps the researchers make a judgement (Kirk, 2008). Effect size is defined as follows:

$$d = \sqrt{\frac{\chi^2}{N}} \tag{2}$$

Cut-off points for a scale can be determined by two-step cluster analysis. In this context, the sample of the present study was divided into three clusters and two cut-off points were estimated. These cut-off points were 63.5 and 82.5, respectively. The point interval 0-63.5 represents a negative attitude towards social media, 63.5-82.5 represents a moderate attitude towards social media and 82.5-115 represents a positive attitude towards social media. Also, as it has been stated in literature (Pehlivan, 2006), CHAID analysis is highly effective in classifying categorical dependent variables; total

scale scores that are continuous in this study were transformed into a categorical variable by two step cluster analysis.

The reasons CHAID analysis has been the most preferred tree diagram in practice are as follows: it provides highly powerful estimations by benefiting from large samples, it handles missing data in the variables, and because it is exempt from assumptions that are determined in the structural form of the model, it is an alternative non-parametric tree diagram to binary and multinominal logistic models (Dogan and Ozdamar, 2003). In addition, CHAID analysis can better reveal the effects of missing data in the model than other decision tree methods (Tanhan and Kayri, 2012). SPSS 21.0 and LISREL 8.80 packaged programs were used for data analysis. The SPSS Package program was used for descriptive statistics, CHAID analysis, and two step cluster analysis and the LISREL package program was used for confirmatory factor analysis.

Results

The level of high school students' attitudes towards social media and the variables that were effective in these attitudes were determined by CHAID analysis. In the model, the outcome variable is the categories (positive, moderate/threshold, negative) converted and obtained from the scale scores, and the independent variables are student's age, the frequency of mobile phone use, student's gender, student's grade level, monthly income, student's dream job, state of computer use, purpose of computer use, parent's educational level and parent's job.

CHAID analysis results are given in Figure 1. The percentages of having negative attitudes, moderate attitudes and positive attitudes were found to be 23.1%, 49.9% and 27.0%, respectively ($\chi 2$ (df) = 82.331(6), *adjusted-p = 0.000*). According to the tree diagram, the attitude level of the students was moderate.

The most effective variable on the attitudes of high school students towards social media was student's dream job ($\chi 2$ (df) = 82.331(6), *adjusted-p = 0.000*). According to analysis, students whose dream jobs were doctor and engineer had a more positive attitude towards social media than other students whose dream jobs were different. Furthermore, the attitudes of students whose dream job was lawyer were less positive towards social media than other students.

Figure 1. Tree diagram for the variables that have an effect on the attitudes of high school students towards social media.

According to Figure 1, it was established that a student's future job was one of the independent variables and having a personal computer, student's grade level, and mother's educational level were related. It was ascertained that the variable of having a personal computer had an effect on attitudes of the students whose dream job was categorized as "other" ($\chi 2$ (df) = 25.224 (2), *adjusted-p = 0.000*). It was also established that student's grade level had an effect on the attitudes of students who wanted to be a doctor or an engineer in the future and the students in ninth and twelfth grade had a positive attitude towards social media that the students in the tenth and eleventh grade ($\chi 2$ (df) = 18.925 (2), *adjusted-p = 0.001*). Likewise, it was determined that student's grade had an effect on the attitudes of the students who wanted to be architects, teachers, soldiers, and policemen ($\chi 2$ (df) = 29.901 (4), *adjusted-p = 0.000*). The variable that affected the attitudes of the individuals who wanted to be a lawyer was the mother's educational level ($\chi 2$ (df) = 14.316 (2), *adjusted-p = 0.012*).

In classification models, classification and risk tables provide an evaluation of how well the model works. The estimate value in a risk table is used to determine whether the classification model is appropriate/valid. As a result of the CHAID analysis, the risk estimate is 0.162 and its standard error is 0.009. Consequently, the model classifies approximately 83.8% of the students correctly.

CHAID analysis uses F and chi square statistics if the dependent variable is continuous and categorical, respectively. Chi square statistics was used because the dependent variable was a categorical variable in this study. Effect size and relation

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

31

degree are calculated for practical significance where chi square statistics is used. These coefficients can help a researcher make a judgment on whether the results are practically significant (Kirk, 2008, s.474). Chi square statistics can be used to estimate effect size and relation degree between variables. Relation degree and effect size were calculated by Equations 1 and 2, respectively. The relation degrees and effect sizes for the significant variables were given in Table 5.

Table 5

The effect sizes of significant variables and relation degrees

Variables	Effect Size (d)	Relation Degree (C)	Interpretation Criteria
Student's dream job	0.22	0.23	Large
State of computer use	0.28	0.29	Medium
Student's grade	0.28	0.29	Medium
Mother's educational level	0.28	0.29	Medium
Purpose of computer use	0.31	0.33	Medium
Father's job	0.35	0.37	Medium
Father's educational level	0.30	0.32	Medium

According to Table 5, it was found that a student's dream job that was the most effective significant variable for the attitudes of students towards social media had a large effect size and high relation degree.

Discussion, Conclusion and Recommendations

The purpose of the study was to determine the variables that had an effect on students' attitudes towards social media. In accordance with this purpose, the "Social Media Attitude Scale (SMAS)" developed by Otrar and Argin (2013) was used.

According to the findings of the study, it was determined that high school students had a moderate attitude towards social media. In literature, it has been established that students have a positive attitude towards social media and they use them actively (Ari, Yilmaz and Bektes, 2016; Aydin, 2016; Sahin, Kaynakci and Aytop, 2016; Calisir, 2015; Otrar and Argin, 2014; Vural and Bat, 2010; Buyuksener, 2009). In this context, the findings obtained from the study are in accordance with the literature.

The independent variables that were effective in the high school students' attitudes towards social media were as follows: student's dream job, having a computer, student's grade level, mother's educational level, the purpose of computer use, father's job and father's educational level. It was determined that the effect sizes were large and medium and relation degrees were high and moderate for the significant variables. The increase in mother's educational level causes students' attitudes towards social media to decrease. Karan (2006) emphasized that parents should encourage their children to make real friends and raise awareness of unwanted conditions on virtual platform. According to Berson (2000), having a positive attitude

towards social media depends on the guidance of the parents. In the study in which teachers' opinions on internet addiction were considered, Ulusoy (2017) stated that parents complained and they had the opinion that the students' education was negatively affected because students spent most of their time on the internet. It was emphasized that the parents especially should be included in the development of the habit of internet use (Olagunju, 2009). In a compilation study, Balci and Tezel-Sahin (2016) expressed that social media tools could be an effective way for teachers and students' parents to communicate. In this context, it was established that the findings of the research were supported by literature.

In terms of the purpose of computer use, it was found that the students who used computers to study and to share film, music and video had a more positive attitude towards social media than the students who used computers to prepare homework and to join social networking sites. In the research done by Togay, Akdur, Yetisken and Bilici (2012), it was stated that most of the university students claimed they found social media tools necessary, useful and easy and they increased their academic achievement. Similar results were also found by Kalafat and Goktas (2012).

The present study has some limitations. First, it would not be true to generalize the findings of the research to every high school student in Turkey because the sample was obtained from a single province. Second, the outcome variable and attitude in the study was a continuous variable and it was converted into a discrete variable via two-step cluster analysis. Within this context, it should be taken into consideration in future research. Third, there were a limited number of variables handled in the study. Therefore, more variables could be added for the future research.

In summary, the order of the effect size from highest to lowest for the variables that were effective in students' attitudes towards social media were sorted as follows: student's dream job, having a computer, student's grade level, mother's educational level, the purpose of computer use, father's job and father's educational level. It was determined that the purpose of computer use was an effective variable on students' attitudes towards social media.

Taking into account the attitude levels of high school students, social media should be a part of education and should be turned into an opportunity. Teachers can share their materials via social media, can assign homework through social media, and control its usage in the classroom. In this study, it has been found that there is no difference between students' attitudes towards social media in terms of their gender and age. The main reason for this finding may result from the fact that the number of male and female students is not equal. For future research, it is advised to use a balanced sample size in terms of gender.

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

33

References

Ajjan, H., & Hartshorne, R. (2008). Investigating faculty decisions to adopt Web 2.0 technologies: Theory and empirical tests. *The internet and higher education*, 11(2), 71-80. https://doi.org/10.1016/j.iheduc.2008.05.002

Albion, P. R. (2008). Web 2.0 in teacher education: Two imperatives for action. *Computers in the Schools*, 25(3-4), 181-198. https://doi.org/10.1080/07380560802368173

Altunay, M.C. (2010). Twitter: Gundelik yasamin yeni rutini Pit Pit Net [The Everyday Life And Social Network Sites: Twitter or "Pit Pit Net"]. *Galatasaray Universitesi Iletisim Dergisi*, 12, 31-56.

Ari, E., Yilmaz, V., & Bektes, R. (2016). Universite ogrencilerinin sosyal ag kullanimina iliskin davranislarinin teknoloji kabul modeli ile arastirilmasi [Researching the behaviours of the university students regarding social networking by technology acceptance model]. *International Journal of Management Economics & Business,* 12(27), 67-82.

Aydin, I. E. (2016). Universite Ogrencilerinin sosyal medya kullanimlari uzerine bir arastirma: Anadolu Universitesi Ornegi [A Study on usage of social media among univerty students: Anadolu University Case]. *Selcuk Universitesi Sosyal Bilimler Enstitusu Dergisi*, (35), 373-386.

Balci, A., & Tezel-Sahin, F. (2016). Sosyal medyanin aile katiliminda kullanilabilirligi uzerine bir incelemen[A study on the usability of social media in family participation]. *Kastamonu Egitim Dergisi*, 24(5), 2309-2322.

Berson, M. J. (2000). The computer can't see you blush. *Kappa Delta Pi Record*, 36, 158-162. https://doi.org/10.1080/00228958.2000.10518777

Brown, T. A. (2015). *Confirmatory factor analysis for applied research.* New York: Guilford Press.

Buyukozturk, S., Cakmak, E.K., Akgun, O.E., Karadeniz, S., & Demirel, F. (2013). *Bilimsel arastirma yontemleri* [Scientific research methods]. Ankara: Pegem A Publishing.

Buyuksener, E. (2009, December). *Turkiye'de Sosyal Aglarin Yeri ve Sosyal Medyaya Bakis* [The Place of Social Networks in Turkey and a Look at Social Media]. Paper presented in XIV. Internet Conference in Turkey, İstanbul.

Calisir, G. (2015). Kisilerarasi iletisimde kullanilan bir arac olarak sosyal medya: Gumushane universitesi iletisim fakultesi ogrencilerine yonelik bir arastirma [Social media as a means used in interpersonal communication: a research oriented onto the students of Gumushane University faculty of communication]. *E-Journal of New World Sciences Academy, NWSA Humanities Sciences,* 10(3), 115-144.

Cesur, Z., Memis, S., Gokkaya, O., & Demir, B. (2016). Sosyal medya uzerinde gosterilen reklamlarin ahlak kavrami acisindan degerlendirilmesi: universite ogrencileri uzerine bir arastirma [Evaluating the advertisements shown through social media in terms of the concept of morality: a review on university students]. *TURAN Center for Strategic Researches*, 8(32), 274-287.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Lawrence Erlbaum Associates. https://doi.org/10.1002/bs.3830330104

Cokluk, O., Sekercioglu, G., & Büyüköztürk, Ş. (2014). *Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları*. Ankara: Pegem A Publishing.

Dogan, N., & Ozdamar, K. (2003). CHAID analizi ve aile planlamasi ile ilgili bir uygulama [CHAID analysis and an application related with family planning]. *T. Klin. Tip Bilimleri,* 23, 392-397.

Ferdig, R. (2007). Examining social software in teacher education. *Journal of Technology and Teacher Education*, 15, 5-10.

Gulsecen, S., Gursul, F., Bayraktar, B., Cilengir, S., & Canim, S. (2010, February). *New generation mobil learning tool: Podcast*. Paper presented at 13. Akademik Bilisim Konferansi, Mugla.

Kalafat, O., & Goktas, Y. (2012). *Usage of social networks in higher education: university of Gumushane, sample of facebook.* Proceedings of the 5th International Computer & Instructional Technologies Symposium, 948-953.

Karademir, T., & Alper, A. (2012). *Standarts of social networks being a learning environment.* Proceedings of the 5th International Computer & Instructional Technologies Symposium, 731-743.

Karan, G. (2006). *Anne babalar icin internet rehberi* [Internet guide for parents]. İstanbul: Epsilon Publishing.

Kayri, M. (2007). Arastirmalarda iki asamali kumeleme analizi ve bir uygulamasi [Two-Step Clustering Analysis in Researches: A Case Study]. *Eurasian Journal of Educational Research,* 28, 89-99.

Kayri, M., & Boysan, M. (2007). Arastirmalarda Chaid Analizinin kullanimi ve bas etme stratejileri ile ilgili bir uygulama [Using Chaid Analysis in Researches and an Application Pertaining to Coping Strategies]. *Ankara Universitesi Egitim Bilimleri Fakultesi Dergisi*, 40(2), 133-149.

Kim, M. (2009). Two-stage logistic regression model. *Expert Systems with Applications*, 36, 6727–6734. https://doi.org/10.1016/j.eswa.2008.08.063

Kirk, R. E. (2008). *Statistics an introduction*. Belmont: Thomson Higher Education.

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

35

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. New York: Guilford press.

Kenanoglu, R., & Kahyaoglu, M. (2012). *Relationship between the cognitive, emotional and social behavior of pre-school children and the internet use.* Proceedings of the 5th International Computer & Instructional Technologies Symposium, 649-654.

ODTU., & TIB., (2011). *Cocuklarin sosyal paylasim sitelerini kullanim aliskanliklari arastirmasi raporu.* Ankara: Ulastirma Bakanligi Internet Kurulu.

Olagunju, A. O. (2009). Harmonizing the interests of free speech, obscenity, and child pornography in cyberspace: the new roles of parents, technology, and legislation for Internet safety. *The Scientific World Journal*, 9, 1260-1272. https://doi.org/10.1100/tsw.2009.147

Oncel, M., & Tekin, A. (2016). Ortaogretim ogrencilerinin Facebook bagimlilik duzeyleri ve kullanim amaclarinin farkli degiskenler acisindan incelenmesi [Investigation of secondary school students' Facebook addiction levels and usage purposes in terms of different variables]. *Adiyaman Universitesi Egitim Bilimleri Dergisi*, 6(1), 179-197.

Osburn, H. G. (2000). Coefficient alpha and related internal consistency reliability coefficients. *Psychological methods*, 5(3), 343-355. https://doi.org/10.1037/1082-989X.5.3.343

Otrar, M. & Argin, F. S. (2014). Ogrencilerin sosyal medyaya iliskin tutumlarini belirlemeye yonelik bir olcek gelistirme calismasi [A scale development study to determine the attitude of students' towards social media]. *Egitim ve Ogretim Arastirmalari Dergisi*, 4(1), 391-403.

Otrar, M. & Argin, F.S. (2013, June). *A scale development study to determine the attitude of students' towards social media.* Paper presented in 7th International Computers and Instructional Technologies, Ataturk University, Erzurum.

Ozmen, F., Akuzum, C., Sunkur M., & Baysal, N. (2012). Sosyal ag sitelerinin egitsel ortamlardaki islevselligi [Functionality of Social Networks in Educational Settings]. *e-Journal of New World Sciences Academy,* 7(2), 496-506.

Ozturk, M. F., & Talas, M. (2015). Sosyal medya ve egitim etkilesimi [Interaction of social media and education]. *Journal of World of Turks*, 7(1), 101-120.

Pehlivan, G. (2006). *CHAID analizi ve bir uygulama.* Yuksek Lisans Tezi, Yildiz Teknik Universitesi Fen Bilimleri Enstitusu, İstanbul.

Selwyn, N. (2007). *Web 2.0 applications as alternative environments for informal learning - a critical review.*

Solmaz, B., & Gorkemli, H. N. (2012). Yeni bir iletisim alani olarak sosyal medya kullanimi ve Konya Kadin Dernekleri ornegi [Use of social media as a new

communication tool: the case of Konya Woman Associations]. *Selcuk Universitesi Sosyal Bilimler Enstitusu Dergisi*, 28(1), 183-189.

Sahin, A., Kaynakci, C., & Aytop, Y. (2016). Ziraat Fakultesi Ogrencilerinin Sosyal Medya Kullanim Aliskanliklarinin Belirlenmesi [Determination of Social Media Usage Habits of Students in The Agriculture Faculty]. *JOTAF/Tekirdag Ziraat Fakultesi Dergisi*, 13(2), 34-45.

Tan, S. (2009). Misuses of KR-20 and Cronbach's Alpha reliability coefficients. *Education and Science,* 34(152), 101-112.

Tanhan, F., & Kayri, M. (2012). Ogretmen adaylarinin universite ogretim elemanlarina yonelik algilarini etkileyen faktorlerin CHAID Analizi ile incelenmesi [An Examination of the Factors Affecting Prospective Teachers' Perceptions of Faculty Members Using Chaid Analysis]. *Kuram ve Uygulamada Egitim Bilimleri*, 12(2), 807-821.

Togay, A., Akdur, T. E., Yetisken, İ. C., & Bilici, A. (2013, January). *The Usage of Social Networks in Education Processes: A Vocational High School Experience.* Paper presented at Akademik Bilisim Konferansi, Akdeniz University, Antalya.

Traub, E. R. (1994). *Reliability for the social sciences: Theory and Applications. Measurement methods for the social sciences*. London: Sage Publications.

Ulusoy, K. (2017). Tarih ogretmenlerinin madde bagimliligi ve medya bagimliligi ile ilgili dusunceleri [Opinions of history teachers related to drug addiction and media addiction]. *Kastamonu Egitim Dergisi*, 25(1), 385-400.

Vural, B. A., & Bat, M. (2010). Yeni bir iletisim ortami olarak sosyal medya: Ege Universitesi iletisim fakultesine yonelik bir arastirma [Social media as a new communication environment: a research on Ege University Faculty of Communication]. *Journal of Yasar University*, 20(5), 3348-3382.

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

37

Lise Öğrencilerinin Sosyal Medyaya Yönelik Tutumlarını Etkileyen Faktörlerin CHAID Analizi ile İncelenmesi

Atıf:

Özet

Problem Durumu: İletişim teknolojilerinin gelişmesi ve özellikle internet, gündelik yaşamın akışında değişimin en fazla yaşandığı popüler iletişim ortamlarından biri olmuş ve çok erken yaşlarda hayatımızı etkilemeye başlamıştır. Eğitimden sağlığa, savunma alanından bilimsel çalışmaya ve eğlenceye kadar hemen her alanda yer alan internet sosyal, ekonomik ve kültürel yaşamı olduğu kadar kişilerarası etkileşimi de yeniden şekillendirmiştir. İnterneti paylaşım ortamına çeviren web 2.0 akımı ile giderek yaygınlaşan sosyal medya, iletişim ve etkileşimi üst düzeye çıkarması ile potansiyel bir güç meydana getirmektedir. Sosyal paylaşım sitelerinin sadece günlük yaşantıda değil aynı zamanda eğitim alanında kullanılması alanyazında sıkça ele alınan bir konu haline getirmiştir. Alanyazın incelendiğinde genç bireylerin sosyal medyayı etkin bir şekilde kullandığı ve günlük hayatın vazgeçilmez bir parçası olarak gördükleri belirtilmiştir. Bu bağlamda ortaöğretim düzeyindeki öğrencilerin sosyal medyaya yönelik tutumları ve bu tutum üzerinde etkili olan değişkenlerin belirlenmesi önemli görülmektedir. Bu çalışma kapsamında alanyazında dikkate alınmamış değişkenlerin sosyal medyaya yönelik tutum üzerinde etkisine bakılması önemli görülmüştür. Ayrıca tutum üzerinde etkili olan değişkenlerinin etki düzeylerinin dikkate alınması açısından ilgili çalışma alana katkı sağlamaktadır.

Araştırmanın Amacı: Bu çalışmanın amacı, lise öğrencilerinin sosyal medyaya yönelik tutumlarının hangi düzeyde olduğu ve bu tutum üzerinde etkili olan değişkenlerin neler olduğunu belirlemektir.

Araştırmanın Yöntemi: Lise öğrencilerinin sosyal medyaya ilişkin tutum düzeyleri ve bu tutumu etkileyen faktörlerin neler olduğunu ortaya koymaya yönelik olduğundan araştırmada betimsel yöntemlerden ilişkisel tarama deseni kullanılmıştır. Araştırmanın evrenini 2015-2016 eğitim öğretim yılında Batman ili milli eğitim müdürlüğüne bağlı liselerde okumakta olan öğrenciler, örneklemini ise seçkisiz olmayan örnekleme yöntemlerinden maksimum çeşitlilik örnekleme yöntemi ile seçilen 640 kız ve 954 erkek öğrenci olmak üzere toplam 1594 öğrenci oluşturmaktadır. Araştırmanın verisi "Sosyal Medya Tutum Ölçeği" kullanılarak toplanmıştır. Mevcut çalışma için ölçeğin güvenirlik ve geçerlik analizleri ilgili

örneklem için yeniden yapılmıştır. Ölçeğin geçerlik çalışması için doğrulayıcı faktör analizi (DFA) kullanılmıştır. Gerçekleştirilen analizler sonucunda, modelin veri ile kabul edilebilir bir uyum sağladığı ve ölçme modelinin ölçeğin dört faktörlü yapısına ilişkin geçerli ölçümler yaptığı söylenebilir. Ölçeğin 23 maddelik formundan elde edilen ölçümlerin iç tutarlılık anlamında güvenirlik düzeylerine ilişkin kanıtlar elde etmek üzere, ilgili ölçümler için McDonald ω, Cronbach α ve tabakalı Cronbach α değerleri hesaplanmıştır. Ölçümlere ilişkin hesaplanan güvenirlik katsayıları incelendiğinde, kabul edilebilir değerlerin elde edildiği söylenebilir. Bu çalışmada hem McDonald ω hem de Cronbach α katsayısının kullanılmasındaki temel neden ölçeğin orjinalinde elde edilen güvenirlik değerlerinin karşılaştırılması ve konjenerik ölçümlerde McDonald tarafından önerilen güvenirlik katsayısının daha tutarlı sonuçlar ürettiğinin gösterilmesi amaçlanmıştır. Veri analizinde, sırasıyla; ölçekten elde edilen toplam sürekli puanını kategorik puanlara dönüştürmede kullanılan iki aşamalı kümeleme analizi ve veri madenciliği yöntemlerinden biri olan CHAID analizi kullanılmıştır. Sürekli veri setinin kategorik veri formatına dönüştürülmesinin temel nedeni, CHAID analizinin kategorik veri setlerinde daha güçlü ve tutarlı sonuçlar üretme durumu dikkate alındığından bu yapılmıştır.

Araştırmanın Bulguları: Lise öğrencilerin sosyal medyaya yönelik tutumlarının düzeylerini ve bu tutumları üzerinde etkili olan değişkenleri belirlemek için CHAID analizi kullanılmıştır. Kurulan modelde, bağımlı değişken ölçekten elde edilmiş toplam puanın kategorik hali (olumlu-orta/eşik-olumsuz şekli) bağımsız değişkenler ise öğrencinin yaşı, cinsiyeti, aylık gelir durumları, telefon kullanma sıklığı, sınıf düzeyi, gelecekteki meslek hayali, bilgisayar kullanma durumu, bilgisayar kullanma amacı, ebeveynin öğrenim durumu ve ebeveynin meslek durumları dahil edilmiştir. CHAID analizi kullanılarak yapılan analizler neticesinde, lise öğrencilerinin sosyal medyaya yönelik tutumları üzerinde etki düzeyi en yüksek olan bağımsız değişkenin "öğrencilerin gelecekteki meslek hayalleri" olduğu tespit edilmiştir. Analiz çıktıları incelendiğinde, gelecekte doktor ve mühendis olmak isteyen bireylerin diğer meslek gruplarından biri olmak isteyenlere göre daha olumlu bir tutuma sahip oldukları belirlenmiştir. Sosyal medyaya yönelik tutum üzerinde etkili olan diğer bağımsız değişkenler ise sırasıyla şu şekildedir; öğrencinin sınıf düzeyi, öğrencinin bilgisayarının olup olmaması, bilgisayarı kullanım amacı, annesinin öğrenim düzeyi, babasının öğrenim düzeyi ve babasının mesleği olduğu tespit edilmiştir. Ayrıca öğrencilerin orta düzeyde bir tutuma sahip oldukları belirlenmiştir. Anlamlı bulunan bağımsız değişkenler için etki büyüklüğü ve ilişki derecesi hesaplanmıştır.

Araştırmanın Sonuç ve Önerileri: Yapılan analizler sonucunda elde edilen bulgular incelendiğinde, lise öğrencilerinin sosyal medyaya yönelik tutumlarının orta/eşik düzeyde olduğu ve gelecekte hayal ettikleri meslek dallarının bu tutum üzerinde etkili bir değişken olduğu görülmüştür. Ayrıca öğrencinin sınıf düzeyi, öğrencinin bilgisayarının olup olmaması, bilgisayarı kullanım amacı, annesinin öğrenim düzeyi, babasının öğrenim düzeyi ve babasının mesleği bağımsız değişkenlerinin de sosyal medyaya yönelik tutum üzerinde etkili olduğu görülmüştür. Sonuçların

Ismail KARAKAYA - Mehmet SATA - Ergun Cihat CORBACI - Bayram CETIN
Eurasian Journal of Educational Research 76 (2018) 21-40

39

pratik anlamlılığı için anlamlı bulunan bağımsız değişkenlere ilişkin etki büyüklükleri ve ilişki dereceleri hesaplanmış ve sonuçlar raporlanmıştır. Yapılan hesaplamalara göre etki büyüklükleri büyük ve orta etkiye sahip iken, ilişki dereceleri de yüksek ve orta düzeyde ilişkili olarak bulunmuştur. Sosyal medyanın lise öğrencileri üzerinde ciddi bir etkisinin olduğu göz önüne alındığında, ebeveynlerin, öğretmenlerin ve diğer eğitim paydaşlarının eğitim ortamında bunları etkili bir şekilde kullanmaları gerektiği düşünülmektedir. Ayrıca lise öğrencilerinin sosyal medyaya yönelik tutumlarının olumlu yönde gelişmesi için anne ve babalarının yeterli bilgi birikimine sahip olmaları gerektiği düşünülmektedir.

Anahtar Kelimeler: CHAID analizi, etki büyüklüğü, tutum, sosyal medya.

Eurasian Journal of Educational Research
www.ejer.com.tr

Investigation of School Effects on Student Achievement in Primary Education Using Value-Added Assessment[1]

Filiz KOC[1]

ARTICLE INFO	ABSTRACT

Purpose: The purpose of this study is to assess the contribution of primary schools within the metropolitan municipality of the province of Ankara, Turkey to the achievement outcomes of 7th grade students using the results of the end of year Level Determination Exam. **Research Methods:** Carried out using a casual comparative study framework, the population of this study consists of the primary schools in the province of Ankara. The sample includes the 24 primary schools whose students have attended to the same school in both the 2007-2008 and 2008-2009 educational years within the territorial jurisdiction of the metropolitan municipality of capital of Turkey. The value-added effects of the schools in the sample on the student growth are assessed using a simple fixed-effect model. Moreover, in order to determine whether or not there exists a statistically significant relationship between the rankings of the schools according to average student achievement levels and the rankings according to the value-added effects on student growth, Kendall tau rank correlation coefficients are calculated. **Findings**: The results of this study indicate that there are significant inconsistencies between the rankings of the schools according to their value-added effects on student improvement and the rankings according to the average student achievement, the latter being the method frequently used to assess the performance of the schools in Turkey. Moreover, the results demonstrate that the value-added effects of the schools on student improvement differ drastically from subject to subject. **Implications for Research and Practice:** It is expected that this research will lead to a more balanced evaluation of schools particularly given the likely emergence of more data over the years. In addition, this is the first value added assessment study carried out in Turkey. It points out that the way Turkish schools are assessed is problematic and suggests that value added methods should be considered in evaluating the effects of schools.

Introduction

In today's information society, organizations have constantly reinvent???
themselves because of the international competition brought by the globalizing world.
As a social institution, it is necessary for schools to adapt to changing social conditions
in order to increase the quality of education and to improve their students both
academically and sociall. For this reason, educating individuals who are able to use
knowledge, adapt to the times and develop themselves has been the main task of
schools. It is known that there are differences in the success levels of schools in terms
of fulfilling these aims. Identification of the causes of the resulting success and the
promotion of less successful schools to the level of successful schools are important to
increase the quality of education as a whole. The beginning of this process is the
determination of what makes a successful school.

Families want to enroll their children in schools that offer good education. But how
is a "good school" determined? Is there a right method of identification? For many the
method used is closely connected to student achievement, which in Turkey is
associated with the success of the nationwide exam. This may be the case where an
association and an appraisal method may not be fair and objective. This is because
student success relies not only on the contribution made by the school and the teacher,
but also depends on such factors as the socio-economic level of the family, the
academic background of the student, and the knowledge of the student. The
evaluation of the contributions of schools and teachers by scientific methods without
the influence of such factors is an important research topic in educational sciences.

It is arguable that what a teacher or school is actually responsible for is the change
in student attainment during their period within the school community. In this
context, the real responsibility of the school or teacher should be "how much has the
school/teacher contributed to student success" regardless of the student's
socioeconomic level, intellectual background. It is arguable that the main objective of
education is to determine the academic performance of the students and to maximize
their achievement (Evergreen Freedom Foundation [EFF], 2008).

Taking this aim into consideration, successful education maximizes the academic
performance of the students. When evaluating schools in this context, it would be more
appropriate to assess schools based on their contribution to academic success of
student (US Department of Education, 2009).

In the United States and England, the leading countries in the field of school and
teacher evaluation in educational sciences, Value Added Assessment (VAA)
approaches have been developed so that school and teacher influences can be assessed
scientifically and objectively without any other effects. In these countries, the success
of the students is followed by the year-end exams similar to Level Determination Exam
(LDE, SBS in Turkish) in Turkey; but school and teacher evaluations use VAA
approaches instead of average achievement scores. VAA is the general name of
complex statistical techniques aimed at predicting the individual causal effects of

teachers or schools on student achievement, using test results from more than one year (McCafrey and Hamilton, 2007).

In VAA methods, each student forms his own control group. The extracurricular factors that influence the success of the pupils (socioeconomic status -SES, the educational status of the family, etc.) do not significantly change over subsequent years. The scores that a student receives from two equivalent exams (in two successive years) are influenced by both school-related and non-school factors. However, non-school factors can be considered to be equally effective in both exams, as non-school factors such as SES, do not change significantly over a couple of years. In this case, the amount of change in student achievement (development) is the result of school-related factors. This approach constitutes the basic principle of the VAA. Students are followed through repeated exams over the years and the school is held responsible for the change in student's success.

The idea of evaluating teacher effectiveness based on the development of students was first proposed by Eric Hanushek in 1971 (Hanushek, 1971). Subsequent studies have shown that the difference of teacher effectiveness is statistically significant (Hanushek, 1972, Murnane, 1975). Following the 1980s, work on VAA was accelerated, particularly in the US and UK educational sciences literature (eg, Department of Education and Science, 1983, 1984, Gray and Jesson, 1987, Gray et al., 1984, 1986; Woodhouse and Goldstein, 1988; Willms, 1987). Studies in this field show that although socioeconomic status of pupils is a significant factor for student success, it does not influence the increase in the success level significantly (Sanders, 1997). The effects of the teacher and the school are much more dominant than the student related background factors including the socioeconomic status (Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997). Studies examining the effectiveness of schoosl and teachers using the VAA approach have continued to take place after 2000 (eg Kane and Staiger, 2008; Ishii and Rivkin, 2009; Hanushek and Rivkin, 2010). The most well-known example of CRB is the Tennessee Value Added Assessment System (TVAAS) (Sanders and Horn, 1994, 1998). The Dallas Independent School District has a system that examines teacher influences and uses their results in an official teacher assessment (Webster and Medro, 1997). It has become one of the most common methods used to assess the effectiveness of teachers and schools in the US, especially with the Race to The Top program, which has distributed upwards of $ 4.5 billion of federal aid to States who use VAA to assess teacher and school performances (US Department of Education, 2009).

According to McCaffrey et al. (2003), there are two reasons why more and more attention is paid to value added assessment studies. First, VAA distinguishes the impact of school and teacher, irrespective and independent of the other strong factors in student success. These non-school related factors include, but are not limited to, the socioeconomic status of the family, race, and family history. This differentiation is very important for schools and teachers to be assessed in a robust and fair way. The second reason is that VAA studies show that the teacher effect can change a lot from one teacher to another. If these differences can be embodied and some characteristics of the teachers can be connected within a cause-effect relationship, significant improvements

can be made in the education system. The success of schools in Turkey is generally determined by their performance in the exams they enter the transition from one school level to an upper level (Basaran and Cınkır, 2013). It is therefore difficult to say that school, teacher and student achievements have been correctly and fairly determined because there are different factors that influence the success of students.

In Turkey, discussions are continuing on both the high school and university entrance systems (SBS, OKS, TEOG, LGS for high school; OSS, OYS, LYS, YGS for university). Political parties offer promises in their education programs that they will either cancel these exams or apply alternative exams. It is arguable that what Turkey really needs is that as the alternative systems being developed, the assessment of teachers and schools should also be taken into account. It is envisaged that this study will contribute to the evaluation of school and teacher assessment in the search for alternative systems.

To date, in Turkey, schools are assessed with respect to the average of their students' successes in the nationwide exams. When evaluating schools in this context, it would be more appropriate to assess school students based on their contribution to academic development (US Department of Education, 2009). Many states in the United States have made it mandatory to use VAA in evaluating teachers and schools, as assessments using VAA are thought to be fairer. There is no study (in academia or practice) related with VAA in Turkey, except for the masters thesis of the author of this study. In this context, this is the first study carried out in Turkey for Turkish educational system. This study is intended to be a cornerstone for more VAA studies to carry out in Turkey.

Common Problems in Measurement and Evaluation

One of the most common problems that educators encounter and discuss in the education system is the evaluation of student achievement in the school and the objective assessment of schools by means of correct measurement.

The score that the pupil gets from a test is not simply due to the academic accumulation he earns in his current institution. A student's success in a test may be based on two different accumulations: 1) the academic background of the student, and 2) the academic background of the school in which the student is enrolled during the test. Moreover, the accumulation of the student's previous academic life may have determined the school in which he is registered. In Turkey in order for a student to qualify for enrollment in some school (e.g., Science high schools), he/she must demonstrate a certain level of success in the nationwide high school entrance examinations. Those who have high scores in these tests are enrolled in Science high schools. In other words, the academic background of students enrolled in Science high schools is much better than those studying in other high schools. This is because Science high schools use academic achievement as a prerequisite for enrollment. Therefore, in the university entrance exam Science high school students have much higher achievement than ordinary high school graduates. In this context, it is not true that the achievements of the students are assessed as a school-based achievement and that these schools provide a more successful education. This type of assessment is

largely attributed to the enrollment conditions of schools and to the high pre-knowledge of students.

Turkey has been part of many international student assessment programs including PISA (Program for International Student Assessment), TIMSS (Trends in International Mathematics and Science Study), and PIRLS (Progress in International Reading Literacy Study. These projects enable participating countries to evaluate their own education systems and to follow the development of the knowledge and skills of students in the fields of mathematics, science and reading skills according to years. It is not intended for a competition amongst countries. It is expected from the countries to follow the effects of these reforms by ensuring the necessary education reforms and the regular participation in the projects in the country by means of the conclusions (Ministry of National Education of Turkey, 2010). Many studies have been conducted by the MoNE on contextual indicators that relate student and school characteristics to performance outcomes using the data of these studies.

There are many studies that evaluate and interpret the results of these internationally held exams in this way, and compare how the achievements of schools change with respect to each other, or over the years. In all these studies, the average of the scores of students in that school on a national or international (eg PISA) exam is used as a sign of the success of the schools. The common result of studies comparing different school types with each other is that the most successful schools are Science high schools, which are followed by the various groups of high schools and the lowest in the order of success is the vocational high schools (eg Berberoğlu and Kalender, 2005, Farmer 2006, Demir, Kilic and Depren, 2017). The main point that is ignored in these studies is that students are compared to each other without paying attention to socio-cultural characteristics, students pre-knowledge. There is no control group in these evaluations. A group of students is compared with another group of students who are completely different. The main difficulty in evaluating the schools is the fact that the students are being evaluated and sorted, and the evaluation is done by ignoring some major factors that affect student achievement. However, the most important factor in student success is the student himself or herself, with all background knowledge, SES, accumulated educational history etc.

While such a school assessment approach is applied in this way in Turkey, there are different practices in other countries. There are still debates in the US on how exactly to implement an assessment strategy that uses standardized test results. One possible approach is the cohort-to-cohort comparison method, which has gained increasing interest among people engaged in educational policy for the last two decades. In this approach, successes of students in consecutive cohorts are compared against each other (eg 6th graders in the 2017-2018 school year versus 6th graders in the 2018-2019 year) and it is expected that the success of the educators will show an increasing tendency when passing from one cohort to another. In this method students are not followed individually; instead, all of a cohort is compared with past cohorts that have already enrolled in the same class (Ehlert et al., 2013). The method of the Public Schools Accountability Act adopted in California in 1999 includes such a benchmarking approach (McCaffrey, Lockwood, Koretz and Hamilton, 2003).

Early examples of studies for school and teacher effect on student achievement through scientific methods are seen in the 1950s, when trends such as science management and behaviorism gained strength in the United States (Ellett and Teddlie, 2003). Over the next 20 years, American educational scientists have conducted intensive research to identify effective teaching methods. These studies are usually based on observable teacher behaviors, observable student outputs, and a causal relationship. As in the United States, the most commonly used toolkit for determining school effectiveness in the UK since the 1980s is VAA (Department of Education and Science, 1983, 1984; Gray and Jesson, 1987; Gray et al., 1984, 1986; Woodhouse and Goldstein, 1988, Willms, 1987, Kurtz, 2018).

Value Added Assessment Method

Arguably the most common tool used to determine the contribution of schools to student development in recent years is the value added assessment method (VAA). In this approach, the rate of students' success in the school system is monitored, and various statistical methods are used to estimate the contribution of the teacher or school to the increase in the success of students. These methods have been associated with the value-added term borrowed from the literature of economics, as it relates to the contribution of the teacher or school from the moment the student enters the school system (McCaffrey et al., 2003). In the field of value-added assessment, the school or teacher is used to refer to the cumulative effect of education or to the original contribution of student development, independent of the student's own socioeconomic status and environment.

As mentioned before, although there are many factors that affect the academic success of the students, at least it can be said that it would not be fair to hold only schools or teachers accountable for success or failure. For example, the average academic success of a class of 25 students will increase in a nationwide exam if the least successful 5 students drop out the class. Similarly, if the top 5 students leave, this reduces the average score of the class. This change in the average cannot be regarded as success or failure of the teacher or school. This system, which shows student success as a school success, has led to years of unfair evaluation of schools. As a result, schools and teachers have been held unfairly accountable for the academic success or failure of students.

Research findings show that the main factors in student *development* are school and teacher, and other characteristics (socioeconomic level, background knowledge of the student, environment, etc.) are also statistically insignificant (EFF, 2008). For example, Sanders (1997) statistically determined that low socioeconomic level and academic development amounts are irrelevant.

In recent years, VAA has also become a focus of attention particularly in the United States, Canada, the United Kingdom and Australia. Because the evaluation of education as a system is demanded important stakeholders such as parents, educators, academicians, and politicians. To make this assessment, tools that can accurately and qualitatively measure the quality of education are needed. It is not enough to measure students individually in order to find a logical way of evaluating the effectiveness of

the education system. Objective and standardized tools are needed to measure how successfully schools and teachers are meeting their objectives with regards to student attainment. Examinations designed to evaluate the quality of education should be applied regularly and provide equivalent information. The results should provide an objective comparison. VAA is a candidate for hosting all of these features (EFF, 2008).

Studies in the literature show that the effect of teachers is more dominant than the socioeconomic level and other student growth factors (Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997).

Different VAA studies measured the effects of individual teachers and shared these measurements with the school management to improve teacher performance. Among these studies, the most known is TVAAS (Tennessee Value Added Assessment System). Sanders and Horn (1994, 1998) examined the effects of teachers in the 4th and 8th grades of the state of Tennessee since 1996. The Dallas Independent School District also has a system that examines teacher influences and uses their results in the evaluation of official teachers (Webster and Medro, 1997). In a study conducted by William Sanders in the state of Tennessee, state-wide achievement tests conducted across the state found that students trained by effective teachers for 3 years had a score of 50 percent higher than students who attended in the class for 3 years.

It is possible to examine the literature on teacher effects under three main headings:

1) The influence of teachers is very important, and it is indeed arguable that the most important educational input on student development is the effectiveness of the teacher (Rivkin et al., 2000; Rowan, Correnti and Miller, 2002; Wright et al., 1997, Paufler, 2014).

2) Teacher influence is long-term and cumulative (Kain, 1998; Mendro, Jordan, Gomez, Anderson, and Bembry, 1998; Rivers, 1999;

3) Teacher efficacy differs from student to student according to achievement levels of students (Sanders and Rivers, 1996, Loeb, Soland and Fox, 2014).

The studies in the first group prove teacher effectiveness using different methods. The size of the impact and the relative prevalence have been compared with other factors. The studies in the second group are about the permanence of the student's influence in the later education periods of previous teachers. The fact that all of these findings are consistent with each other is proof that the teacher's influence is permanent. However, the magnitude of this persistence may be exaggerated (McCaffrey et al., 2003). The study in the third group shows that the students with the lowest success are the ones who benefit the most from a more effective teacher.

In Turkey, the schools are assessed by the performance of their students in a nationwide exam for passing to a higher level institution. However, both the results of international success projects (eg MoNE, 2003, 2010) and studies in the VAA field literature (eg EFF, 2008; Sanders, 1997) have shown that evaluating schools in this way reflects the socioeconomic status of many students in the success of schools. VAA is a type of method to be used to objectively set the success of schools in Turkey.

VAA distinguishes the contribution of the teacher or the school from the contribution of the student. Based on the academic background of each student in VAA practice, an assessment of this year's success is forecasted. The level of success predicted by a student at a normal level of education determines the level of success. It is important to note that achievement scores that are equal to or higher than the predicted grade are effective (Hershberg, 2008, Everson, 2017). On the contrary, teaching is not effective if the actual results are below the predicted values.

As an example, if a student is in the highest 15% in a nationwide exam in the last two years, it can be assumed that she will be in the upper 13% -17% this year (with a deviation of 2%, for example). If the student becomes in the upper 10% of the cohort, it can be said that the student has developed above normal. If all students in a school or classroom have made a development above normal, it can be said that the school or teacher achieves this successful outcome. We can say that the source of development that is above or below the expectation is the school or teacher. Because the demographic impacts on student achievement are the same for all years of comparison (Hershberg, 2004, 2008, Everson, 2017).

VAA cannot, by itself, explain the reason for student failure. However, when there is as much data, the educators can ask themselves questions about the failure or where the successes are observed. VAA helps teachers to measure the results of their teaching, to help teachers understand what their teaching is focused on (which students have benefited most), and their impact (how effective the year's development is in communicating to students). The development of student achievement may show different patterns according to class, subject, school and region. It is possible to examine these patterns from the results of VAA (Hershberg, 2004, Loeb, Soland and Fox, 2014).

Like every method, VAA has strengths and weaknesses. It takes many years to use the information obtained with VAA. At least a few years of data accumulation is needed to make a sound assessment of schools and teachers. It is very difficult to make a comment about the first year as it allows evaluation based on the development values of the VAA. Despite the strong scientific bases of VAA, it is difficult to apply because the statistical models it is based on are not easy to understand and explain. There are some computer programs available for the use of VAA in the countries where VAA is used and they are evaluating students, teachers and schools using these computer programs. For this reason, schools are required to allocate money and time in order to implement the VAA.

Despite all these difficulties, however, VAA provides very useful data. Managers can make plans for the future through the data provided by the VAA. In addition, successful schools and teachers can be identified, the key factors for their success can be searched, and these factors can be made available to other schools and teachers. If schools more developed schools are compared to other schools and the difference is located in the underlying hardware, school administrations may act to complete the lack of equipment. The results of the VAA allow teachers to be evaluated and followed

up. In some states in the United States, the VAA-based reward systems are used by the school districts (EFF, 2008).

The problem of this research is to determine the contribution of primary schools to student success by means of value added methods.

The purpose of this research is to determine the added value of the education given to the students who are studying in the 7th grade of the primary schools within the borders of the Ankara Metropolitan Municipality with the data obtained from the year-end Level Determination Examinations (SBS). For this purpose, we try to answer the following questions.

1. How much relative value is added by primary schools to their students?

2. Is the order of schools consistent in terms of student success and student development?

The aim of the study is to prove that there are inconsistencies between how the schools are assessed by the authorities and the public and the actual value added by the schools to its students.

Method

Research Design

The research uses a causal comparison model. In the study, the effects of primary schools in Ankara on student performance measured by the SBS is assessed by means of a Value Added Method.

Population and sample

The population of this research is the primary schools in Ankara- capital of Turkey, and the sample is composed of 24 primary schools within the borders of the Metropolitan Municipality, in which the students who studied both in the 2007-2008 and 2008-2009 years are in the same school.

Provinces that formed a sample in the first phase were Çankaya, Altındağ, Yenimahalle, Mamak, Keçiören, Gölbaşı, Sincan and Etimesgut. One of the purposes of the research is to understand the relationship between socioeconomic status of the students and the value added by the schools to these students. In the second phase, the three socioeconomic levels in these provinces are low, medium and high, determined by random selection method. In the third stage, the schools were ranked according to the socioeconomic level. In this order, schools with low socioeconomic level 3, schools with medium socioeconomic level 2, and schools with high socioeconomic level are numbered 1. The selection of the socioeconomic levels of the schools was made with the help of an expert working in the provincial directorate of national education in Ankara. One of the main reasons to select this city was that it fairly represents Turkey. Data for the 24 aforementioned schools have been requested,

with 22 schools complying. The study was conducted on data from these 22 schools. The names of the schools are provided in Table 6 in the Appendix.

The Level Determination Exam (SBS)

In Turkey, the level determination exam (SBS) is a central exam applied by the ministry of national education. Students are allocated to high schools with respect to the weighted average of multiple SBS exams applied at the end of 6th, 7th, and 8th grade. SBS exams started in 2008 to be applied only to 6th and 7th graders, and finalized in 2013. Other than the SBS system, high school entrance exams consist of a single exam at the end of 8th grade and the data is not suitable to use in a VAA study. The SBS exam consists of five subject areas: Turkish, mathematics, science, social sciences, and foreign Language-English. The number of right answers in each subject area is refereed to as *number right (NR)* score and the NM score which is calculated as NR minus the number of wrong answers/3 for each subject area is referred to as *negative marking (NM)* score.

Data Collection

The data for this study comprises the students who enrolled in the 6th and 7th grades in the schools within the borders of Ankara Metropolitan Municipality over the years 2008 and 2009. This data has been taken in the electronic environment with the necessary permission as a result of the correspondences made with the Ministry of National Education, EĞİTEK.

In VAA, data are not experimental but observational (McCaffrey et al., 2003). For this reason, the observational data are not encountered except for the problems to be experienced in collecting. Observational data is the fact that the largest problem data is not complete. Despite the fact that 24 of the records were requested from the EĞİTEK, 22 studies could be reached. The research group was selected as students who were in grades 6 and 7 in the same school. In this respect, data incompleteness problems that may arise from the students who are transported are prevented. It is also considered that the actual contributions of schools can be determined independently of the effect of school change on the student.

Analysis of the data was done using "SPSS 13.0 for Windows" and "Microsoft Excel 2007" programs. There are known true and false numbers in the SBS scores and subtests that all students in each school receive from two exams. The change in NM scores and true / NM scores of these students are calculated using Microsoft Excel. Then the average scores of the students in each school were calculated, which was stated as the school average. Some outliers were detected using z-scores approach. For example, students who did not attend one of the two exams did not participate in the analysis. A student may not have been able to raise any sub-test questions in the second test and may have shown very low success because he did not use the exam. This change is not caused entirely by the influence of the school.

Data Analysis

The main statistical models used in the VAA are examined under three main headings: auto-regressive approaches (Hanushek, 1972, Murnane, 1975, Rowan et al., 2002, Sanders and Rivers, 1996, Webster and Medro, 1997), development-based approaches (Allison, 1990, Bryk and Weisberg, 1976; Lord, 1969; Rogosa, 1995; Rowan et al., 2002; Thum, 2001) and multivariate approaches (Sanders, Saxton and Horn, 1997; Ballou, Sanders and Wright, 2004).

In the most well-known examples of VAA, a multivariate model that uses many years of accumulated data is used. In this study, we only used two years' data. This was mainly due to difficulty in obtaining data in Turkey.

In this study, a development based approach: simple fixed effect model is used in the study. The reasons for selecting this approach is because:

- this model contains fewer variables,

- the solution is easy to understand and explain,

- some studies (Tekwe et al., 2004) show that more complex models with this model give similar results,

- the availability of data that does not allow the use of more detailed models

There are many factors, both in-school and out-of-school, that affect a student's score from exams or subtest scores. The objective of the VAA approach is to determine the contribution of the school by eliminating the effects of extracurricular factors. A student's score can be modeled linearly as follows.

$$y_{ijst} = b_{st} + m_{ijs} + SED_{ij} + O_{ist} + e_{ijst} \tag{6}$$

In this model, the score of student j in the i-th school in subtest s at year t is related to b_{st} (a constant for test s, year t), to the influence of the socioeconomic level (SED), the school effect (O), and an error term (e), as a constant number, the effect of the incremental knowledge of the student on this subject (m). The error term here assumes a normal distribution with a mean of 0 and a standard deviation of σ_{ijst}. The following is the score obtained by a student in the sequential tests conducted in 2008.

$$y_{ijs,2008} = b_{s,2008} + m_{ijs} + SED_{ij} + O_{is,2008} + e_{ijs,2008} \tag{7}$$

$$y_{ijs,2009} = b_{s,2009} + m_{ijs} + SED_{ij} + O_{is,2009} + e_{ijs,2009} \tag{8}$$

It is not expected that the background knowledge and socioeconomic level of a student will change very much from one year to the next, and such a change is considered as an outlier. Information on the determination of such outliers and their discarding from the analysis is given in the section on analysis of data. Apart from this, the difference between the scores of students who have taken two years of top-up examinations will be as follows.

$$y_{ijs,2009} - y_{ijs,2008} = \left(b_{s,2009} - b_{s,2008}\right) + \left(O_{is,2009} - O_{is,2008}\right) + \left(e_{ijs,2009} - e_{ijs,2008}\right) \tag{9}$$

In the above equation, the SED and the incremental contributions (m) in the two-year period are not included in this equation. The reason is that it is assumed that the SED and the incremental contributions of the students do not change from year to year. Based on the knowledge that the increasing knowledge of students and their socioeconomic levels have not changed, it is observed that this difference is not related to the SED or the information (m) when the students score difference in the following years. Thus, each student forms his own control group and the contribution of the school is revealed in a way free from SED and student background. The above equation can be edited as follows.

$$\left(O_{is,2009} - O_{is,2008}\right) = \left(y_{ijs,2009} - y_{ijs,2008}\right) - \left(b_{s,2009} - b_{s,2008}\right) - \left(e_{ijs,2009} - e_{ijs,2008}\right) \tag{10}$$

By solving this equation system for all students, the contribution each school makes to its students in 2009 can be calculated. Because, in the above equation, the left side of the equation is the contribution of the school up to 2009 with the contribution up to 2008, which is the contribution the school has made to its students in 2009.

By expanding this method to include more years, a multi variable structure can be obtained. However, in this study, the simple fixed effect model, which is the most appropriate method, has been used.

Fixed Simple Effects Model:

In this method, in a sample of n schools, the change in student scores is modeled as follows:

$$d_{ijs} = \beta_{0s} + \sum_{k=1}^{n-1} \beta_{1ks} s_{kij} + e_{ijs} \tag{11}$$

Here,

$$d_{ijs} = y_{ijs2} - y_{ijs1} \tag{12}$$

y_{ijst}, is the score from test s of student j in school i at year 2009 (t=2)

$$s_{kij} = \begin{cases} 1, & i = k, i \neq n \\ 0, & i \neq k, i \neq n \\ -1, & i = n, k = 1,2,\cdots,n-1 \end{cases} \tag{13}$$

$e_{ijs} \sim N\left(0, \sigma_{es}\right)$, d_{ijs} is the score difference for student j at school i from test s.

The β_{0s} constant at Equation (11) corresponds to $\left(b_{s,2009} - b_{s,2008}\right)$ in Equation (9), and β_{1ks} is the value added to students of school k for test s, which corresponds $\left(O_{is,2009} - O_{is,2008}\right)$ in equation (9). In summary, if equations (11) are to be written

explicitly, the following equation system is obtained, assuming that the average contribution of all schools is 0.

$$
\begin{aligned}
d_{1js} &= \beta_{0s} + \beta_{1,1s} + e_{1js} \\
d_{2js} &= \beta_{0s} + \beta_{1,2s} + e_{2js} \\
&\vdots \\
d_{22js} &= \beta_{0s} + \beta_{1,22s} + e_{22js}
\end{aligned}
\tag{14}
$$

Adding all equations side by side and taking the expected value results in

$$
\beta_{0s} = \bar{d}_s - \mathrm{E}[e_s]
\tag{15}
$$

And as the expected error value is 0, $\beta_{0s} = \bar{d}_s$ .In other words, it is found as the mean changes of all schools in the sample in *s* test. In this case, the average contribution of all schools indicates the relative contribution of the school to this average. In other words, the contribution of some schools will be positive, others will be negative; because the average contribution of all schools is assumed to be zero. This is mainly because this method calculated the contribution of individual schools relative to the others. Hence, as the analysis is in relative terms there is no need for vertical scaling.

The average score change of each school and their averages were calculated by Microsoft Excel program and the variables in the simple fixed effect model were calculated after eliminating the outliers. Then, the standard deviation of the school contributions is calculated, and the "z-score" indicating how many standard points each school has deviated from the average is calculated for each school and subtest as follows.

$$
z_{ks} = \frac{\beta_{1ks}}{\sigma_s}
\tag{16}
$$

Here, $\square$ is the school effect for test *s*, and $\square_s$ is the standard deviation of school effects. The z-scores are then used to classify the schools according to the following table (Tekwe et al., 2004).

Table 1

Using z-scores for Classification

z-score	Classification
$z \geq 2$	A
$2 \geq z \geq 1$	B
$1 \geq z \geq -1$	C
$-1 \geq z \geq -2$	D
$-2 \geq z$	F

After the calculation of the school contributions, the schools were ranked according to these contributions and the relationship between success rankings and contribution (development) rankings was analyzed. For this purpose, the degree of correlation between achievement orders and contribution (development) orders according to the 6th and 7th grade NR / NM numbers for each course was calculated according to kendall's tau rank correlation coefficient. SPSS 13.00 for Windows program was used to draw scatter diagrams that reflect the relationship between achievement rankings (NR / NM change for 6th and 7th grade) and contribution rankings (NR / NM change) for all courses separately.

Results

In this section, the analysis made based on the questions that are sought in the general purpose frame of the research, the findings and interpretations belonging to these analyzes are included. In the presentation of the findings, the order of the questions to be answered in the study was followed and the results of the analysis were given according to the courses.

How Much Relative Value is added by Primary Schools with respect to Their Students?

In this section, the interpretation of findings and interpretations are handled according to Turkish and Mathematics courses, respectively. There are two test scores for a student, 1: the number of question that the student *correctly* answered (NR score) and 2: the net score which is calculated by subtracting 1/3 of the wrong answers from the correct answers (NM score). The average NR and NM scores of schools for 2008 and 2009, calculated separately for each course, β (Beta) values calculated according to the fixed effect model, transformed into z-points and classified. Findings about social studies, foreign language and science courses can be found in the author's masters thesis.

Turkish Course

In Table 2, the NR and NM scores for Turkish test are calculated and analyzed. The first column of the table shows the socioeconomic status (SES) of the region where the schools are located. It was determined as 1: high, 2: medium and 3: low in this column. In the second and third columns, the students average NR and NM scores for 6th grade test is provided. The NR and NM scores in the 7th grade SBS are given in the fourth and fifth columns. The score difference between grade 7 and grade 6 was calculated as the score of change and was given in columns six and seven. Beta (β) school contributions are given in the eighth and ninth columns, and z-scores in the tenth and eleventh columns. The z-scores have been translated into the classification as given in Table 2 and shared in the last two columns of classification of schools in terms of NR and NM scores for Turkish course. The last two lines of the chart provide averages of school average scores and standard deviations.

Table 2

Turkish Course Statistics and Classification

SED	School number	6th grade		7th grade		Change		(β) School Contribution		z-score		Classification	
		NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	16.934	16.325	16.231	14.791	-0.702	-1.534	0.682	0.914	1.762	1.807	B	B
2	School 2	14.511	13.179	13.538	11.240	-1.027	-2.011	0.357	0.438	0.923	0.865	C	C
3	School 3	12.127	10.135	10.363	7.208	-1.747	-2.903	-0.363	-0.454	-0.938	-0.898	C	C
1	School 4	13.801	12.362	11.590	8.768	-2.200	-3.581	-0.816	-1.133	-2.109	-2.239	F	F
2	School 5	12.560	10.716	11.197	8.250	-1.387	-2.498	-0.003	-0.049	-0.007	-0.097	C	C
3	School 6	10.160	7.510	8.532	4.728	-1.515	-2.603	-0.131	-0.154	-0.339	-0.305	C	C
1	School 7	13.775	12.308	12.050	9.500	-1.725	-2.808	-0.341	-0.360	-0.881	-0.711	C	C
2	School 8	16.236	15.455	15.554	13.878	-0.755	-1.676	0.629	0.772	1.625	1.526	B	B
1	School 9	15.370	14.353	14.096	12.075	-1.312	-2.327	0.072	0.121	0.187	0.240	C	C
2	School 10	13.582	11.941	11.963	9.280	-1.456	-2.464	-0.072	-0.016	-0.185	-0.031	C	C
3	School 11	11.706	9.833	10.486	7.352	-1.235	-2.490	0.149	-0.042	0.385	-0.082	C	C
1	School 12	15.379	14.422	14.235	12.183	-1.113	-2.197	0.271	0.251	0.700	0.497	C	C
2	School 13	12.347	10.427	11.180	8.322	-1.176	-2.055	0.208	0.394	0.539	0.778	C	C
1	School 14	13.931	12.561	12.502	10.040	-1.416	-2.503	-0.032	-0.055	-0.082	-0.108	C	C
2	School 15	11.789	9.649	10.895	7.886	-0.895	-1.763	0.489	0.685	1.265	1.355	B	B
3	School 16	13.527	11.938	11.450	8.592	-2.081	-3.347	-0.696	-0.899	-1.800	-1.777	D	D
1	School 17	14.339	13.018	12.875	10.423	-1.464	-2.595	-0.080	-0.147	-0.207	-0.290	C	C
2	School 18	12.391	10.465	11.166	8.300	-1.258	-2.205	0.126	0.243	0.327	0.481	C	C
3	School 19	13.378	11.685	11.824	9.069	-1.537	-2.515	-0.152	-0.067	-0.394	-0.132	C	C
1	School 20	15.386	14.426	14.261	12.285	-1.091	-2.102	0.293	0.346	0.758	0.684	C	C
2	School 21	14.638	13.438	12.813	10.401	-1.827	-3.037	-0.443	-0.588	-1.144	-1.163	D	D
3	School 22	12.301	10.393	10.829	7.824	-1.533	-2.650	-0.149	-0.202	-0.384	-0.399	C	C
Average of Averages		13.644	12.115	12.256	9.654	-1.384	-2.449	0.000	0.000				
Std. dev. of Averages		1.646	2.135	1.822	2.363	0.387	0.506	0.387	0.506				

NM Score: Correct answers, NR Score: Correctly answered

In Table 2, the 6th grade average Turkish NR score is 13.644 and the 7th grade average Turkish NR score is 12.256. There is a decrease of 1.384 points in the Turkish NR score for the sample. Similarly, while the 6th grade Turkish NM score was 12.115, it dropped to 9.654 with a decrease of 2.449 points. The standard deviation of the change scores is 0.387 and 0.506 for the NR and NM scores. The classification of schools is summarized in Figure 1. As shown in the figure, many schools are considered to be the average accepted class C by looking at their z-scores. While the average NR of schools in the sample decreased by 1.384, the average NR score of students with high SED School #1 decreased by only 0.702, which caused the school to be classified as class B.
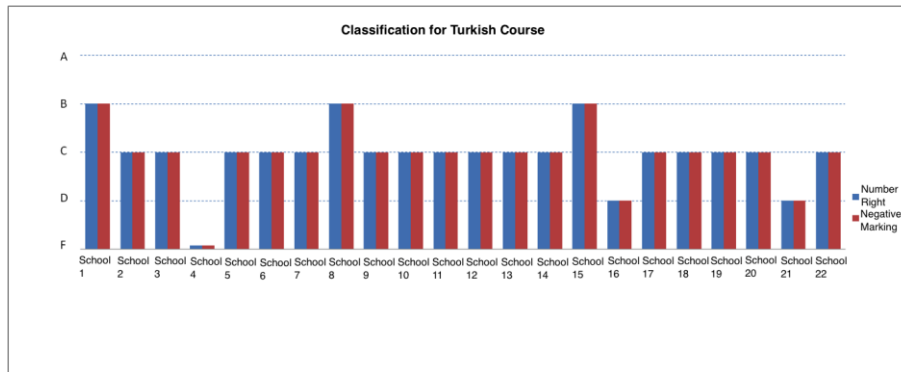
Figure 1. Classification for Turkish Course

Likewise, School #8 and #15, both of which are composed of students with high SES, have been classified as B. For School #21, which is low in SES and School #16 with medium SES, NR and NM scores in Turkish showed a decline more than the average. For these reasons, these schools were classified as class D. School #4, who had a high SES, was evaluated as class F. This school did not provide adequate improvement for its students with a NR average score of 13.801 and NM average score of 12.362 in the 6th grade. The NR score dropped by 2.720 in the 7th grade and to a 11.590 average (3.581 NM point reduction to 8.768). This decline was attributed to the fact that the school was classified as class F, with a considerable decline compared to all the schools in the sample. School #14, which has similar average 6th grade scores and same SES, provided an average improvement (-0.032 in the NR score, -0.055 in the NM score) whereas the contribution of the School #4 was -0.816 for the NR score and -1.133 for the NM score, respectively.

Mathematics Course

Table 3 corresponds to results of Mathematics course. The average NR score was 7.778 (NM score of 5.436) for the 6th grade and it is reduced by 2.396 points to 6.065 in the 7th grade. The average NM score reduced by 1.712 points to a level of 3.051. The

standard deviation of the change scores was 0.440 for the NR score and 0.499 for the NM score. The classification of schools is summarized in Figure 2. As shown in the figure, many schools are considered to be the average class C by looking at their z-scores. The average NR of all schools decreased by 1.712. The average NR score of School #6 students dropped by only 0.889, which led to this school being classified as class A.
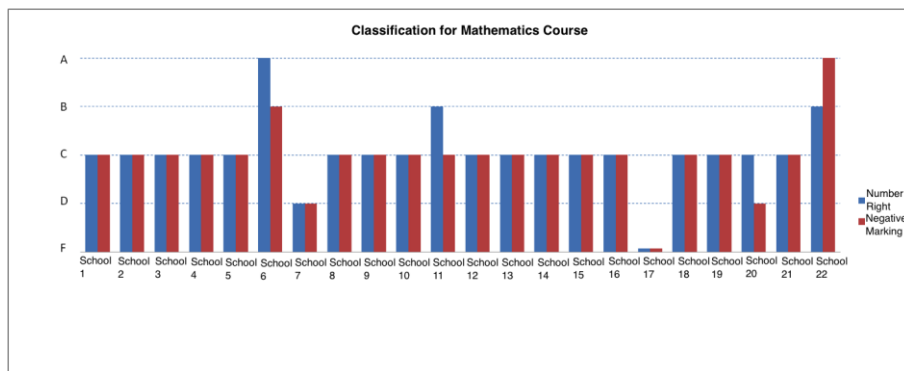
Figure 2. Classification for Mathematics Course

Likewise, School #11 and School #22, whose socioeconomic status are low, have been evaluated as A or B class in terms of their NR or NM scores. For School #7 and School #20, who have high SES, the NR and NM scores in Mathematics has declined above the average. For this reason, these schools were classified as class D. School #17, which is composed of students with high SES, is considered as F class. This school did not provide adequate contribution to the students in the 6th grade. The average NR score reduced from 9.018 to 6.268, and NM scores reduced from 7.042 to 3.494 in the 7th grade. These reductions (2.750 for NR and 3.548 points for NM scores) are at least two standard deviations away from the mean. This extremely high reduction in the points resulted in an F classification for this school in Mathematics course. The average scores of School #9, which is composed of students with the same SES, was close to School #17, in the 6th grade test for mathematics test. This school provided an average improvement for its students, whereas School #17, couldn't provide this improvement.

Is the Order of Schools Consistent in terms of Student Success and Student Development?

The differences between the NR and the NM scores of the schools in 2008 and 2009 for Turkish and Mathematics courses were calculated. In addition, a correlation analysis was performed between 2008 and 2009 student scores. The aim of the analysis is to investigate if successful students in first year continue to be successful in the second year. We also analyzed the relationship between average student success and average improvement of the students. Moreover, we compared the order of schools in terms of their average points in the first and second year with the average improvement they provide to their students.

The correlation between Turkish test NR scores for all the students in the 6th and 7th grades is 0.754. The correlation for the NM Turkish scores of both classes is found to be 0.756. These results show that there is a high positive correlation between Turkish NR and NM scores for the two consecutive years. The students who have high NM or NR score in the first year have a high NM or NR score in the second year. This is expected as students with relatively better background (who scored high in 6th grade) tend to be more successful in the 7th grade.

In the correlation analysis to determine the relationship between grade 6 points and change points, the change score of each student was found by subtracting grade 6 from grade 7 score. Correlation between students' sixth grades and change points (grades 7 - 6) was found to be -0.26 for the Turkish NR score and -0.28 for the Turkish NM score. This indicates that there is almost no (if not slightly negative) relationship between students' change scores and grade 6 scores. When combined with the high correlation between 6th and 7th grade scores, this result is extremely important. This shows that the change points are independent of students background.

In Figure 3, scatter diagrams are given for all pairs of Turkish scores 6th and 7th grade NR and NM averages and change averages. As can be seen in the figure, while there is a high level linear relationship between NR and NM scores in both 6th and 7th grade, there is almost no relation between those with the average change (improvement).

Table 4 examines the schools' consistency in terms of student average student success and student improvement. For the Turkish course, each school is ranked in the order of the 6th grade NR and NM average score, the 7th grade NR and NM average score, and the average NR and NM score improvement among the 22 schools in the sample.
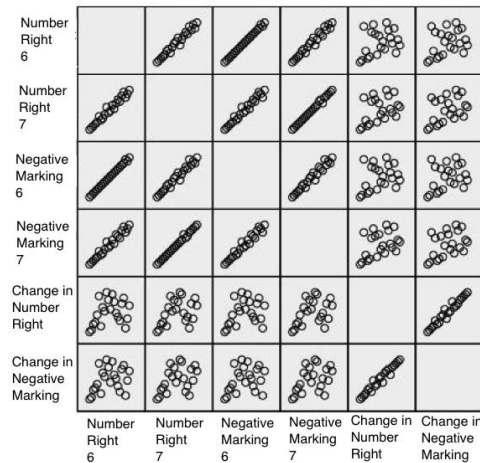
Figure 3. Average success and improvement for Turkish Course (all pairs)

Table 3

Mathematics Course Statistics and Classification

SED	School number	6th grade		7th grade		Change		(β) School Contribution		z-score		Classification	
		NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	11.702	10.499	9.901	7.774	-1.802	-2.725	-0.090	-0.329	-0.223	-0.659	C	C
2	School 2	7.823	5.418	6.237	3.222	-1.643	-2.268	0.068	0.127	0.170	0.255	C	C
3	School 3	6.722	3.937	5.05	1.563	-1.646	-2.367	0.066	0.029	0.164	0.057	C	C
1	School 4	7.719	5.311	6.021	2.889	-1.677	-2.393	0.035	0.002	0.087	0.005	C	C
2	School 5	6.813	4.169	5.276	2.048	-1.392	-2.098	0.320	0.298	0.795	0.597	C	C
3	School 6	5.570	2.453	4.642	0.890	-0.889	-1.458	0.823	0.938	2.045	1.879	A	B
1	School 7	8.300	6.200	6.150	3.292	-2.410	-3.239	-0.699	-0.844	-1.736	-1.691	D	D
2	School 8	10.15	8.543	7.993	5.631	-2.094	-2.826	-0.382	-0.430	-0.951	-0.863	C	C
1	School 9	9.014	7.065	7.023	4.275	-1.991	-2.790	-0.279	-0.394	-0.694	-0.790	C	C
2	School 10	7.203	4.662	5.519	2.436	-1.582	-2.148	0.129	0.248	0.322	0.497	C	C
3	School 11	6.176	3.725	4.886	1.562	-1.294	-2.167	0.418	0.229	1.038	0.459	B	C
1	School 12	8.919	6.936	7.167	4.572	-1.739	-2.34	-0.027	0.056	-0.068	0.112	C	C
2	School 13	7.096	4.487	5.314	2.107	-1.817	-2.439	-0.105	-0.043	-0.262	-0.087	C	C
1	School 14	7.318	4.973	5.712	2.730	-1.598	-2.227	0.114	0.168	0.282	0.337	C	C
2	School 15	6.776	4.149	5.368	2.114	-1.408	-2.035	0.304	0.361	0.755	0.723	C	C
3	School 16	7.363	4.838	5.777	2.546	-1.545	-2.234	0.167	0.162	0.415	0.324	C	C
1	School 17	9.018	7.042	6.268	3.494	-2.750	-3.548	-1.038	-1.152	-2.581	-2.309	F	F
2	School 18	6.913	4.265	5.279	2.098	-1.693	-2.246	0.019	0.150	0.047	0.301	C	C
3	School 19	7.128	4.543	5.430	2.364	-1.695	-2.169	0.017	0.227	0.041	0.455	C	C
1	School 20	9.109	7.170	7.037	4.285	-2.075	-2.918	-0.363	-0.522	-0.902	-1.046	C	D
2	School 21	8.307	6.144	6.523	3.542	-1.787	-2.685	-0.076	-0.290	-0.188	-0.580	C	C
3	School 22	5.967	3.054	4.854	1.686	-1.131	-1.385	0.581	1.010	1.443	2.025	B	A
Average of Averages		7.778	5.436	6.065	3.051	-1.712	-2.396	0	0				
Std. dev. of Averages		1.445	1.879	1.213	1.554	0.402	0.499	0.402	0.499				

NM Score: Correct answers, NR Score: Correctly answered

When Table 4 is examined, it is clear that there are serious discrepancies between the successes and improvement orders of schools. For example, School #11, School #13 and School #15, which are left behind in the average success order, are in the higher rankings in terms of improvement. On the other hand, for the schools with high order rank with respect to average student success (School #9, School #17 and School #21), their order in terms of student improvement is very low. Some schools, such as School #1 and School #8, did not show any difference in terms of success and improvement. To analyze the difference of ranking on success and improvement, Kendall tau rank correlation coefficient between success and improvement rankings was calculated as 0.290 in terms of the 7th grade NR scores and 0.325 for NM scores. These values also indicate that there is a statistically low level of correlation between average student success and improvement ranks, and therefore these ranks are not consistent with each other.

Table 4

Turkish Course Success and Improvement Comparison

SED	School number	6th grade		7th grade		Change	
		NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	1	1	1	1	1	1
2	School 2	7	7	6	6	4	4
3	School 3	19	19	21	21	19	19
1	School 4	10	10	13	13	22	22
2	School 5	15	15	15	17	11	12
3	School 6	22	22	22	22	15	16
1	School 7	11	11	10	10	18	18
2	School 8	2	2	2	2	2	2
1	School 9	5	5	5	5	10	9
2	School 10	12	12	11	11	13	10
3	School 11	21	20	20	20	8	11
1	School 12	4	4	4	4	6	7
2	School 13	17	17	16	15	7	5
1	School 14	9	9	9	9	12	13
2	School 15	20	21	18	18	3	3
3	School 16	13	13	14	14	21	21
1	School 17	8	8	7	7	14	15
2	School 18	16	16	17	16	9	8
3	School 19	14	14	12	12	17	14
1	School 20	3	3	3	3	5	6
2	School 21	6	6	8	8	20	20
3	School 22	18	18	19	19	16	17

NM Score: Correct answers, NR Score: Correctly answered

The correlation between the Mathematics course NR scores of all the students in the 6th and 7th grades is 0.651, and the correlation between the Mathematics course NM scores is 0.659. These results show that there is a moderate positive correlation between mathematics NR and NM scores made in both years. Students with high NR or NM scores in the first year have had a high NR or NM scores in the second year. This is expected as student background is one of the main drivers of student success.

In the correlation analysis to determine the relationship between grade points and change points, the change score of each student was found by subtracting grade 6 score from grade 7 score. Correlation of the change points with grade 6 scores was examined. The correlation between students' sixth grades and their exchange scores (grade 7 - grade 6) was found to be -0.437 for the mathematics course NR score and -0.474 for the NM scores. These value indicates that there is a moderate negative correlation between the change scores of the students and the 6th grade points. This suggests that students with high academic achievement in the 6th grade showed more decline than students with low academic achievement.

In Figure 4, scatter diagrams are given for all pairs of Mathematics course 6th and 7th grade NR and NM averages and 7th grade NR and NM improvement averages of schools. As seen in Figure 4, there is a high linear relationship between success levels in 6th and 7th grade for both NR and NM scores. However, there is a moderate *negative* linear relationship between 6th grade scores and improvement (change) scores. Similarly, there is a negative linear relationship between 7th grade scores and improvement scores.

Table 5 examines whether schools are consistent in terms of student achievement and student improvement. In the Mathematics class, each school is ranked in the order of the 6th grade NR and NM score average, the 7th grade NR and NM score average, and the average NR and NM change scores.

On examining Table 5, it is clear that there are serious inconsistencies between success and development sequences of schools. For example, School #3, School #5, School #6, School #10, School #11, School #15 and School #22 are in higher rank in terms of change, while School #1, School #8, School #9, School #12, School #20 and School #21 are in lower orders in terms of change. Some schools such as School #2 and School #13 have not observed a serious difference in terms of their success and improvement. The statistical comparison of the rankings with respect to each other, the Kendall tau rank correlation coefficient between success and development ranks in terms of the 7th grade NR averages was calculated to be -0.576. The same coefficient is calculated as -0.550 between success and improvement orders in terms of 7th grade NM score averages. These values also indicate that there is a moderate negative relationship between success and improvement. Therefore, there is a serious inconsistency between the improvement and success sequences of schools.
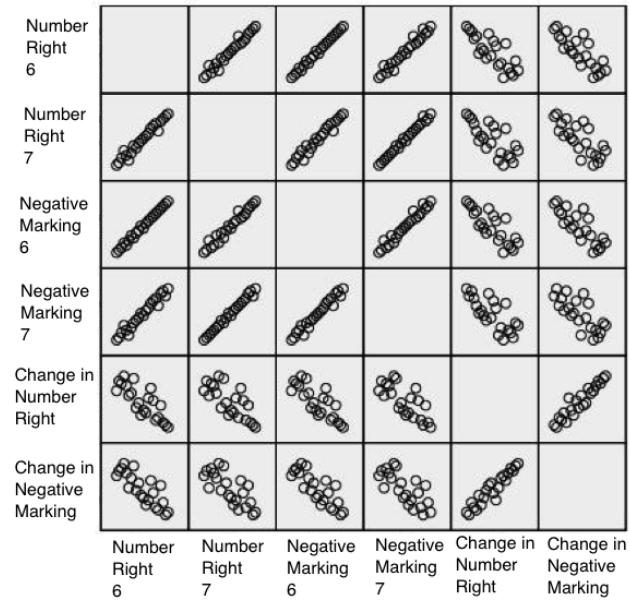
Figure 4. Average success and improvement for Mathematics Course (all pairs)

In summary, the added value of schools for the first research question is given under the Beta (β) column in the relevant schedule for each course. Schools are also categorized in terms of the contribution they provide for each course, and in the same charts, this classification is shared with the reader.

As regards the second research objective, it has been observed that schools with high student achievement are classified as F class in terms of student development in some courses. It was also determined that schools with low student achievement contributed significantly to some courses. In this case, the order of schools in terms of student success and student development is inconsistent. The level of this inconsistency varies according to the courses. However, there is no clear relationship and alignment between the success and the development order of the schools.

These findings do not provide information on the *reasons* for the differences in the development of schools indeedthere can be many reasons why the schools are different from each other. These include, for example, the physical conditions of schools, lack of teachers, the personal abilities of teachers in teaching, and out-of-school assignments of teachers. It can be determined whether there are problems with schools by using the VAA method, but the type or solution method of the problem is not of interest to the VAA. Detailed and long-term monitoring of schools is required to clarify the causes of differences between schools.

Table 5

Mathematics Course Success and Improvement Comparison

SED	School number	6th grade		7th grade		Change	
		NR score	NM score	NR score	NM score	NR score	NM score
1	School 1	1	1	1	1	16	17
2	School 2	9	9	8	9	9	11
3	School 3	19	19	19	20	10	13
1	School 4	10	10	10	10	11	14
2	School 5	17	17	18	18	4	4
3	School 6	22	22	22	22	1	2
1	School 7	8	7	9	8	21	21
2	School 8	2	2	2	2	20	19
1	School 9	5	4	5	5	18	18
2	School 10	13	13	13	13	7	5
3	School 11	20	20	20	21	3	6
1	School 12	6	6	3	3	14	12
2	School 13	15	15	16	16	17	15
1	School 14	12	11	12	11	8	8
2	School 15	18	18	15	15	5	3
3	School 16	11	12	11	12	6	9
1	School 17	4	5	7	7	22	22
2	School 18	16	16	17	17	12	10
3	School 19	14	14	14	14	13	7
1	School 20	3	3	4	4	19	20
2	School 21	7	8	6	6	15	16
3	School 22	21	21	21	19	2	1

NM Score: Correct answers, NR Score: Correctly answered

Discussion, Conclusion and Recommendations

In this study, we assessed the value-added effects of the education received by 7th graders that attended primary schools within the territorial jurisdiction of the metropolitan municipality of the province of Ankara using the results of the end of year Level Determination Exam.

We determined that the contribution of the school to student success is independent of the academic achievement of the students in the school. Moreover, school contribution is also independent of the socioeconomic status of the students. It has been determined that high academic achievement and schools with students with SES do not always make a high contribution to these students. These results are consistent with many VAA studies (see for example Sanders, 1997; Sanders and Rivers, 1996; Wright, Horn and Sanders, 1997).

In some countries including the US, UK, Canada, and Australia, VAA methods are used to evaluate school and teacher performance. A similar approach can also be

applied in Turkey by the Ministry. VAA is a statistical approach that allows schools to be evaluated in a fairer and objective manner.

Efficient schools and their activities will be a model for less effective schools, which will lead the education system to improve as a whole. Hershberg (2004) draws an analogy between VAA methods for education and a stop watch for a running team. A stop watch cannot make a runner faster; however, it can help the team to understand which members should run in different races. Likewise, VAA methods can be used to improve the effectiveness of an educational system. In order for this to happen, first of all effective and less effective schools need to be identified. With this in mind, in this study, we classified the schools with respect to their contribution to the student success using VAA methods.

We have shown that some schools which are known to be successful (School #4 in Turkish course, School #17 in Mathematics course,), were not able to contribute as much to their students as schools that are thought to be unsuccessful. There are also cases where low academic achievement and SES schools contribute greatly to their students. This also supports the findings in the literature in saying that the effect of the teacher differs with respect to the achievement levels of students (Sanders and Rivers, 1996, Loeb, Soland and Fox, 2014).

For example, School #6, School #11 and School #22 have low academic achievement and SES but have been contributed more to their students than other schools in the Mathematics course. In addition to these, there are also cases where high academic achievement and high SES schools contribute greatly to their students. For example School #1 and School #8 have made serious contributions to their students in Turkish course.

It has also been determined that the change in student scores does not depend on previous academic achievement and SES. Although the academic success of the students is low, it has been seen that a good education may make a great contribution to the student.

Successful and less successful schools have been identified by the VAA method. Less successful schools should benefit from the experience of high contributing schools. The positive impact of the facilities of teachers and schools who contribute more to their students and the experience of these teachers can also be used by other teachers. Letting less effective schools to analyze and imitate the more effective schools is one of the significant benefits of VAA methods (McCaffrey et al, 2003).

While VAA methods are used heavily for accountability reasons, there are opposing views on the value added methods. Haertel (2013) criticizes VAA models in terms of validity and reliability in evaluating teacher VAM estimates. They conclude that scores must be based in appropriate and sound test scores, comparisons should be made amongst homogeneous group of teachers, and VAM should be used flexibly to evaluate teachers. They also mention that users should be well trained to interpret the scores and there should be a clear understanding of the inherent uncertainty in the analysis.

Analysis in future studies should be expanded to include many more years, and exams. Teachers and schools that provide superior added value for a certain period of time can be rewarded according to the amount of contribution calculated by the VAA method. This reward system can also increase the effort of other teachers and schools, which can lead to an overall increase in the quality of education. Complex models can be used if more detailed information is provided. In these complex models, it can be determined whether the contributions of the school or teacher are dependent on other variables, and if so, how these dependencies are shaped. In addition, analyzes can be done on a school-by-teacher basis. This ensures that both school and teacher contributions can be distinguished in student development.

For a more successful VAA application, a large database of long-term exam results of students, including school, classroom and teacher changes is needed. In order for schools and teachers to be judged fairly, it is necessary to establish a standardized examination system that will last for many years. The exams should be used to evaluate not only students, but also teachers and schools.

As a result of this study, there are important decisions that need to be taken in terms of educational policies. It is inevitable that investments for schools with a low level of development should be increased rapidly, and education opportunities should be distributed equally among all schools.

One of the main limitations of this study is that it uses only two years' data. Using more data which is spread over multiple years would be helpful to estimate the schools' effects in a better way. Moreover, we assume that the schools are the only responsible parties for the change in the student scores. However, we did not calculate the effect of teachers independent of the schools. This is mainly because data provided by the authorities includesß only the school information. There was no data available for teacher and/or class codes.

References

Allison, P. D. (1990). Change scores as dependent variables in regression analysis. *Sociological Methodology*, *20*, 93-114.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics, 29* (1), 37-65.

Başaran, İ., & Çınkır. Ş. (2013). *Türkiye Eğitim Sistemi ve okul yönetimi.* Ankara: Ekinoks Yayınları.

Batyra, A. (2017). Gender gaps in student achievement in Turkey.

Berberoğlu, G. (2006). *Sınıf İçi ölçme değerlendirme teknikleri.* İstanbul: Morpa.

Berberoğlu, G., & Kalender, İ. (2005). Öğrenci başarısının, yıllara, okul türlerine, bölgelere göre incelenmesi: ÖSS ve PISA analizi. *Eğitim Bilimleri ve Uygulama, 22,* 21-35.

Bryk, A., & Weisberg, H. (1976). Value-added analysis: A dynamic approach to the estimation of treatment effects. *Journal of Educational Statistics*, *1*, 127-155.

Ciftci, A. (2006). *PISA 2003 Matematik alt testi sonuçlarına göre Türkiye'deki öğrencilerin başarılarını etkileyen bazı faktörlerin incelenmesi.* Yayımlanmamış Yüksek Lisans Tezi, Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara.

Demir, İ., Kılıç, S., & Depren, Ö. (2009). Factors affecting Turkish students' achievement in Mathematics. *US-China Education Review, 6* (6), 47-53.

EFF. (2008). *School directors' handbook.* Evergreen Freedom Foundation. Web: http://www.effwa.org/pdfs/Value-Added.pdf adresinden 11 Eylül 2008'de alınmıştır.

Ehlert, M., Koedel, C., Parsons, E., & Podgursky, M. (2013). Selecting growth measures for school and teacher evaluations: Should propor- tionality matter? *National Center for Analysis of Longitudinal Data in Education Research*, 21, 1–33. [1,5,7]

Ellett, C. D., & Teddlie, C. (2003). Teacher evaluation, teacher effectiveness and school effectiveness: Perspectives from the USA. *Journal of Personnel Evaluation in Education*, *17*(1), 101-128.

Everson, K. C. (2017). Value-added modeling and educational accountability: Are we answering the real questions? *Review of Educational Research*, *87*(1), 35-70.

Gray, J., & Jesson, D. (1987). Exam results and local authority league tables. In A. Harrison & J. Gretton (ed.), *Examination and training UK,* (pp. 33-41).

Gray, J., Jesson, D., & Jones, B. (1984). Predicting differences in examination results between local education authorities: Does school organisation matter?. *Oxford Review of Education 10*(1) 45-68.

Gray, J., Jesson, D., &Jones, B. (1986). The search for a fairer way of comparing schools' examination results. *Research Papers in Education, 1*(2) 91-122.

Haertel, E. H., (2013) *Reliability and validity of inferences about teachers based on student test scores*. ETS Research and Development Princeton, NJ 08541-0001

Hanushek, E. (1971). Teacher characteristics and gains in student achievement: Estimation using micro-data. *American Economic Review, 61*(2), 280-288.

Hanushek, E. (1972). *Education and race.* Lexington, MA: D.C. Heath and Company.

Hanushek, E., & Rivkin, S. G. (2010). Generalizations about using value-added measures of teacher quality. *American Economic Review, 100*(2), 267-271.

Hershberg, T. (2004) *Operation public education*. Retrieved from

http://www.cgp. upenn.edu/ope_value.html last access: September 11, 2008.

Hershberg, T. (2008). *An overview of value-added assessment.* Web: http://www.cgp.upenn.edu/pdf/Value-Added%20for%20Web.pdf   last access September 11, 2008.

Ishii, J., & Rivkin, S. G. (2009). Impediments to the estimation of teacher value added. *Education Finance and Policy, 4*(4), 520-536.

Kain, J. F. (1998). *The impact of individual teachers and peers on individual student achievement.* New York: paper presented at the Association for Public Policy Analysis and Management 20th Annual Reseach Conference.

Kane, T. J., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research Working Paper 14607.*

Kurtz, M. D. (2018). Value-added and student growth percentile models: What drives differences in estimated classroom effects?. *Statistics and Public Policy, 5*(1), 1-8.

Lord, F. M. (1969). Statistical adjustments when comparing preexisting groups. *Psychological Bulletin, 72* (5), 336-337.

Loeb, S., Soland, J., & Fox, L. (2014). Is a good teacher a good teacher for all? Comparing value-added of teachers with their english learners and Non-English learners. *Educational Evaluation and Pol- icy Analysis*, 36, 457–475. [1,4]

Maxwell, W. S. (1987). Teachers' attitudes towards disruptive behaviour in secondary schools. *Educational Review, 39*(3), 203-216.

McCaffrey, D. F., & Hamilton, L. S. (2007). *Value-added assessment in practice: Lessons from the Pennsylvania value-added assessment system pilot project*. RAND Corporation.

McCaffrey, D. F., Lockwood, J. R., Koretz, D. M., & Hamilton, L. S. (2003). *Evaluating value-added models for teacher accountability.* RAND Corporation.

MoNE. (2003). *TIMSS 1999 Üçüncü Uluslar Arası Matematik ve Fen Bilgisi Çalışması Ulusal Rapor*. Ankara: T.C Millî Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

MoNE. (2010). *PISA 2006 Projesi Ulusal Nihaî Rapor.* Ankara: T.C Millî Eğitim Bakanlığı Eğitimi Araştırma ve Geliştirme Dairesi Başkanlığı.

Murnane, R. J. (1975). *The impact of school resources on the learning of inner city children.* Cambridge, MA: Ballinger Publishing Co.

Rivers, J. C. (1999). *The impact of teacher effect on student Math competency achievements.* Knoxville: University of Tennessee.

Rivkin, S. G., Hanushek, E. A. & Kain, J. F. (2000). *Teachers, schools, and academic achievement.* Cambridge, MA: National Bureau of Economic Research, NBER Working Paper #W6691.

Rogosa, D. R. (1995). Myths and methods: "Myths about longitudinal research" plus supplemental questions. In J. M. Gottman (Ed.), *The analysis of change* (pp. 3-66). Mahwah, NJ: Lawrence Erlbaum Associates.

Rowan, B., Correnti, R., & Miller, R. J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the "prospects" study of elementary schools. *Teachers College Record, 104*, 1525-1567.

Sanders, W. L. (1997). *Graphical summary of educational findings from the Tennessee Value-Added Assessment System (TVAAS).* Knoxville: University of Tennessee Value-Added Research and Assessment Center.

Sanders, W. L., & Horn, S. P. (1994). The Tennessee Value-Added Assessment System (TVAAS): Mixed-model methodology in educational assessment. *Journal of Personnel Evaluation in Education, 8*, 299-311.

Sanders, W., & Horn, S. (1998). Research findings from the Tennessee Value-Added Assessment System (TVAAS) Database: Implications for educational evaluation and research. *Journal of Personnel Evaluation in Education, 12* (3), 247-256.

Sanders, W., & Rivers, J. C. (1996). *Cumulative and Residual Effects of Teachers on Future Student Academic Achievement.* Knoxville, TN: University of Tennessee Value-Added Research Center.

Sanders, W., Saxton, A., & Horn, B. (1997). The Tennessee Value-Added Assessment System: A Quantitative Outcomes-based Approach to Educational Assessment. In M. J (Dü.), *Grading teachers, grading schools: Is student achievement a valid evaluational measure?* (pp. 137-162). Thousand Oaks, CA: Corwin Press Inc.

School Directors' Handbook. (2008). Evergreen Freedom Foundation: 11 09, 2008 Retrieved from

http://www.effwa.org/pdfs/Value-Added.pdf

Tekwe, C. D., Carter, R. L., Ma, C.-X., Algina, J., Lucas, Roth, J., Ariet M., Fisher, T., & Resnick, M. B. (2004). An empirical comparison of statistical models for value-added assessment of school performance. *Journal of Educational and Behavioral Statistics, 29,* 11-36.

Thum, Y. M. (2001). Measuring progress towards a goal: Estimating teacher productivity using a multivariate multilevel model for value-added analysis. *Sociological Methods and Research, 32*, 153-207.

US Department of Education (2009). *Race to the top fund: Final rule.* Web: http://edocket.access.gpo.gov/2009/pdf/E9-27426.pdf last access August 21, 2010.

Webster, W., & Medro, R. (1997). The Dallas Value-Added Accountability System. In J. Millman (Dü.), *Grading teachers, grading schools: Is student achievement a valid evaluation measure?* (pp. 81-99). Thousand Oaks, CA: Corwin Press, Inc.

Wikipedia. (2008). No Child Left Behind. 11 09, 2008 retrieved from http://en.wikipedia.org/wiki/NCLB

Willms, J. D. (1987). Difference between Scottish educational authorities in their educational attainment. *Oxford Review of Education 13*(2), 211-232.

Woodhourse, G., & Goldstein, P. (1988). Educational performance indicators and LEA league tables. *Oxford Review of Education, 14*(3), 301-320.

Wright, S. P., Horn, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personnel Evaluation in Education, 11* (1), 57-67.

İlköğretimde Öğrenci Başarısında Okulun Etkisinin Katma Değer Belirleme Yöntemiyle İncelenmesi (Ankara İli Örneği)

Atıf:

Özet

Problem Durumu: Türkiye'de okulların başarısı genellikle öğrencilerin bir öğrenim düzeyinden üst öğretim düzeyine geçişte girdikleri sınavlarda gösterdikleri başarıya göre belirlenmektedir. Bu haliyle okul, öğretmen, öğrenci başarılarının doğru ve nesnel olarak belirlendiğini söylemek güçtür. Oysa okul, öğretmen, öğrenci başarılarına etki eden değişik faktörler bulunmaktadır. Türkiye'de gerek ilköğretim düzeyinde SBS gerekse ortaöğretim düzeyinde ÖSS sınavları üzerine tartışmalar süregelmektedir. Siyasi iktidarlar eğitim programlarına bu sınavları kaldıracakları ya da alternatif sınavlar uygulayacakları yönünde görüşler koymaktadır. ABD, İngiltere, Kanada ve Avustralya gibi gelişmiş ülkelerde okul, öğretmen, öğrenci başarılarının belirlenmesinde Türkiye'de yapılan sınavlara alternatif olan KDB yaklaşımları kullanılmaktadır. Bu ülkelerde öğrencilerin yıllık gelişimlerinin tespitinde ve bu gelişime okul, öğretmenin katkılarının hesaplanmasında SBS'ye benzer sınavlar kullanılmaktadır. Ancak okul, öğretmen değerlendirmelerinde ortalama başarı puanları yerine KDB yaklaşımları kullanılmaktadır. Benzer yaklaşım MEB tarafından Türkiye'de de uygulanabilir. Çünkü, KDB okulları adil, akla yatkın ve nesnel bir biçimde değerlendirmeyi sağlayan istatistiksel bir yöntemdir.

KDB yönteminde, öğrenci okul sistemine girdiği andan itibaren öğretmen veya okulun katkısıyla ilgilendiği için, ekonomi literatüründen ödünç alınan katma değer terimiyle ilişkilendirilmiştir (McCaffrey, Lockwood, Koretz, & Hamilton, 2003). Katma değer eğitim literatüründe okulun veya öğretmenin, eğitimin birikimli

etkisinden veya öğrencinin kendi sosyo-ekonomik durumu ve çevresinden arınmış olarak, öğrenci gelişimindeki özgün katkıları anlamında kullanılmaktadır. Eğitimde katma değer gerek akademik çevrelerde, gerekse eğitim politikalarını şekillendiren eğitim çalışanları ve siyasetçiler arasında rağbet gören bir kavramdır. Birleşik Devletlerdeki eğitimde reform çabalarının en önemli parçalarından biri de okulların, öğretmenlerin ve öğrencilerin performanslarının standardize sınavlarla ölçülmesidir. Hemen hemen bütün eyaletlerde eğitim alanında kullanılan takip sistemleri (accountability system) o ya da bu şekilde standardize sınavların sonuçlarını kullanır. Özellikle bütün öğrencilerin akademik başarımının asgari bir seviyenin üstünde olmasını amaçlayan "Hiçbir Çocuk Geride Kalmasın" (No Child Left Behind – NCLB) yasasının yürürlüğe girdiği 2001 yılından itibaren standardize sınav tabanlı takip sistemleri, Amerikan eğitim politikasının merkezine oturmuştur (Wikipedia, 2008). Devre devreye karşılaştırma ve Yıllık Yeterli Gelişim yöntemlerinin takip sistemlerindeki hızla artan baskınlığına karşın Katma Değer Belirlemesi yöntemi hem eğitim politikacıları hem de akademik çevrelerde giderek artan bir ilgi çekmeye başlamıştır. Örneğin Ohio, Pennsylvania ve Tennessee eyaletlerinde kabul edilen bazı yasalar, eğitim bölgesi yöneticilerinin, okul müdürlerinin ve öğretmenlerin, KDBdeki başarılarına göre terfi ve maaş anlamında ödüllendirilmesine veya cezalandırılmasına yönelik hükümler içermektedir (School Directors' Handbook, 2008).

Araştırmanın Amacı: Bu araştırmanın amacı, Ankara İl Büyükşehir Belediyesi sınırları içindeki ilköğretim okullarının, 7. sınıf öğrencilerinin katkısını, yıl sonu Seviye Belirleme Sınavı sonuçlarını kullanarak değerlendirmektir. Bu amaç doğrultusunda aşağıdaki sorulara yanıt aranmıştır. (1) İlköğretim okulları, öğrencilerine ne kadar katma değer sağlamıştır? (2) Öğrenci başarısı ile öğrenci gelişimi açısından ilköğretim okullarının sıralamaları tutarlı mıdır?

Araştırmanın Yöntemi: Nedensel karşılaştırma çalışma çerçevesine göre yürütülen bu çalışmanın popülasyonu, Ankara ilindeki ilköğretim okullarından oluşurken, örneklemi 24 ilköğretim okulunu kapsamaktadır. Bu okullardaki öğrencilerden 2007-2008 ve 2008-2009 eğitim yıllarında aynı okula giden öğrenciler analize dâhil edilmiştir. Örneklemdeki okulların öğrenci gelişimi üzerindeki katma değer etkileri, basit bir sabit etki modeli kullanılarak değerlendirilmiştir. Bununla birlikte okulların ortalama öğrenci başarıları ve okul katkılarının sıralaması arasında istatistiksel bir ilişki olup olmadığı da incelenmiştir.

Araştırmanın Bulguları: Bu çalışmanın sonuçları, okulların sıralamasında öğrenci gelişimine kattığı katma değer ve öğrenci başarısına göre sıralamalar arasında önemli bir tutarsızlık olduğunu göstermektedir. Bu, Türkiye'deki okulların bugüne kadar performansın değerlendirilmesinde sıklıkla kullanılan yöntemin eksikliğini göstermektedir. Ayrıca, sonuçlar, okulların öğrenci gelişimi üzerindeki katma değer etkilerinin konudan konuya büyük ölçüde farklı olduğunu göstermektedir. Bu araştırmada okulun öğrenci başarısına olan katkısının okuldaki öğrencilerin geçmiş akademik başarılarından ve bulundukları bölgenin sosyoekonomik düzeyinden bağımsız olduğu saptanmıştır. Yüksek akademik başarıya ve sosyoekonomik düzeye sahip öğrencilerden oluşan okulların her zaman bu öğrencilere yüksek katkı sağlamadıkları belirlenmiştir. Ayrıca, Başarılı olarak bilinen bazı okulların aslında

öğrencilerine düşük başarılı olarak düşünülen okullar kadar katkı sağlayamadıkları saptanmıştır. Tersine düşük akademik başarıya ve sosyoekonomik düzeye sahip okulların öğrencilerine yüksek katkı sağladığı durumlar da vardır. Öte yandan, öğrenci puanlarındaki değişimin önceki akademik başarıya ve sosyoekonomik düzeye bağlı olmadığı saptanmıştır. Öğrencilerin akademik başarısı düşük olsa da iyi bir eğitimin öğrenciye büyük katkı sağladığı görülmüştür.

Araştırmanın Sonuç ve Önerileri: Bu araştırmada yapılan çözümlemelerle başarılı ve daha az başarılı okullar KDB yöntemiyle belirlenmiştir. Daha az başarılı okullar yüksek katkı sağlayan okulların birikimlerinden faydalanmalıdırlar. Üst düzey il ve ilçe Milli Eğitim Müdürlüklerinin koordinasyonuyla öğretmenlerin bilgi alışverişinde bulunmaları sağlanabilir. Öğrencilerine daha fazla katkı sağlayan öğretmenlerin sahip olduğu olanakların olumlu etkisi ve bu öğretmenlerin deneyimlerinden diğer öğretmenlerin de faydalanması sağlanabilir. Bu araştırmanın, yıllar boyunca daha fazla verinin ortaya çıkmasıyla okulların adil bir değerlendirmesine yol açması beklenmektedir. Okul öğrencilerinin ortalama başarısı Türkiye'deki okulların performansını değerlendirmek için sıklıkla kullanılan yöntemdir. Bu çalışma, Türk okullarının değerlendirilme biçiminin sorunlu olduğunu ve okulların etkilerinin değerlendirilmesinde katma değer yöntemlerinin dikkate alınması gerektiğine işaret etmektedir. Ayrıca, bu çalışma Türkiye'de gerçekleştirilen ilk katma değer değerlendirmesi çalışması olması sebebiyle önem arz etmektedir. İleriki çalışmalarda çözümleme hem tüm Türkiye'yi içerecek şekilde hem de daha uzun yılları ve sınavları içerecek şekilde genişletilmelidir. KDB yöntemiyle hesaplanan katkı miktarına göre belirli bir süre üstün katma değer sağlayan öğretmenler ve okullar ödüllendirilebilir. Benzer şekilde KDB uygulamaları için uzun yıllara ait sınav sonuçlarının bulunduğu, öğrencilerin okul, sınıf ve öğretmen değişikliklerini de içeren geniş bir veritabanına ihtiyaç vardır. Okul ve öğretmenlerin adil bir şekilde değerlendirilebilmesi için uzun yıllar devam edecek standart bir sınav sisteminin oluşturulması gerekir. Sınavlar öğrencilere ek olarak öğretmen ve okulları değerlendirmek amacıyla da kullanılmalıdır.

Anahtar Kavramlar: Okul performansının değerlendirilmesi, katma değer belirlenmesi, basit sabit etki modeli, seviye belirleme sınavı.

APPENDIX

Table 6

Municiaplities and Names of Schools in the Study

Region	School Name
Çankaya	Yasemin Karakaya İlköğretim Okulu
	Turhan Feyzioğlu İlköğretim Okulu
	Mohaç İlköğretim Okulu
Altındağ	Cebeci İlköğretim Okulu
	Atıfbey İlköğretim Okulu
	Polis Amca İlköğretim Okulu
Yenimahalle	Türkkonut İlköğretim Okulu
	Refika Aksoy İlköğretim Okulu
Mamak	29 Ekim İlköğretim Okulu
	Şehitlik İlköğretim Okulu
	Fatih Sultan Mehmet İlköğretim Okulu
Gölbaşı	T.E.K İlköğretim Okulu
	Atatürk İlköğretim Okulu
Keçiören	Çizmeci İlköğretim Okulu
	Atapark İlköğretim Okulu
	Şenlik İlköğretim Okulu
Sincan	Andiçen İlköğretim Okulu
	Burak Reis İlköğretim Okulu
	Özkent Akbilek İlköğretim Okulu
Etimesgut	Hasan Ali Yücel İlköğretim Okulu
	Atatürk İlköğretim Okulu Layika Akbilek İlköğretim Okulu

Who is the Most Effective Agent When Giving Indirect Written Corrective Feedback?

Hanife Bensen BOSTANCI[1], Fatma SENGUL[2]

ARTICLE INFO	ABSTRACT
	Purpose: Who is the most effective agent when giving indirect written corrective feedback (IWCF) to English as a foreign language (EFL)? The answer is ambiguous, and factors such as gender have been neglected. For these reasons, this study attempts to reveal the most effective agent when giving IWCF and seeks to highlight the impact of gender when receiving IWCF from different agents. **Method:** A quasi-experimental study was carried out in which the participants were three classes of EFL learners studying at a private university's preparatory school. One of the classes was named class A, which only recieved instructor IWCF, another class B, which only recieved peer IWCF, and the last class C, which only recieved collaborative IWCF for a five-week period.

Each group produced five written texts regarding the same topic each week at the same time. The data, or the participants' texts, were analyzed quantitatively. **Findings:** It was revealed that class C—who received only collaborative IWCF—significantly improved their writing skills compared to the other classes that received teacher and peer IWCF. In terms of gender, it was revealed that the male participants performed better than the female participants in class A, and the female participants in classes B and C produced better written texts compared to the male participants. **Implications for Research and Practice:** Pre-service and in-service teachers should provide IWCF to their EFL learners collaboratively, and they should consider the gender factor. It is suggested that future research focuses on other factors (i.e. age, proficiency). It is also suggested that researchers focus on the other type of feedback, namely direct written corrective feedback.

---

[1] Corresponding Author: Near East University, North Cyprus. e-mail: hanife.bensen@neu.edu.tr., ORCID: https://orcid.org/0000-0001-5451-8698
[2] Near East University, North Cyprus. e-mail: fatmasengul3@gmail.com, ORCID: https://orcid.org/0000-0002-9925-6449

Introduction

Written errors are unavoidable in our first language (L1), let alone a foreign/second language (L2). English as a foreign language (EFL) learners commit errors, and as teachers (instructors), we tend to correct our students in some way to help them develop their writing skills. Giving indirect written corrective feedback (IWCF) is undeniably beneficial for EFL students to improve their writing skills. It is through IWCF that learners see and work on their errors. However, it is still unknown whether the agent through whom learners receive constructive IWCF and thereby produce better writing matters. In addition, to the best of our knowledge, no study has indicated the impact of gender when receiving IWCF in the Turkish Cypriot context.

As the feedback process gains a great deal of importance in the worldwide context, English Language Teaching (ELT) instructors in the Turkish Cypriot context have started to pay more attention to the feedback process. In this context, very few studies have focused on the IWCF process (Bensen & Cavusoglu, 2017). In contrast, a few researchers have focused on the effects of the agent factor in the feedback process (Marzban & Shrjami, 2014; Mowlaie, 2014). As a result, instructors are unaware of the importance and effects of the agent in the IWCF process (Bensen & Cavusoglu, 2017).

Particularly in the Turkish Cypriot context, instructors pay little or no attention to providing effective and constructive feedback to their students (Bensen & Cavusoglu, 2017). In general, EFL students receive traditional teacher feedback, where direct corrections (see literature review section) are provided by the instructor. Nowadays, some institutions have started to provide IWCF to their EFL students by using their departments' own writing correction codes. Additionally, some have started to provide peer feedback, and a few have started to provide collaborative feedback. In order to universalize agents during the feedback process, we need to pinpoint the effects on EFL students' writing performance and to take gender into account when doing so. In short, more focus should be given to the issues and dilemmas related with the feedback process in the Turkish Cypriot context.

Currently, writing is the prominent skill of language-learning instruction for evaluating academic achievement in the education sector. Students spend more time studying their writing skills in order to have full fluency, and thus they learn to act as a decision maker in their writing process during their language learning classes (Jahin, 2012). As a result, a good command of writing skills is crucial in language learning. In terms of teaching writing, not many changes and advancement have been made, particularly when giving indirect written corrective feedback (IWCF) to students' errors (Marzban & Sarjami 2014).

Writing plays a key role in the cognition process, as it leads, composes, and communicates one's opinions (Yugandhar, 2015). In fact, a great number of researchers believe that writing is one of the most difficult skills, both to be learned and to be taught (Su, 2011). One important element of writing and writing improvement is errors. Errors can be viewed as an indispensable part of the foreign language-learning process; they are unavoidable for instructors, who are the first agents that reply to students' inappropriate language productions (Li & He, 2017). Instructors "find these

as guiding stars reflecting upon the learning process of learners", that is, "the teacher today is conscious of the fact that errors reflect how much the learners have learned and guide the teachers what next to focus on in language teaching" (Raja, Albasher, & Farid 2016, p. 5). For these reasons, error correction is viewed as inevitable in the language-learning process (Raja, Albasher, & Farid 2016). They indicate to teachers the areas that they need to focus on in order to improve their students' language proficiency.

According to Zaman and Azad (2012), feedback is a major element of the foreign language writing pedagogy; students wait for feedback on their work, having ideas about their strong and weak points**.** Therefore, providing students with constructive and effective feedback will foster their learning. The term *written corrective feedback* (WCF) can be identified as a way of "accommodating the new knowledge into the existing knowledge and prompts the learners to stick the learnt stuff in their long term memory" (Maleki & Eslami, 2013, p. 1255). Two distinctive types of WCF have been introduced into the literature: direct and indirect WCF. The process of *direct written corrective feedback* (DWCF) can be defined as a language tool that is used to help students see their errors and the corrected version of their errors; linguistic forms are provided by the agent. On the other hand, *indirect written corrective feedback* (IWCF) can be defined as a clue that points out to the student that an error has been committed (Hosseiny, 2014). Furthermore, IWCF "indicates in some way that an error exists but does not provide the correction, thus leaving it to the student to find" (Eslami, 2014, p. 446). In this process, the students are provided codes, i.e. WW (meaning 'wrong word'), underlining, circling, etc. that highlight that an error exists. Guenette (2012) emphasized that "for linguistic notions that the learners have been exposed to or are expected to know, it makes sense to provide indirect corrections with or without brief explanations or simply reminders to consult their notes or other resources". Still, "for features that are clearly outside the realm of the learners or that have not been the subject of instruction, direct corrections can be used, or errors can simply be ignored" (pp. 121-122). Many researchers revealed that IWCF is efficient in developing students' language and writing abilities (Alhumidi & Uba 2016, p. 366).

Information gathered from surveys has indicated that IWCF is the most common type of feedback strategy in which students and instructors engage, and prominent distinctions exist among the learners' predilections for direct and indirect written corrective feedback (Li & He, 2017, p. 71). A great deal of research regarding the impact of DWCF and IWCF on grammatical accuracy advancement not only concluded that IWCF was more effective than DWCF in promoting grammatical accuracy, but also found that IWCF "led to either greater or similar levels of accuracy over time" (Phiewma & Padgate, 2017, pp. 1-2). Another research finding showed that students at higher proficiency levels prefer to receive IWCF, while students at lower proficiency levels prefer to receive DWCF (Zareil & Rahnama, 2013, p. 10).

With regard to gender, different findings have been revealed. In light of Li and He's (2017) research findings, it could be said that "gender difference significantly influences learners' preference" during the feedback process (p. 63). In contrast, some studies indicated no significant difference between the preferences of female and male

EFL learners regarding the feedback process (Kahraman & Yalvac, 2015, p. 78). Moreover, no prominent distinction has been found between male and female students with regard to their predilections for WCF, apart from their preference of the requirement of error correction and the 'no corrective feedback' option. In other words, female and male students showed significant differences in evaluating requirement of error correction "and also choosing no corrective feedback as a viable option" (Khorshi & Rassaei, 2013, p. 71). Furthermore, the research findings of Khorshidi and Rassaei (2013) indicated that 88 of the students in their study stated that teacher WCF was more beneficial, and 51 of the learners stated that peer WCF was beneficial—that is, students preferred teacher WCF to peer WCF. On the other hand, if the gender variable is taken into consideration, no statistically-significant difference was found between male and female students in delivering the specific agents of peer and teacher WCF (Khorshidi & Rassaei, 2013, p. 80).

The agent who gives feedback is also important. Azevedo et al. (2012) mentioned that "pedagogical agents can adequately and correctly detect, track, model, and foster learners' self-regulatory processes" (p. 212). Additionally, the Business Dictionary (2016) defines the term 'factor' as "a constituent or element that brings about certain effects or results, or indicates a specific multiple, number, or quantity" (p. 1). Thus, in this study, the term *agent factor* refers to people such as a teacher or peers who are the source of feedback that the students receive; this study examines their effects on the IWCF process.

One of the agents who gives feedback is the teacher. As Marzban and Sarjami (2014) pointed out, "there is no doubt that teacher written feedback plays an essential role in English writing classes" (p. 293). Thus, as can be understood from the name, *teacher feedback* is the process through which the teacher provides feedback to the learners.

Another agent is the peers in the class. Bijami, Kashef and Nejad (2013) defined the term *peer feedback* as the use of learners as sources of information and interactants for each other in such a way that learners assume roles and responsibilities normally taken on by a formally trained teacher, tutor, or editor in commenting on and critiquing each other's drafts in both written and oral formats in the process of writing. (p. 93)

Peer feedback refers to the process where students provide feedback to each other.

A more recent agent is collaboration. Barnawi (2010) defines the term *collaborative feedback* as the "collaboration between students and students or students and teachers who are engaged in the act of explaining, arguing, and negotiating their ideas with their peers" (pp. 211-212). Indeed, collaborative feedback is the process in which students and teachers collaborate with each other in order to discover the errors; they share their knowledge about these errors and participate in the learning process together.

Not many studies have explicitly pointed out which specific agent helps develop writing skills, nor which gender reaps more benefits when receiving IWCF from a

particular agent (Sengul & Bostanci, 2018). For this reason, the aim of this study is to explore and reveal the most effective agent in giving IWCF to EFL students studying at a preparatory school in North Cyprus. In addition, this study aims to reveal the effects of gender on receiving IWCF from the teacher, from peers, and collaboratively.

To be able to address this topic, the following research questions are posed:

1. Which agent improves English as a foreign language (EFL) learners' writing skills the most when indirect written corrective feedback is employed?
2. What is the effect of gender when receiving indirect written corrective feedback from different agents?

Research Hypothesis

1. When the proficiency level of the learners and this specific context are taken into account, the EFL learners will improve their writing performance more when receiving instructor IWCF.
2. Female students will improve their writing skills in all three groups more than the male students.

Method

Research Design

This study employed a quasi-experimental research design in which three classes of EFL students received IWCF from different agents: teacher, peer, and collaborative. The design is quasi-experimental in the sense that the participants were not selected by the researchers. In quasi-experimental designs, the samples are not randomly assigned (Cook & Campbell, 1979). The class that received IWCF from their teacher acted as the control group, while the two other groups (peer and collaborative) acted as the experimental groups.

Research Sample

Forty-eight EFL learners constituted the sample of this study. Three classes were selected, each class consisting of 16 EFL learners. All the participants of this study were preparatory school students, with an elementary proficiency level in English, studying at a private university in North Cyprus. The research participants' ethnic origins were Turkish, Turkish Cypriot, Arabic, and Kurdish, and all participants were over the age of 18.

Research Instrument and Procedure

To be able to determine the answers to the aforementioned research questions, five written text scores of the participants were compared, both within the group and with the other groups. In addition, the written text scores were compared again in terms of gender, both within the same group and with the other groups.

Three classes of EFL students were selected. These three classes followed the exact same writing syllabus with different instructors for one semester (16 weeks in total: 14 weeks of instruction, a midterm, and final exam week). However, only five of these weeks were used to carry out this study. At the end of each of the five weeks, the students of each group produced a written text. These written texts were selected by the institution before classes commenced. The topics of the written texts were as follows: week 1: Introducing myself; week 2: My everyday routines; week 3: Advertisement; week 4: My bedroom; and week 5: My favorite restaurant.

Initially, the three classes were named class A, B, and C. Class A received instructor IWCF, class B received peer IWCF, and class C received collaborative IWCF. The instructor in class A, the peers in class B, and the learners and instructor in class C were all trained on how to give IWCF before the classes commenced, in order for the IWCF agents to follow the same process. The agents employed the written code criteria and assessment criteria of the preparatory school when correcting errors. Every week on the same day at the same time, the EFL students of each group were expected to write a composition.

Validity and Reliability

In terms of reliability, the five written text tasks were scored out of ten in each group. Interrater reliability was employed to ensure that the assessment and evaluation of the participants were significant. The researchers of this study, along with the lecturers of the groups, discussed and scored each paper. A Pearson product-moment correlation coefficient (Pearson's *r*) was computed to assess the relationship between the instructors' and the researchers' assessments given to the five written texts produced by the students of each of the three classes. A Pearson's *r*, also known as a Pearson product-moment coefficient, was employed, as it is "used with variables that have a curvilinear relationship, the resulting correlation is an *underestimate* of the true relationship between these variables" (Ravid, 2011, p. 119). This procedure was carried out to confirm the reliability of the lecturers' assessments.

A strong relationship appeared between the two sets of results, as the Pearson *r* is very close to one (see Appendix) (Kahn, 2010; Ravid, 2011). This means that changes in one variable are strongly correlated with changes in the second variable. The 2-tailed significance tests show that the variables positively correlate and that the relationship is statistically significant. For this reason, it could be concluded that the assessments of the texts by the instructors were reliable.

Before this research was conducted, a consent form was designed, and an ethical review application was created for the study. As required by the academic research etiquette, upon approval from the Ethics Committee of the Institute of Education Sciences of Near East University, the ethical clearance letter, the information sheet for participants, and the consent form were shared with the participants of this study. The participants were also notified that the data collected from this study would not be used for any purpose other than for analysis. Confidentiality of all data was assured.

Data Analysis

Data were analyzed quantitatively through the written works of the samples. The scores received from each task were entered into the Statistical Package for Social Sciences (SPSS) software program version 20 to be analyzed. Descriptive statistics were employed to find out the effects of the aforementioned agents on the writing performances of the samples. Then, each group's written texts were compared among each other, adopting an Analysis of Variance test (ANOVA). Finally, to be able to determine gender differences, an independent samples t-test was conducted. The mean scores and standard deviation of the participants were then entered into a table (see Table 1).

Results

In order to answer the research questions, each class (class A, teacher IWCF; class B, peer IWCF; class C, collaborative IWCF) completed five writing tasks in which they wrote a composition and received IWCF.

The Most Effective Agent. In order to answer whether the agent factor affects the students' performance, and which agent improves EFL students' writing skills the most when employing IWCF, the participants' five written texts were scored out of ten and entered into SPSS.

Table 1

The Most Effective Agent

Tasks	N	Teacher		Peer		Collaborative	
		M	SD	M	SD	M	SD
1	16	2.38	1.74	1.38	2.84	2.31	2.12
2	16	2.31	2.33	2.38	2.18	3.25	1.57
3	16	2.19	2.13	0.63	2.70	2.69	1.74
4	16	2.69	1.53	1.50	1.93	2.31	1.95
5	16	1.69	1.58	0.13	2.50	2.56	1.99
Valid N (listwise)	16						

Key: M - Mean Score   SD - Standard Deviation

Instructor IWCF. As can be seen in Table 1, the mean scores indicate that the samples of class A did not improve their writing skills during their first, second, and third tasks. They seemed to have performed better in their fourth task compared to their first, second, and third tasks. It could be seen that the participants' writing performance seemed to decrease during the fifth task.

Peer IWCF. Table 1 indicates that class B did not improve their writing skills in general. The findings show the performances of class B's samples' writings from task 1 to task 2 significantly improved, but rapidly decreased in their third task. Nevertheless, the samples increased their writing performance during their fourth

task. In comparison to their performance during the fourth task, there was a rapid decrease in their fifth task.

Collaborative IWCF. The findings related to class C indicate that collaborative IWCF helped the samples improve their writing performance (see Table 1). It was observed that class C's writing performance significantly improved from task 1 to task 5.

The Effect of Gender. To be able to reveal whether gender affects the participants' writing performances, participants' gender was compared among the groups.

Instructor IWCF. According to Table 1, the students who received IWCF from their instructor (class A) seemed not to have improved their writing skills in general. However, it was revealed that the male participants in class A performed better than the female participants (see Table 2).

The results of the written-text analysis of the female participants in class A show that the female participants performed better in their first task than the second and third tasks. Moreover, it was also observed that these participants also increased their writing performance in task 4, while their writing performances decreased in their final task. In the same way, the male samples in class A performed better in their second, third, and fourth tasks compared to their first task. Similar to the female participants, male participants' writing performances decreased in the fifth task. In short, the male students performed better than the female students during their second, third, and fourth tasks, while the female students performed better than the male samples in task 5. There was no difference between the performances of the female and male participants in group A in task 1.

Table 2

Gender and IWCF

Tasks	G	N	Teacher		Peer		Collaborative	
			M	SD	M	SD	M	SD
1	female	8	2.38	1.84	2.38	2.32	3.00	1.41
	male	8	2.38	1.76	0.38	3.11	1.63	2.56
2	female	8	1.13	2.53	3.00	1.92	3.63	.74
	male	8	3.50	1.41	1.75	2.37	2.88	2.10
3	female	8	1.50	2.07	0.50	2.67	3.38	.74
	male	8	2.88	2.10	0.75	2.91	2.00	2.20
4	female	8	2.63	1.40	1.50	2.07	2.88	1.24
	male	8	2.75	1.75	1.50	1.92	1.50	2.43
5	female	8	2.00	1.06	1.00	2.39	3.75	1.03
	male	8	1.38	1.99	-0.25	2.43	1.38	2.06
Valid N (listwise)		88						

Key: G: Gender   M: Mean Score   SD: Standard Deviation

Peer IWCF. According to the average grade of class B's samples, group B's female and male participants were believed to not have improved in their writing skills. Moreover, it was discovered that the female participants in group B performed better than the male participants.

It is evident from the results of the written-text analysis of the female participants in class B that they performed better in their second task in comparison to their first task, while their writing performance rapidly decreased in their third, fourth, and fifth tasks (see Table 2). Similarly, the results of the written-text analysis of the male participants in class B revealed that they performed better in their second, third, and fourth tasks compared to their first task. It was also revealed that the male participants in class B rapidly decreased their performance in the last task. As a result, the female participants performed better during tasks 1, 2, and 5 compared to the male participants, while the male participants performed slightly better than the female participants in task 3. Both female and male participants in this group received similar scores.

Collaborative IWCF. The findings with regard to class C revealed that the female participants performed better than the male participants in all tasks (see Table 2). The findings of the written-text analysis of the female participants in class C demonstrated that they performed better in their second and third tasks in comparison to their first task, while their writing performance decreased a little in their fourth task and rapidly increased in their fifth task. Meanwhile, the written-text analysis results of the male participants showed that they rapidly increased their writing performance in their second, third, and fourth tasks compared to their first task, while their writing performance decreased in their fifth task. As a result, the female participants performed better in the first, second, third, fourth, and fifth tasks compared to the male participants. In general, the female participants in class C performed better than the male participants in all tasks.

To sum up, the overall findings showed a significant difference between the performances of the female and male participants in classes A, B, and C, who all received IWCF from different agents. It was discovered that the female participants in classes B and C (peer and collaborative) performed better than the male participants, while the male participants in class A (teacher) performed better than the female participants.

Discussion and Conclusion

Drawing on our findings, we could say that—despite the fact that Alharbi (2016) found that teachers' WCF (instructor in this case) had positive effects on students' writing skills—the present study found that the samples in class A, who received IWCF from the instructor, did not improve their performances. Furthermore, class B, who received IWCF from their peers, also did not improve their writing skills. Contradictory to the findings of Yoon (2011)—who indicated that EFL students' performances increased after receiving peer feedback—no improvement was seen in

the writing performance of class B, who received IWCF from their peers. In contrast to the results of classes A and B, class C, who received IWCF collaboratively, showed that collaborative IWCF helped the participants improve their writing performance. This is in line with the findings of Motallebzadeh and Amirabadi's (2013) study, in which the EFL students who received collaborative feedback performed better than the other groups.

In brief, similar to Dang's (2016) findings, the participants who received collaborative IWCF performed better than the groups who received teacher and peer IWCF (classes A and B). This study revealed that the most effective agent for the improvement of EFL students' writing skills was when both the teacher and students gave IWCF. After collaborative IWCF, teacher IWCF was found to be second-most effective, and lastly peer IWCF. This finding contradicted the findings of Kahyalar and Yılmaz (2016), who revealed that the group receiving peer feedback performed better than the group receiving teacher feedback. These results may be due to the fact that the teacher was more acquainted with giving feedback, compared to the students on their own. Moreover, the students were elementary level in proficiency, so they may not have been sufficiently equipped in terms of language proficiency to give effective feedback to their peers in order to help them develop their writing performance.

The students who received IWCF from their instructor (class A) seemed to have not improved their writing skills in general. Moreover, it was shown that the male participants in class A performed better than the female participants. With regards to the average grade of class B's participants, it was noticed that class B's female and male participants did not improve in their writing skills. However, the results related to class C revealed that the female participants performed better than the male participants in all tasks. These results are not in line with the findings of Kahraman and Yalvaç (2015), who found no significant difference related to the gender of their participants when giving WCF.

Drawing on the findings, we could say that the most effective agent when giving IWCF to EFL students is collaboration—both the students and the teacher of the writing class give the student in question IWCF. EFL students' writing performances improve when they receive IWCF collaboratively. When the gender variable was taken into consideration, it was also revealed that female participants who received peer and collaborative IWCF outperformed the male participants, whereas the male students who received teacher IWCF outperformed the female participants. For this reason, we could conclude that gender has an effect on students' writing performances when feedback is given by different agents. Particularly in this specific context, where students are used to receiving no feedback or only feedback from their teachers, female students seem to act more confident and have more say in the writing process when they receive feedback from their peers and collaboratively.

Limitations

This research is limited to the Turkish Cypriot context. Consequently, the research results might be different in other contexts. Therefore, the results should not be generalized to the whole population, both in and out of North Cyprus. Likewise, this

research is limited to adult EFL learners; hence, the research results might be different if carried out with learners from different proficiency levels and age groups. The collected written texts have not provided any information regarding the most frequent error types; they have only focused on the errors committed in general. Moreover, even though the related literature has introduced two types of written corrective feedback, this study has only employed indirect written corrective feedback. In addition, this study only focuses on three possible agents (teacher, peer, collaborative) to correct errors; self-correction is only fostered through indirect error correction. Furthermore, the EFL sample is limited, with only sixteen participants in each group; as a result, the research findings might change if a study were to be carried out with a larger sample.

Implications for Research and Practice

As the findings indicated, there are significant differences between class A, class B, and class C participants' scores when receiving IWCF from different agents, and gender has an effect on the scores of the participants. It is recommended that institutions, teachers, and IWCF providers take these results into account, especially in the Turkish and Turkish Cypriot context, in which students are EFL learners. It is also suggested that pre-service and in-service instructors provide CIWCF to their learners and use more collaborative work and activities in their classrooms, in order to create a friendlier and more positive, supportive, and collaborative atmosphere in the language-learning environment. By including CIWCF in writing tasks, instructors will be better able to help students improve their writing skills.

Further research could be conducted to see if other variables—such as EFL students' years of English study—affect their writing development. Another study might focus on individual differences regarding the motivation or the performance of the participants during the IWCF process. Discovering the attitudes of EFL learners and their teachers towards IWCF could be the focus of another study.

References

Alharbi, H. S. (2016). Effects of teachers' written corrective feedback on Saudi EFL university students' writing achievements. *International Journal of Linguistics, 8*(5), 15-29.

Alhumidi, H. A., & Uba, S. Y. (2016). The effect of indirect written corrective feedback to Arabic language intermediate students' in Kuwait. *European Scientific Journal, 12*(28), 361-374.

Azevedo, R., Landis, R. S., Behnagh, R. F., Duffy, M., Trevors, G., Harley, J. M., . . . Hossain, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with meta-tutor. *Lecture Notes in Computer Science*, 212-221.

Barnawi, O. Z. (2010). Promoting noticing through collaborative feedback tasks in EFL college writing classrooms. *International Journal of Teaching and Learning in Higher Education*, 209-217.

Bijami, M., Kashef, S. H., & Nejad, M. S. (2013). Peer feedback in learning English writing: Advantages and disadvantages. *Journal of Studies in Education, 3*(4), 91-97.

Bensen, H., Cavusoglu, C. (2017). The impact of blended learning in an EFL writing course. LAP LAMBERT Academic Publishing. ISBN: 978-3-659-54027-1.

Business Dictionary. (n.d.). *Factor*. Retreived from http://www.businessdictionary.com/definition/factor.html

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues in field settings*. Boston, MA: Houghton Mifflin.

Dang, T. T. (2016). Vietnamese EFL students' perceptions of noticing-based collaborative feedback on their writing performance. *English Language Teaching, 9*(5), 141-153.

Eslami, E. (2014). The effects of direct and indirect corrective feedback techniques on EFL students' writing. *Social and Behavioral Sciences*, 445-452.

Guénette, D. (2012). The pedagogy of error correction: Surviving the written corrective feedback challenge. *TESL Canada Journal, 30*(1), 117-126.

Hosseiny, M. (2014). The role of direct and indirect written corrective feedback in improving Iranian EFL students' writing skills. *Social and Behavioral Sciences*, 668–674.

Jahin, J. H. (2012). The effect of peer reviewing on writing apprehension and essay writing ability of prospective EFL teachers. *Australian Journal of Teacher Education, 37(*11), 60-84.

Kahn, J. (2010). *Reporting statistics in APA style. American Psychological Association.* Washington, DC: Publication manual of the American Psychological Association (6th ed.).

Kahraman, A., & Yalvac, F. (2015). EFL Turkish university students' preferences about teacher feedback and its Importance. *Procedia - Social and Behavioral Sciences*, 73-80.

Kahyalar, E., & Yılmaz, F. (2016). Teachers' corrective feedback in writing classes: The impact of collaborating with a peer during the editing process on students' uptake and retention. *The Reading Matrix: An International Online Journal, 16*(1), 148-160.

Khorshidi, E., & Rassaei, E. (2013). The effects of learners' gender on their preferences for corrective feedback. *Journal of Studies in Learning and Teaching English, 1*(4), 71-83.

Li, H., & He, Q. (2017). Chinese secondary EFL learners' and teachers' preferences for types of written corrective feedback. *English Language Teaching, 10*(3), 63-73.

Maleki, A., & Eslami, E. (2013). The effects of written corrective feedback techniques on EFL students" control over grammatical construction of their written English. *Theory and Practice in Language Studies*, 1250-1257.

Marzban, A., & Sarjami, S. M. (2014). Collaborative negotiated feedback versus teacher written feedback: Impact on Iranian intermediate EFL learners' writing. *Theory and Practice in Language Studies, 4*(2), 293-302.

Motallebzadeh, K., & Amirabadi S. (2013). Online international feedback in second language writing: Through peer or tutor. *MJAL, 3*(2), 156-174.

Mowlaie, B. (2014). The effect of teacher feedback, peer feedback, and self revision on Iranian EFL learners' form and content revision. *English Language Teaching, 1*(1), 37-57.

Phiewma, W., & Padgate, W. (2017). The effects of teacher indirect feedback and collaborative revision activity on grammatical accuracy of Thai college students' writing. *Journal of Community Development Research (Humanities and Social Sciences), 2*(10), 1-10.

Raja, M. S., Albasher, K. B., & Farid, A. (2016). Error treatment in teaching English to EFL adult learners: A study in current English language teaching practices in native versus non-native divide context in Saudi Arabia. *Journal of Applied Linguistics and Language Research, 3*(5), 1-16.

Ravid, R. (2011). *Practical statistics for educators* (4th ed.). UK: Rowman & Littlefield Publishers.

Sato, M. (2013). Beliefs about peer interaction and peer corrective feedback: Efficacy of classroom intervention. *The Modern Language Journal*, 611-633.

Sengul, F., & Bostanci, H. B. (2018). *Indirect written corrective feedback in EFL*. LAP LAMBERT Academic Publishing. ISBN: 978-613-9-85797-5.

Su, Z. (2011). Peer feedback: A new approach to English writing instruction in a Chinese college setting. *Sino-US English Teaching, 8*(6), 364-368.

Yoon, S. Y. (2013). Students' reflection on feedback in L2 writing in blended learning. *Multimedia-Assisted Language Learning,* 235-261.

Yugandhar, D. K. (2015). Practicing teacher organized peer review to advance EFL students' writing skills. *International Journal on Studies in English Language and Literature, 3*(1), 25-29.

Zaman, M., & Azad, A. K. (2012). Feedback in EFL writing at tertiary level: Teachers' and learners' perceptions. *ASA University Review, 6*(1), 139-156.

Zareil, A. A., & Rahnama, M. (2013). The effect of written corrective feedback modes on EFL learners' grammatical and lexical writing accuracy: From perceptions to facts. *International Journal on Studies in English Language and Literature*, 1-14.

Appendix

Interrater Reliability Test Results

Teacher IWCF

Task 1

		Lecturer's Marks	Researcher's Marks
Lecturer's Mark	Pearson Correlation	1	.968**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Mark	Pearson Correlation	.968**	1
	Sig. (2-tailed)	.000	16
	N	16	

** Correlation is significant at the 0.01 level (2-tailed).

Task 2

		Lecturer's Marks	Researcher's Marks
Lecturer's Mark	Pearson Correlation	1	1.000**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Mark	Pearson Correlation	1.000**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 3

		Lecturer's Marks	Researcher's Marks
Lecturer's Mark	Pearson Correlation	1	.982**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Mark	Pearson Correlation	.982**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 4

		Lecturer's Marks	Researcher's Marks
Lecturer's Mark	Pearson Correlation	1	.966**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Mark	Pearson Correlation	.966**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 5

		Lecturer's Marks	Researcher's Marks
Lecturer's Mark	Pearson Correlation	1	.953**
	Sig. (2-tailed)		.000
	N	16	16
	Pearson Correlation	.953**	1
	Sig. (2-tailed)	.000	
	N	16	16
Researcher's Mark			

** Correlation is significant at the 0.01 level (2-tailed).

Peer IWCF

Task 1

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.992**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.992**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 2

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.994**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.994**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 3

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.983**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.983**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 4

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.980**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.980**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 5

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.954**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.954**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Collaborative IWCF

Task 1

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.971**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.971**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 2

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.977**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.977**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 3

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.945**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.945**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 4

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.974**
	Sig. (2-tailed)		.000
	N	16	16
Researcher's Marks	Pearson Correlation	.974**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Task 5

		Lecturer's Marks	Researcher's Marks
Lecturer's Marks	Pearson Correlation	1	.967**
	Sig. (2-tailed)		.000
Researcher's Marks	N	16	16
	Pearson Correlation	.967**	1
	Sig. (2-tailed)	.000	
	N	16	16

** Correlation is significant at the 0.01 level (2-tailed).

Dolaylı Yönden Yazılara Geri Dönüş Verirken En Etkili Kişi Kimdir?

Atıf:

Özet

Problem Durumu: İngilizce dilini yabancı dil olarak öğrenen öğrencilere dolaylı yoldan yazılı düzeltici geribildirim verilmesinde en etkili etken kişinin kim olduğu sorusunun cevabı halen belirsizliğini sürdürmekte ve cinsiyet gibi faktörlerin etkisi göz ardı edilmektedir.

Araştırmanın Amacı: Yukarıda belirtilen nedenlerden dolayı bu çalışma ile dolaylı yoldan yazılı düzeltici geribildirim verilmesinde en etkili etken kişiyi bulmak ve farklı etken kişilerden dolaylı yoldan yazılı düzeltici geribildirim alınmasında cinsiyet faktörünün etkisini bulmak amaçlanmaktadır.

Araştırmanın Yöntemi: Yarı deneysel yöntem izlenerek yapılan bu çalışmanın katılımcıları, özel bir ünivesitede hazırlık okulunda İngilizce dilini yabancı dil olarak öğrenen üç farklı sınıfın öğrencileridir. Beş haftalık bir süreç boyunca sadece öğretmen tarafından dolaylı yoldan yazılı düzeltici geribildirim alan sınıf, A sınıfı; sadece öğrencilerden dolaylı yoldan yazılı düzeltici geribildirim alan sınıf, B sınıfı; ve hem öğretmen hem öğrencilerin katılımı ile işbirlikçi bir şekilde dolaylı yoldan yazılı düzeltici geribildirim alan öğrencilerin bulundukları sınıf ise C sınıfı olarak adlandırılmıştır. Her grup, toplamda beş haftada beş farklı konulu yazılı metin olmak

üzere her hafta aynı zamanda yazılı metin üretmişlerdir. Öğrencilerin yazmış oldukları bu yazılı metinlerden elde edilen veriler, nicel olarak analiz edilmiştir.

Araştırma Sonuçları: Araştırma sonucuda, hem öğretmen hem de öğrencilerin işbirlikçi bir şekilde dolaylı yoldan yazılı düzeltici geribildirim verdikleri C sınıfındaki katılımcılar, diğer sınıflardaki katılımcılara önemli ölçüde kıyasla yazma yeteneklerini geliştirdikleri bulgusuna varılmıştır. Cinsiyet açısından ise A sınıfındaki erkek katılımcıların, aynı sınıftaki kadın katılımcılara oranla daha iyi bir performans göstermesine karşın, B ve C sınıfındaki kadın katılımcıların erkek katılımcılara oranla daha iyi performans sergilediği ortaya konmuştur.

Araştırmanın Sonuçları ve Önerileri: Öğretmen adayı ve öğretmenlerin, cinsiyet faktörünü göz önünde bulundurarak, İngilizceyi yabancı dil olarak öğrenmekte olan öğrencilere, hem öğretmen hem de öğrencilerin işbirlikçi bir şekilde dolaylı yoldan yazılı düzeltici geribildirim verdikleri bir yöntem izlemeleri sağlanmalıdır. Bu nedenle, ileride yapılacak olan araştırmaların yaş ve dil yeterlik gibi diğer faktörler üzerine yoğunlaşmaları önerilmektedir. Aynı zamanda araştırmacılara doğrudan yazılı düzeltici geribildirim gibi diğer geribildirim çeşitleri üzerine odaklanmaları önerilmektedir.

Anahtar Kelimeler: Cinsiyet, İngilizceyi yabancı dil olarak öğrenen, yazı gelişimi.

An Analysis of Day to Day Activities of a Sample of Primary School Principals in Ireland

Martin STYNES[1], Gerry MCNAMARA[2], Joe O'HARA[3]

ARTICLE INFO	ABSTRACT
	Purpose: There is a vast quantity of research into principalship, mainly concentrating on macro level theorising about concepts such as 'instructional' leadership, 'distributed' leadership and a myriad of other notions of the role. In contrast there is very little work done on the messy and demanding day to day and hour to hour work of school principals and the experience, knowledge and skills that this requires and the intense stresses and strains placed on school leaders. In Ireland, as elsewhere, primary school

principals meet the challenges of teaching, community leadership and on-site management in an era of continual change, in most cases with limited or no formal preparation. The rationale for this work is to balance research in the field by focusing on the micro tasks that make up the bulk of the principal's role and to examine how school leaders cope with the job and how they respond to it. **Research methods:** A diverse group of 31 primary school principals from schools across Ireland generated data from the self-observed minutae of researcher-driven diaries and from a colourful spectrum of personal reflections in follow-up semi-structured qualitative interviews. Coding in NVivo and the querying of emergent themes through conceptual frameworks provided detailed evidence of a myriad of daily activities and experiences. **Findings:** This paper offers an exploration, in narrative form and with supporting evidence, of principals' encounters with the constant minutiae of administration, dealing with the unexpected and interacting with staff. The daily practicalities of school governance and community leadership demand a considerable investment of time and personal interest and often little time for consideration of higher level macro theories of leadership. **Implications for Research and Practice**: As the boundaries between principals' professional and personal lives blur significantly in the narrative, the evidence supports a generally held understanding that life's journey as a school principal is demanding but worthwhile.

---

[1] Centre for Evaluation, Quality and Inspection (EQI) DCU Institute of Education, Ireland, e-mail: mstynes62@gmail.com ORCID: https://orcid.org/0000-0002-4924-2225
[2] Centre for Evaluation, Quality and Inspection (EQI) DCU Institute of Education, Ireland, e-mail: gerry.mcnamara@dcu.ie, ORCID: 0000-0002-9725-9304
[3] Corresponding Author: Centre for Evaluation, Quality and Inspection (EQI) DCU Institute of Education, Ireland, e-mail: joe.ohara@dcu.ie, ORCID: 0000-0003-1956-7640

Introduction

School leadership research has produced a remarkably vast number of studies but while the functioning of schools and the quality of leadership performance in them is studied commonly on a macro level, the microanalysis of lived leadership experiences with evidence of reflective capital is under-researched (Scott et al., 2004). With this in mind, we undertook a case study of the everyday work and life experiences of a group of Irish primary school principals. Our objective was to provide an original and critical analysis of the human experience of school leadership, as opposed to a showcasing of Irish practice or a generalised reasoning of the universal (Kuhn, 1996; Smith, 2009).

Academic research into the daily experiences of Irish primary principals was relatively thin on the ground and it offered only an ad-hoc range of perspectives. Morgan and Sugrue (2008) for example, who yielded significant information from attitudinal surveys, nonetheless concluded that 'we are still left very much in the dark as to what Principals actually do during their working day' (p.13). How they experienced such a working day and how they could voice this personally, gave rise to our particular curiosities.

It is significant that there is no mention of the principal in Coolahan's (1981) authoritative and seminal volume on the history of the Irish education system and that an absence of an historical perspective on the work of principals through the 20th century is abated only in works of biography and fiction. Today in Ireland, we know that 3,262 primary principals organise the education of 553,380 children (DES, 2016). A high proportion of small rural schools in parish-based settings gives rise to a situation where almost three quarters of Irish primary school principals teach full-time. Approximately one quarter, typically those in larger schools in urban settings, occupy full-time administrative posts. The roles of teaching and administrative principals differ significantly in their daily routines but in many respects, policy directives and the general responsibilities of administration, management and leadership apply equally to both.

The generations-old Irish language term 'príomhoide' is a predominant descriptor in common discourse. It translates as 'principal teacher' and as such, it signifies a heritage of instructional collegiality (LDS, 2007). The term 'school leader' on the other hand, while well established in the universal lexicon, is a relatively new arrival to ordinary conversation. It is actively promoted however in professional circles from different, if sometimes contrasting and politicised perspectives. Both the teachers' trade union The Irish National Teachers' Organisation (INTO, 2017) and the Department of Education and Skills (DES) through its various manifestations, organise regional and national meetings, short courses, events, conferences, mentoring and support networks to promote the leadership dimension of the principal's role (IPPN, 2017; PDST, 2017; CSL, 2017).

Under the Education Act (Ireland, 1998), a school's Board of Management has a statutory responsibility to provide an appropriate education for its pupils and to this end, a Board appoints a principal to provide leadership to the school community.

While a Board serves a four-year term of office, a principal is appointed in a permanent capacity and typically acts as a link from one Board to the next.

Using both statutory provision and research, Drea and O'Brien (2003) on behalf of the Irish Primary Principals' Network (IPPN), identified seven categories that encapsulated the primary principal's responsibilities. Only one, that of delivering and developing high quality teaching, was school-specific. The remaining six, including administration, building external relationships and other management specific duties, could reasonably be applied to the CEO of any business or industry.

While such findings demonstrate a significant advancement in the role and responsibility of the Irish principal, they also present somewhat of a dichotomy. While the current national school system, established from 1831 under Victorian structures of hierarchy and station, was intended purposefully to include formal management structures, at no time was it envisaged that the day-to-day manager should actually be the principal teacher (McDonald, 2008).

Another potential dichotomy emerges for Irish principals in a changing and increasingly secularised society where over 90% of primary schools operate under a Catholic ethos management authority (CPSMA, 2016). Potentially, principals face a dilemma in that their stated school leadership ethos is rooted in a sense of vocation and in working for a perceived higher purpose (Adair, 2000; Palmer, 1998; Ryan and Higginbottom, 2017). The practical implications of such an ideological underpinning may prove to be at odds with new forms of public management, given the globalised development of leadership theories and the scientific demands they bring to bear; particularly in relation to emergent forms of professionalisation within the work-force (The Teaching Council, 2016).

Establishing a position between objective and subjective polarities, balancing scales of opposites and intellectualising beyond the binary, provides researchers in school leadership with interesting challenges. Knowledge-work necessitates exploration beyond the influences of the mapped literature as we enter physically into communities of practice where management and leadership events actually happen, impact upon people and form part of life's realities (Potter et al., 2002; Addison, 2009; Fullan ,2014).

It must reasonably be acknowledged that schools are publicly funded and should therefore be accountable for some measurable return, but we would argue, in light of such escalating professionalisation and the associated philosophical redirectioning, that the provision of child-centered learning experiences and the development of fully socialised human beings are largely forgotten as the primary and central purposes of schooling in Ireland. Similarly, Taysum (2013) reflects a problematic dichotomy in the UK whereby the positivist influences of high-stakes testing and other measures of educational accountability are leapfrogged over the less tangible and more interpretive dimensions of education, such as values development and inclusiveness.

School leadership is a critical influencer and in Ireland, as elsewhere, the canons of instructional and transformative models for example have become accepted CDP

templates. Hall and Southworth (1997) and Fink (2005) however urge caution and call for sophisticated reflection on the practical implications of new and potentially superficial leadership practices. Similarly, Starrat (2005) and Saltman and Means (2017) criticises a dependence on leadership tool kits, lists of competencies and on one-size-fits-all checklists. Authors such as Thrupp and Wilmott (2003), Ryan (2012) and Anderson and Lopez (2017) paint the performativity paradigm vividly and warn of the dominance of managerialism in the education discourse. As management becomes the solution of our times they add, we are called upon not only to research a new perceived professionalisation of schools but more importantly perhaps, the influences that are brought to bear on the lives of those who work within them.

Sugrue (2003) traces the redefinition of the principal's role in Ireland to the 1970s and highlights a notable increase in administrative and managerial responsibilities since then. During the recent half century, Ireland's education policies have undoubtedly been influenced by the development of global leadership theory and by an international infatuation with higher performance standards (Hallinger, 2005). Gunter (2001) and Gunter and Fitzgerald (2103) trace a similar evolution in the UK where significant advancement is backed by a Government investment in school leadership that is reportedly the highest in the world (Day and Smethem (2009).

As in other developed countries, organisational theory and its associated lexicon have migrated steadily into educational discourse and practice (Glatter 2014; Hall et al, 2017; Niesche and Thomson, 2017) . The business paradigm shifts that optimise the performances of CEOs and that reap rich financial rewards on global markets prove also to have powerful influences on the child-centered traditions of many schools and classrooms. Notwithstanding such influences, Ireland's schooling system presents with significantly fewer established structures and with school settings that are, as yet, devoid of formalised leadership acculturation. We propose therefore that the Irish schools offer significant potential for the ethnographic exploration of the organic and contextualised grounding of school leadership practices.

The problem is therefore that while macro level research and theorising presents valuable perspectives on school principalship this needs to be balanced by a much better understanding of the day to day challenges of the job and how school leaders react and cope with the endless stream minor but important issues and tasks which take up the vast bulk of their time. That is the focus of this research

Method

Research Design

Notwithstanding a vast corpus of work on the mastery of the principalship, Troman (2006) describes the principal as a shadowy figure and Ribbins and Marland (1994) assert that we know surprisingly little about headteachers as people. What it means for individuals personally to engage with school leadership is rarely if ever asked (Gronn, 2009). Other notable outliers to the mastery literature are Sugrue (2005),

Tomlinson et al. (1999) and Kelchtermans et al. (2011) whose research on school leadership is concerned not only with the formal structure of the role but with the ways in which it is experienced by the people who cope with the daily vicissitudes and multi-directional realities that school-life presents.

Protocol in the world of such knowledge-work demands a philosophical underpinning or rationale where ontological assumptions, concerning the nature of existence and of reality, are made explicit as an assumed starting point in the research (Bakker, 2010). We chose to employ a perspective beyond the objective in our study of principals and instead we sought sense-making in social processes enactive of the environment, those beyond a single apprehensible reality (Weick, 1995). While we acknowledged the existence of important factors that were demonstrably authentic, factually based and tangible within the edifice of the schoolhouse, we contended that the essential understanding of what is ontic, or what is really real, was constructed not within the institution but in the minds of individuals and groups who populated the structure (Eidlin, 2010).

The most fundamental epistemological questions of whether and how we can know anything underpinned our approach to the gathering of data through the use of credible processes (Burgess et al., 2006). A case study approach aimed to construct knowledge about the experiences of principals through personal, subjective and unique interactions (Stake, 1995). Our objectives in the research were not determined by the proof of testable scientific theories. We moved instead, beyond a modernist perspective of totalising theory where the role of the principal might be clearly defined in policy and practice, to seek a post-modern rationalising; one that attempted to understand uniquely human experiences through an engagement in local narratives (Linstead, 2010).

The process of engaging individual principals in the study and of unearthing data of any quality was complicated by a number of factors. Principals operate in both public and private arenas, engaging in a fluid but constant process of decision-making (Janesick, 2010; O'Leary, 2005). Principals may be suspicious or even uncomfortable about providing information, particularly if they perceive any sense of intrusion, scrutiny or potential exposure (Gillham, 2000; Cohen et al, 2007). *Research Sample*

Research Sample

Given therefore, that the principal's office may reasonably be considered as a difficult site to access, an approach was chosen that incorporated multiple forms of data gathering. Case study, with a revelatory rationale, offered the potential for native point of view, naturalistic context and over-rapport. A sample group of 31 principals responded to on-line requests for research volunteers through professional membership sites and all 31 principals participated in all stages of the research thereafter (INTO, 2016; IPPN, 2016). The sample was self selecting but represented a good representation of school types in terms of gender, size and location. The study was built on the cornerstone of ethical assurances. Participants shared information on the understanding that crude reporting categories and micro-aggregation would

ensure complete confidentiality. By proving each completed diary and interview transcript with a code known only to the researchers and being extremely careful that none of the material used could identify a particular school of principal complete anonymity was assured.

A broad span of professional experience and a range of different school settings were represented in the study. The following table provides an overview.

Table 1

Participants' Demographics

Participants' Gender	Experience as a Teacher	Experience as Principal	School Setting	School Staffing	School Type
Male **11**	1-20 yrs **7**	1-10 yrs **19**	Urban **23**	10-20 **11**	Co-Education **22**
Female **20**	20+ yrs **24**	10-20 yrs **12**	Rural **8**	20+ staff **20**	Single gender **9**

Note (n=31)

Data were collected in structured researcher-driven diaries, with two pages per day.

Research Instruments and Procedures

Designed with Hinds' (2000) recommendation in mind, the information sought initially was not complex. To ease them into the recording process we asked them simply to note their noting arrival and departure times, as well as the times spent with colleagues at break and lunch periods. A survey-type item followed in which principals indicated their involvement with daily management duties. An open-response section allowed for elaboration.

Experience sampling prompted action of a different nature. On each day of the study, the principals received a text message at a random time, asking them to report on the activity in which they were directly engaged. Lemmens et al. (1988) and Camburn et al. (2010) demonstrated the effectiveness of experience sampling techniques and how time-related self-reporting strategies reduced reporting errors associated with brief or non-continuous events. Significant and meaningful activities may often be excluded from diary reports if recording is unprompted or left until the end of the school day. Interestingly, principals commented that experience sampling, built purposefully but unobtrusively into the diary, offered significant opportunities for professional self-reflection.

The focus shifted in the remaining sections to principals' experiences of leadership events. A survey-type item allowed principals to consider their leadership

involvement with pupils, staff, families or others. An open-response section prompted comments and personal reflection. Finally, there was an option to comment about work that was taken home. The periodic arrival of chocolate by post, in a process described by Cohen et al. (2007, p.223) as 'polite reminders', helped maintain a focus during the period of the diary keeping.

The complete set of diary entries were collated into a single anonymised volume and then returned to the participants with an invitation to participate in a second round of data gathering. Bryman (2001) observes that diary research is sometimes supplemented by personal interviews in which diarists add significantly to the richness of the data. Interviewing, when employed as a form of guided conversation, generates further insights and offers opportunities for the joint construction of meaning (Yin, 2009; Janesick, 2010).

Appointments were arranged, interviews were conducted by phone and conversations were digitally recorded. The audio files were transcribed verbatim without tidying up and were then returned to participants for member checking (Poland, 2002). The schedule of questions reviewed the narrative of the researcher-driven diaries with basic descriptive, clarifying, structural/paradigmatic and closing questions (Janesick, 2010). The responses, as with Dubin's (2006) conversations with Principals, provided a detailed narrative.

Data Analysis

While credible investigation techniques may generate worthwhile data, credible analysis required that such data were analysed systematically with strict adherence to the chain of evidence and to the case study protocol (Yin, 2009). Patterns were extracted using QSR's NVivo and as codes and code combinations were organised, emerging themes were categorised and information links were drawn. A narrative analysis allowed for explanation building and for theory building (Yin, 2009). Validity and reliability of findings was ensured by constant comparison of the data across respondents. While breaking the data into codes we carefully documented which view of a topic was delivered by which principal, in order to examine the strength of consensus and support, as well as the comprehensiveness of the data. The open coding was followed by axial coding allowing partially overlapping or related codes to be consolidated into one code. Finally the data were re-analysed through the lens of the emerging codes to ensure that saturation had been reached and the codes could be regarded as valid and reliable.

In formulating our findings, Miles and Huberman's (1994) matrices and visual thematic structures enabled us to organise our codes and to cluster them it into sub-categories. In this paper, we elaborate on the evidence that offered insights into the categories which emerged from this data analysis process, namely, administrative work, dealing with the unexpected, interacting with staff, work ethic, out of hours work and a willingness to do whatever is necessary. Each of these is considered in the following section.

Results

Administration

The findings yielded considerable differences in principals' attitudes to school administration and to the many tasks involved. For some, administration was seen as a key component of the role. Principal six declared it 'very positive to tick off items in the diary and to clear the desk' and in general, in-house administration duties such as sorting correspondence and drafting school policies were considered positively. Principal seventeen described such work as 'absolutely vital to the smooth and efficient running of the school'. Some even appeared to enjoy an oasis of calm at times, with the office door closed and the over-busyness of the school corridors kept at bay outside. Others saved administration for the evenings, when the building had emptied.

There was general negative commentary however about certain administration work. Managing pay-roll, organising local teacher-employment clusters and calculating relevant contractors' tax for school repairs were all described typically as frustrating, in that they drew principals' time and energies away from the core functions of teaching and learning. Their sentiment echoes Harris's (2007) description of principals paddling against an administration current that is 'heartbreaking and soul-destroying' (p.2). Principal twenty-six commented: 'I find that I have all the responsibility to deal with whatever is in hand, but effectively I have no power or control'.

Dealing with the Unexpected

Aspects of 'firefighting', as described by principal fifteen, arose throughout the narrative, particularly with principals in schools in socially disadvantaged areas and with those operating integrated Additional Learning Needs (ALN) units. Unpredictable, random and uncontrolled events were a workplace reality for some:

> I suppose, I'd say the worst that can happen in our school is being lifted out of it or being attacked by an angry parent. It is the unannounced attack, the unannounced vicious attack where you are not used to it and not used to being ranted and roared at, that is what I find the worst (principal one).

In ALN units, principals were called upon as the first line of defence in dealing with problematic pupils. In one possibly extreme case in the research, principal eighteen reported on fears for her personal safety and on the risk of being 'beaten, kicked or stabbed' by an out-of-control pupil. There was a general expectation that the principal would physically manage, negotiate outcomes and support traumatised staff where and when necessary. Such encounters, while no-doubt challenging when they occured, actually appeared to enhance authentic leadership-followership relationships at times (Crippen, 2012).

While some principals were quite positive and described dealing with the unexpected as an interesting and appealing aspect of the work, others were less so. The tone of principal thirteen's comment 'God only knows what I'm facing today', expressed a clear sense of personal vulnerability. High levels of unpredictability in some work contexts, combined with high levels of self-expectation, raised tensions when principals were called upon to face challenging circumstances, particularly as the public face of a school.

Schools in Ireland operate as part of a national public service and are subject to national austerity measures known as the Croke Park, Haddington Road and Lansdowne Road Agreements (LRC, 2016). New productivity initiatives and the timetabling of staff for a longer working week posed significant challenges for school leaders in the narrative. Furthermore, as senior staff retired, it was reported that an embargo on recruitment to middle management or special duties posts left principals feeling somewhat isolated and unsupported. Some considered their continued performance in the role unsustainable.

> The Deputy Principal and I are the only two post holders left in a school of 19 teachers, so this makes things very tough in terms of planning and teamwork ... There was an overwhelming sense of being a bit of a headless chicken today: I was trying to manage the school self-evaluation process, resource teaching hours, union issues with an ancillary staff member, plumbing problems, band practice, preparation for First Holy Communion, to name but a few issues. It left me feeling as if I did lots but achieved very little today. Overall, I am feeling very tired, which is worrying after a week off ... The trick is to figure out how to mind my physical and mental health. (principal nineteen).

Interacting with Staff

Our findings supported Bubb and Earley's (2004) assertion that management is characterised by a fragmentation of tasks and a myriad of activities, each one taking up a relatively short period of time; much like a series of spinning plates on poles. Principals' language suggested a preference for a collegial synergy in maintaining an overall momentum. A sense of approachability and trust was deliberately fostered in many cases. Principals advised, supported, explained and mentored. Significantly, when a positive and work-related focus dominated and when school successes ensued, staffs looked to principals to express appreciation and to affirm their achievements. In such instances, staffs were typically 'incredible', 'wonderful', 'positive' and 'professional'.

Some less-than-positive staff experiences were also recorded and in these cases, principals faced an altogether different set of daily realities. Principal thirteen's anxious assertion that 'the buck lands with the principal no matter what happens in the school' may explain the exasperation that led others to describe a minority of unsupportive staff as 'wingers', 'lazy' or 'incompetent'. One principal elaborated:

> The other thing that ... could be upsetting and that could make a day horrible is having to deal with a member of staff who maybe is not doing what they should be ... I mean being constantly late ... always out of their room and it's not on school business … you are not used to correcting adults like that and you don't expect to have to correct them.

Some principals inherited long-standing interpersonal difficulties on appointment. In one such scenario and lacking staff support, principal twenty described himself 'hiding amongst the children' to find satisfaction in the job. Principal eleven reported similar experiences among colleagues:

> Some principals ... working in bigger schools, had to jack in the job. They had been brought in from outside when there were three or four internal candidates who had applied and who didn't get the job … their life was hell I can tell you. It was absolutely hell on earth.

On a somewhat different note, principals acknowledged that staff members can experience their own personal difficulties and may operate out of stressful circumstances at times. In particular, staff illnesses, bereavements and preoccupation with difficult home or personal circumstances effected teachers' performances and were noted in the research. In such accounts, principals demonstrated an overarching complement dualism, that of showing care and concern for colleagues while attempting to successfully manage the daily practicalities (Scarlett, 2015). Principal twenty-four offers the following diary entry:

> A colleague is out on long-term sick leave and is worried about her entitlement. I can't talk to anyone else about her condition and it's not widely known. It is difficult to keep these things inside. I cleaned the floor of the 4th class toilets. Somebody is making a mess. I think it's the unruly pupil but I can't be sure – yet.

Work Ethic

Spindler and Biott (2005) describe principals' enormous personal investment of time and energies in their schools and their willingness to go beyond the call of duty. A strong work ethic was evident in this research also, particularly in interactions with parents and families. In high-achieving school communities, it was noted that principals drove school performance in key areas and actively supported parents' aspirations for successes. In the case of children with ALNs, principals organised meetings with service providers and acted as advocates for child and family supports.

On school discipline issues, they engaged actively in protracted processes and worked towards optimum attendance and school completion targets.

Across the spectrum, principals were cognisant of the cost of their investments and they accepted the physical implications of being constantly available to everyone and of feeling responsible for just about everything. In a role where duties were entirely uncontracted, it appeared that every aspect of the principal's work was taken for granted by everyone, including the principal.

> Principals do put in an awful lot of their own time and we are all the same I think. I treat my school like my home, with how much I respect and care I have for it., It came through with a lot of Principals (in their diaries) that that's how they feel about their schools as well. They would do anything to keep this show on the road even if it does impede on their day (principal thirteen).

Out-of-hours Demands

Evidence here, as elsewhere, indicated that performing a principal's duties involved more than simply taking on a time-consuming occupation. A sense of personal engagement, whether driven by vocation or simply a sense of the common good, appeared to underpin principals' attitudes. In their stories of family life, newborn babies, dinners and evening routines, principals described how important aspects of the home became secondary at times to the mental and physical demands of out-of-hours work. Principal eighteen elaborated:

> I think that you need some kind of discipline because it is all encompassing. … I put all of my energy into the job to the detriment of everything else you know, including family life and everything else. Yes, I think it would be good advice for principals to have a time and to go. Get it off your mind and definitely don't bring physical work home. It will be in your head anyway. You cannot get rid of that but with discipline, you could lessen it.

After-school engagements typically involved administrative duties, CPD and returning to attend evening meetings. Principal seventeen explained that 'someone has to do it'. In some cases, there appeared to be no discernible boundaries and principals prioritised work-place duties alongside those of their family housekeeping. Principal seven reported being in the school attic on a particular afternoon inspecting roof damage. When asked "*why?*" she replied somewhat bemusedly:

> Yes, why was I in the attic? It's a little bit like when I find myself in the supermarket buying toilet rolls. I'm buying 500 toilet rolls and I ask, "how did this happen?" (Laughs).

The most significant after-hours demands were found when principals were occupied mentally with problems of staff relations or with child protection concerns. Some expressed a sense of guilt about returning home to the family and facing into the evening 'worn out' or 'wrecked'. Principal twenty-one told of how she avoids talking about the 'woes' with her family and only reports the 'good bits'. With a similar outlook, principal eighteen was actually pleased that her family was out when she arrived home. She enjoyed the solitude of an empty house:

> I had a relaxing bath when I came home. There was nobody in at all and I was glad, as I had nothing left to give anyone. I sat down in the chair exhausted and I fell asleep. When I awoke I had the bath. Then, I felt human again.

Personal approaches to the role were undoubtedly varied but there was general agreement that school communities demanded certain visible leadership characteristics to be demonstrated. Kelchtermans et al. (2011) describe an expectation that principals should present a good face and appear strong. For some, this experience was highly demanding and feelings of personal vulnerability and isolation became an endemic and chronic condition of leadership life.

> I find the personal end of the role difficult. By nature, I'm more introverted than extrovert and I find the public nature of the job difficult. I find I bury my feelings about it quite a bit or I wouldn't be able to do it at all (principal twenty-one).

Irrespective of their individual personality traits and circumstances, the principals in the research demonstrated a positive not-for-profit work-ethic, accompanied by high levels of self-expectation. This was evident in their commentaries about school governance where a possible over-involvement in day-to-day management responsibilities was noted.

A Willingness to Tackle Everything

The law in Ireland establishes the direct accountability of the principal to the Board of Management and in theory at least, this suggests that a hierarchical model of vertical relationships should exist (Bush, 2003). The evidence found in this research was that a flat organisational structure operated in practice and that volunteer members of Boards of Management were almost entirely dependent upon the professional guidance and expertise of their principals. Principal five offered the following perspective:

> We have a new Chairperson on the Board this year who has no education background. His only experience of it was when he was at school and that was not good. So, I have had to change approaches and try to lead him too. He is a nice man but it is weird trying to lead your boss.

Principals also reported on the difficulties of constituting school Boards, a difficulty highlighted nationally by the LDS (2007). Volunteers, particularly those in areas of social disadvantage, were reluctant to take on the long-term and significant managerial responsibilities. One such responsibility was school finances and in cases where Boards failed to actively engage an appointed Treasurer, principals undertook the duties and responsibilities themselves (CPSMA, 2016). Principals wrote cheques, calculated taxes, negotiated insurance premiums, organised the payment of staff and advised Treasurers on the preparation of statements of accounts for meetings. In cases where finances fell short, it was the principal who organised fundraising events and who prioritised the fundraised spending.

> I feel a reduction in funding to schools, through the capitation grant and in particular the minor works, has led to a constant need for fundraising with the Principal leading this process. Over the last 8 years I have been involved in fundraising for (sports, literacy and ICT resources). In all, this can be extremely demanding in addition to the normal job (principal twenty-nine).

Principals also prepared and submitted detailed and time-consuming funding applications to the Department for summer building renovations and for school extension works. In the event of trouble-shooting, they acted as the point of contact for architects, engineers, solicitors and even rogue-builders. For general maintenance, they assumed responsibility for tendering, organising and overseeing painting, landscaping, heating, electrical, security and minor construction projects. It could reasonably be asserted that principals were inclined not only to lead and manage, but to assume the full working responsibility and role of the entire Board of Management at times. As principal eighteen explained:

> By the time the BOM meeting comes around ... issues may sort themselves. What inevitably happens is that I solve them with the help of the Chairperson of the BOM and they need never take up time on the BOM.

Discussion, Conclusion and Recommendations

Gunter's (2001) description of school leadership as a difficult but worthwhile journey is worth considering afresh. Despite an overwhelmingly positive response from principals in the study, some difficulties are undeniable. Principals in this study spoke at times of frustration, exhaustion and stress. They described the health impacts of sleep problems, of weight gain, of feelings isolated and of lacking support in daily

management and leadership experiences. Southworth (2002) and Kelchtermans et al. (2011) reported similar difficulties with principals' work-life balance in their depictions of the implicit loneliness of the school gatekeeping role.

The difficulties highlighted in this Irish study were undoubtedly exacerbated by contextual factors. Principals juggled the adverse effects of pay-cuts, increased workload and worsening conditions of service (LRC, 2016). Principals reported that a perceived 'battering' of teachers in the national media only compounded their demoralisation.

Yet, while all such difficulties were recorded, more worthwhile and positive aspects of the journey dominated the narrative. Southworth (2002) argues that for many, being a Principal is more an identity than a role, one that is socially constructed by the expectations of others. Communication, occupational socialisation and interactions with children influence the development of this identity.

Principals perceived their leadership role positively and provided details of leadership engagement under clearly discernible headings. These involved motivating staff, improving the fabric and appearance of the school and enabling the school to play an active and public role in local community life:

We support Iszatt-White's (2011) assertion that leadership must be seen through ethnomethodology and through emphasising the situated real-time nature of action. Leadership is only leadership because the members of a particular setting inter-subjectively negotiate a shared understanding and appreciate collectively that certain actions undertaken by a chosen individual in a setting constitute leadership.

We conclude by acknowledging a key limitation of the paper. Journal wordcount allowed us provide only a partial treatment of a broad case study. Within the limitations of a narrative synopsis, we selected only some of our findings. We concentrated our attentions on those that offered possible commonalities in the experiences of principals in Ireland and those elsewhere and we sought, with the inclusion of diary and interview comments, to provide a focus on the material conditions of principals' lives (Thomson, 2009).

The full text of the narrative, including the content of the researcher-driven diaries as appendices, is available for further study through the University portal (Stynes and McNamara 2014). Our aim in the broader case study was to shift the focus of leadership research from the anti-intellectual philosophical dominance of a value-for-money economic imperative and to focus instead on case-study naturalistic generalization.

In an occupation where contextualised and sustainable leadership is required to maintain the person-centeredness of the endeavour, we concur with Taysum (2003), Ribbins (2008), Gunter (2001, 2016) and Bottery (2016) who stress the need for a deep, intellectual, ethical and holistic treatment of what it is to be a school leader.

A challenging but worthwhile career journey is aptly drawn in Green's comment; 'it's a great life, if you don't weaken' (2000, p.10). In Ireland, McGovern (2016)

identifies that the career trajectory from teacher to principal is dependent largely on the influences of serving principals who recognise and promote leadership talent. A paucity of formalised leadership development structures leaves aspirants reliant therefore upon circumstance, personality and life experience. While this has been sufficient in times past and may be sufficient for some within the contexts of supportive school communities, concerns arise when the range and complexity of external influences on schools are considered. The pressures they exert pose significant challenges to principals who juggle professional, societal and personal expectations in a hurricane of accountabilities (Moos et al., 2011). It is appropriate to conclude with the very striking of words of one principal

> I am consistently called upon each day to demonstrate leadership in a myriad of situations. I love the opportunities this presents but sometimes find myself exhausted from the emotion of it all. I believe this exhaustion and emotion build barriers in my personal life. The current climate has pushed leadership to its limits (principal thirty one).

Amid a vast sea of information on how to successfully enact the principalship, Wood (2002) argues that the education community neglects the investigation of the affective dimensions of the work and of what the job does to those who undertake it. As principals engage in daily cycles of perpetual motion with a willingness and desire to do all that is humanly possible and as they manage and lead in circumstances where they place children first in the universal order of importance over everything else imaginable, what of the principals themselves? We propose that this is a question of significant importance to us all.

The implication of this research is, in our view, that a fuller understanding of the reality of the role of the principal requires closely grounded study of the actuality of their daily life. Principals understand the nature of the role and are deeply committed to it but policy makers and education authorities must be careful about the ever increasing expectations and burdens paced on the shoulders of school leaders. A greater emphasis on administrative support and the encouragement and development of distributed leadership in a genuine sense are vital ways of easing the burden.

References

Adair, J. (2000). *How to find your vocation*. Norwich: Canterbury Press.

Addison, B. (2009). A feel for the game – a Bourdieuian analysis of principal leadership: A study of Queensland secondary school principals. *Journal of Educational Administration and History,* 41(4), 327-341.

Anderson, G. L., & López, A. (2017). Resisting and reclaiming the global discourse of leadership. In D. Waite & I. Bogotch (Ed.), *The Wiley international handbook of educational leadership* (pp. 157-174). Hoboken, NJ: John Wiley & Sons.

Bakker, J.I. (2010). Ontology In A.J. Mills, G. Durepos, & E. Wiebe (Ed.), *Encyclopedia of case study research volume 2* (pp. 628-631). Thousand Oaks, CA: Sage.

Bottery, M. (2016). *Educational leadership for a more sustainable world*. London: Bloomsbury.

Bryman, A. (2001). *Social research methods*. Oxford: Oxford University Press.

Bubb, S., & Earley, P. (2004). *Managing teacher workload: Work-life balance and wellbeing*. London: Paul Chapman Publishing.

Burgess, H. Sieminski, S., & Arthur, L. (2006). *Achieving your doctorate in education*. London: Sage Publications.

Bush, T. (2003). *Theories of educational leadership and management*. London: Sage.

Camburn, E.M., Spillane, J.P., & Sebastian, J. (2010). Assessing the utility of a daily log for measuring principal leadership practice. *Educational Administration Quarterly, 46*(5), 707-737.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education*. London: Routledge.

Coolahan, J. (1981). *Irish education: History and structure*. Dublin: Institute of Public Administration.

CPSMA. (2017). Catholic primary school management association. Retrieved November, 7, 2017, from https://www.cpsma.ie/

Crippen, C. (2012). Enhancing authentic leadership–followership: Strengthening school relationships. *Management in Education* 26(4), 192–198.

CSL. (2017). *The centre for school leadership*. Retrieved November, 7, 2017, from http://www.cslireland.ie/

Day, C. (2007). What being a successful principal really means: An international perspective. *Educational Leadership and Administration, 19,* 13-24.

Day, C., & Smethem, L. (2009). The effects of reform: have teachers really lost their sense of professionalism? *Journal of Educational Change, 10*(2-3), 141-157.

DES Department of Education and Skills. (2016). *Annual statistics report 2015/2016*. Retrieved August, 17, 2016, from https://www.education.ie/en/Publications/Statistics/Statistical-Reports/

Drea, E., & O'Brien, J. (2003). *Defining the role of the primary principal in Ireland*. Dublin: HayGroup Management Consultants.

Dubin, A.E. (2006). Conversations with principals: Issues, values and politics. *Educational Leadership and Administration, 18,* 85-97.

Eidlin, F. (2010). Reality In A.J. Mills, G. Durepos, & E. Wiebe (Ed.), *Encyclopedia of case study research volume 2* (pp. 781-783). Thousand Oaks, CA: Sage.

Fink, D. (2005). *Leadership for mortals.* London: Paul Chapman Publishing.

Fullan, M. (2014). *Leading in a culture of change: Personal action guide and workbook.* Hoboken, NJ: John Wiley & Sons.

Gillham, B. (2000). *Case study research methods.* London: Continuum.

Glatter, R. (2014). Educational administration 1969-2014: Reflections on pragmatism, agency and reform. *Journal of Educational Administration and History, 46*(3), 351-366. doi:10.1080/00220620.2014.919904

Green, F. (2000). *The head teacher in the 21st century: Being a successful school leader.* London: Pearson Education.

Gunter, H. (2001). *Leaders and leadership in education*. London: Paul Chapman publishing.

Gunter, H. M., & Fitzgerald, T. (2013). *New Public Management and the modernisation of education systems 1.* London: Taylor & Francis.

Gunter, H. (2016). *An intellectual history of school leadership practice and research*. London: Bloomsbury.

Hall, V. & Southworth, G. (1997). Headship. *School Leadership and Management, 17*(2), 151-170.

Hall, D., Møller, J., Schratz, M., & Serpieri, R. (2017). From Welfarism to Neo-Liberalism: Conceptualizing the diversity in leadership models in Europe. In D. Waite & I. Bogotch (Ed.), *The Wiley international handbook of educational leadership* (pp. 311-334). Hoboken, NJ: John Wiley & Sons.

Hallinger, P. (2005). Instructional leadership and the school principal: A passing fancy that refuses to fade away. *Leadership and Policy in Schools, 4*(3), 221-239.

Harris, B. (2007). *Supporting the emotional work of school leaders*. London: Paul Chapman Publishing.

Hinds, D. (2000). Research instruments. In D. Wilkinson, *The researcher's toolkit: A complete guide to practitioner research (p.41-54).* London: RoutledgeFalmer.

INTO. (2017). *Irish national teachers' organisation*. Retrieved November, 1, 2017, from https://www.into.ie/ROI/

IPPN. (2017). *Irish primary principals' network*. Retrieved November, 1, 2017, from http://www.ippn.ie/

Ireland. (1998). *The education act.* Retrieved November, 7, 2017 from http://www.irishstatutebook.ie/1998/en/act/pub/0051/index.html

Iszatt-White, M. (2011). Methodological crises and contextual solutions: An ethnomethodologically informed approach to understanding leadership. *Leadership*, 7(119), 119-135.

Janesick, V.E. (2010). *Oral history for the qualitative researcher: Choreographing the story.* New York: The Guilford Press.

Kelchtermans, G., Piot, L. & Ballet, K. (2011). The lucid loneliness of the gatekeeper: Exploring the emotional dimension in principals' work lives. *Oxford Review of Education,* 37(1), 93-108.

Kuhn, T.S. (1996). *The structure of scientific revolutions.* Chicago: University of Chicago Press.

LDS Leadership Development for Schools. (2007). *Improving school leadership: country background report – Ireland. OECD activity.* Ennis: Clare Education Centre.

Lemmens, P., Knibbe, R.A., & Tan, F. (1988). Weekly recall and daily estimates of alcohol consumption in a general population survey. *Journal of Studies on Alcohol and Drugs* 53, 476-486.

Linstead, S.A. (2010). Postmodernism In A.J. Mills, G. Durepos, & E. Wiebe (Ed.), *Encyclopedia of case study research volume 2* (pp. 694-700). Thousand Oaks, CA: Sage.

LRC Labour Relations Commission. (2017). The Lansdowne Road Agreement. Retrieved January, 29, 2017 from http://www.into.ie/ROI/LansdowneRoadAgreement/Downloads/Landsdowne RdAgreement.pdf

McDonald, V. (2008). Leadership issues for principal teachers. *Oideas, 53,* 27-40. Dublin: Department of Education and Science.

McGovern, E.M. (2016). *From aspiration to succession and transition: An exploration study of the fundamental principles of sustainable school leadership from the perspectives of newly appointed principals in Irish primary education.* Unpublished Doctoral dissertation. Dublin: Dublin City University.

Miles, M.B., & Huberman, M. (1994). Qualitative data analysis: An expanded sourcebook. London: Sage.

Moos, L., Johansson, O., & Day, C. (2011). *How school principals sustain success over time: international perspectives.* Dordrecht NL: Springer.

Morgan, M., & Sugrue, C. (2008). The Seven challenges and four rewards of being a school principal: Results of a national survey. *Oideas 53,* 8-26. Dublin: Department of Education and Science.

Niesche, R., & Thomson, P. (2017). Freedom to what ends? — School autonomy in neoliberal times. In D. Waite & I. Bogotch (Eds.), *The Wiley international handbook of educational leadership* (pp. 193-206). Hoboken, NJ: John Wiley & Sons

O'Leary, Z. (2005). *Researching real world problems: A guide to methods of inquiry.* London: Sage Publications.

Palmer, P.J. (1998). *The courage to teach: Exploring the inner landscape of a teacher's life*. San Francisco: Jossey-Bass.

PDST. (2017). *Professional development service for teachers*. Retrieved January, 29, 2017 from http://www.pdst.ie/

Poland, B.D. (2002). Transcription quality. In J. Gubrium, & J. Holstein (Ed.), *Handbook of interview research: Context and method* (pp. 629-651). Thousand Oaks: Sage.

Potter, D., Reynolds, D. & Chapman, C. (2002). School improvement for schools facing challenging circumstances: A review of research and practice. *School Leadership & Management,* 22(3), 243–256.

Ribbins, P., & Marland, M. (1994). *Headship matters: Conversations with seven secondary school headteachers.* Harlow: Longman.

Ryan, J., & Higginbottom, K. (2017). Politics, activism, and leadership for social justice in education. In D. Waite & I. Bogotch (Eds.), *The Wiley international handbook of educational leadership* (pp. 103-124). Hoboken, NJ: John Wiley & Sons

Ribbins, P. (2008). A life and career based framework for the study of leaders in education: problems, possibilities and prescriptions. In J. Lumby, G. Crow & P. Paschiardis (Ed.), *Educational management: refining theory, policy and practice* (pp. 61-81). London: Paul Chapman.

Ryan, J. (2012). *Struggling for inclusion: Educational leadership in a neo-liberal world*. Charlotte, North Carolina: IAP.

Saltman, K. J., & Means, A. J. (2017). From "data-driven" to "democracy-driven" educational leadership: Navigating market bureaucracy and new technology in a 248 Post-Fordist era. In D. Waite & I. Bogotch (Eds.), *The Wiley international handbook of educational leadership* (pp. 125-138). Hoboken, NJ: John Wiley &Sons

Scarlett, W.G. (2015). *The sage encyclopedia of classroom management*. London: Sage.

Scott, D., Brown, A., Lunt, I., & Thorne, L. (2004). *Professional doctorates: Integrating professional and academic knowledge.* Maidenhead: Open University Press.

Smith, N.J. (2009). *Achieving your professional doctorate.* Maidenhead: Open University Press.

Southworth, G. (2002). School leadership in English schools: Portraits, puzzles and identity. In A. Walker & C. Dimmock (Ed.) *School leadership and administration: Adopting a cultural perspective* (pp. 187-205). New York: RoutledgeFalmer.

Spindler, J. & Biott, C. (2005). Immediacy and distance: Intergenerational professional learning. In C. Sugrue (Ed.) *Passionate principalship: Learning from the life histories of school leaders* (pp. 123-135). London: Routledge.

Stake, R. (1995). *The art of case study research.* London: Sage.

Stynes, M. & McNamara, G. (2014). *Walk a mile in my shoes: A case study of the everyday lives and work experiences of a group of Irish primary school principals*. Unpublished Doctoral dissertation. Dublin: Dublin City University. Retrieved November, 7, 2017 from http://doras.dcu.ie/20154/

Starrat, R.J. (2005). Foreword. In C. Sugrue (Ed.) *Passionate principalship: Learning from the life histories of school leaders* (pp. xi-xiii). London: Routledge.

Sugrue, C. (2003). Principals' professional development: Realities, perspectives and possibilities. *Oideas,* 50, 8-39. Dublin: Department of Education and Science.

Sugrue, C. (2005). (Ed.) *Passionate principalship: Learning from the life histories of school leaders.* London: Routledge.

Taysum, A. (2003). In search of the holistic leader. *Management in Education* 17(5), 9-12.

Taysum, A. (2013). The impact of doctoral study on educational leaders' work for students' participation in education systems and society. *Educational Review* 65(4), 432-446.

The Teaching Council. (2016). *The teaching council*. Retrieved November, 7, 2017 from http://www.teachingcouncil.ie/

Thomson, P. (2009). *School leadership: Heads on the block.* New York: Routledge.

Thrupp, M., & Willmott, R. (2003). *Educational management in managerialist times.* Maidenhead: Open University Press.

Tomlinson, H., Gunter, H., & Smith, P. (1999). *Living headship: Voices, values and visions.* London: Paul Chapman.

Troman, G. (2006). Headmasters, collaborative school cultures and school improvement: A changing relationship? *Educational action research,* 4(1), 119-144.

Weick, K.E. (1995). *Sense making in organisations*. London: Sage.

Woods, R. (2002). *Enchanted headteachers: Sustainability in primary school headship*. Nottingham: NCSL.

Yin, R.K. (2009). *Case study research: Design and methods.* Thousand Oaks, CA: Sage.

Gaps between Acquired and Required Teacher Education Graduate Attributes: Does Accreditation Influence in Pakistan?

Muhammad Ayub BUZDAR[1], Hina JALAL[2], Muhammad Naeem MOHSIN[3]

ARTICLE INFO	ABSTRACT
	Purpose: The major purpose of the present study was to identify the gaps between acquired and required graduate attributes of pass out graduates of teacher education programs in Pakistan and investigate the influence of program accreditation on these gaps. **Research Method:** A quantitative research method was used to achieve the research objectives. 160 graduates of the accredited and non-accredited master level teacher education programs associated with 24 different universities in Punjab and Islamabad Capital Territory participated in the study. A questionnaire based on national professional standards for teachers (NPSTs) in Pakistan was developed to collect the data. **Findings:**

The study demonstrates that the participants' acquired attributes do not match the required attributes of their job. The gaps between acquired and required graduates' attributes are comparatively smaller among the graduates of accredited programs. The study supports the positive effects of accreditation on the output of teacher education programs. **Implications for Research and Practice:** The findings of this study strengthen the idea that a well framed accreditation mechanism positively influences the products of professional programs. Further research may be conducted on making teacher education accreditation programs more effective in Pakistan with respect to the graduate attributes.

---

[1] Corresponding author; Government College University Faisalabad, PAKISTAN, Email: ayubbuzdar@gmail.com, ORCID: http://orcid.org/0000-0001-8719-3295
[2] Government College University Faisalabad, PAKISTAN, Email: hinansari23@yahoo.com, ORCID: http://orcid.org/0000-0002-8273-5195
[3] Government College University Faisalabad, PAKISTAN, Email: mnmohsin71@gmail.com, ORCID: https://orcid.org/0000-0001-8386-053X

Introduction

At present, higher education institutions (HEIs) are being transformed due to the expansion of the higher education provision, internationalization, advance technology, and expansion of employability. These changes are influencing the requirements of stakeholders. The massive growth of HEIs and students has compelled the HEIs to make arrangements for meeting the requirements of all stakeholders, including employers and investors, to ensure that their graduates are equipped with necessary graduate's attributes (Oliver & Jorre de St Jorre, 2018). However, researches demonstrate that it is difficult to predict the development of graduate attributes in students (Glover & Hope, 2015). The universities are supposed to assure the proficiency of graduates by developing the necessary professional and personal attributes. Several studies been done to explore the development of graduate attributes in association with the institution and program accreditation. Accreditation agencies judge the quality of programs and institutions to the quality assurance of graduates as per pre-defined standards. Howbeit, less attention is paid to exploring graduate attributes development and assessment in the field of teacher education in Pakistan. In the present study, researchers explore the perceptions of university graduates about their acquisition and requirement of attributes for work place. The value of required graduate attributes development through accredited teacher education programs is also discussed in the study.

The term "graduate attributes" is interchangeably used to refer competencies, qualities, employability skills, generic attributes, transferability, and learning outcomes (Nassef, 2016; Shivoro, 2018). Researchers, industrialists, and practitioners are more concerned about the relevance of higher education with the job market through contemporary graduate attributes. Since that the students' education and training in tertiary institutions are emphasized more than before. The productive development of graduates for the dynamics of the working world is the first priority of universities (Oladokun & Olaleye, 2018). Knowledge, skills, and dispositions along with good citizenship behaviors are considered valuable as program outcomes (Marsh & De Luca, 2016).

The development and implementation of graduate attributes is based on several commonalities including organizational culture, resources, strategies, curriculum design, teaching, and general institutional supports available to the students (Bond et al., 2016). In the teacher preparation, the attributes of an exemplary teacher entail professional and personal attributes (including knowledge, skills, values, competencies, and devotion) that are needed to fulfil the requirements of the teaching profession. The range of graduate attributes in teacher education has been subject of research (McInerney, Korpershoek, Wang, & Morin, 2018). The main purpose of teacher education is to promote professional development, training, and the education of educators. The area of teachers' knowledge and its practice in the classroom is generally neglected (Czerniawski et al., 2018). Since the quality of education directly depends on the quality of educators, the quality of teaching can be effectively developed preparing competent and potential graduate teachers. For this purpose, the

professional standards of teachers are directly linked with the curriculum of teacher education (Révai, 2018).

The employers' expectations are not always aligned with the curriculum of higher education (Kaushal, 2016). Most of the time, the employers complain that the graduates lack attributes that are necessary for their professional development (Grant Thornton Consulting, 2016; Nassef, 2016). Employers expect that graduates to be equipped with technical competencies, and a broader range of relevant attributes. Currently, the HEIs pay much attention to producing distinctive quality attributes in graduates, such as, qualification, discipline knowledge, discipline skills, innovation, competency, and intellectuality from global point of view.

Concerning the conditions in higher education, it is need that the quality assurance organizations evaluate the relevancies of higher education, and how it contributes to developing the desired graduate attributes. Similarly, the accreditation agencies should monitor the quality of programs to foster the accepted graduate attributes. Accreditation refers to a third-party peer review to validate educational institutions and programs. It is a self-regulatory mechanism of institutional accountability. It is an apparatus to evaluate the institutions and programs (Sin, Tavares, & Amaral, 2017). Different countries have enacted legislation to categorize the graduate attributes. Such legislation also empowers the quality assurance agencies to evaluate the program outcomes (Sonnenschein, Barker, & Hibbins, 2017). In the United Kingdom, the quality assurance agency (QAA) evaluates the development of graduate attributes in higher education (Quality Assurance Agency for Higher Education, 2016). In the case of Pakistan, 189 HEIs produce a large number of graduates annually (Khattak, 2016). The burgeoning growth of tertiary institutions and graduates has placed more responsibility on the Higher Education Commission (HEC) of Pakistan and the relevant accreditation agencies. In the Pakistani model, accrediting agencies, generally evaluate programs periodically to assure the production of graduates as required at national/international level. The major objective of the HEC is to modulate appropriateness and consistency of accreditation.

The National Accreditation Council for Teacher Education (NACTE) in Pakistan places less emphasize on evaluating graduate attributes development in accrediting teacher education programs. Consequently, despite all the endeavors of the NACTE and HEC, the status of teacher education graduates is not as high as it is required to be. Research shows that 76% of employers were not satisfied with the performance of Pakistani graduates (Dodhy, 2016). Some 96% of the recruiters stated that the knowledge of graduates does not match their academic subject (Shahbaz, 2016). Despite the massive growth of higher education institutions, the quality of graduates is falling (Naveed, 2017). The present study aims to identify the gaps between acquired and required attributes of education graduates and assess the influence of the accreditation status of programs on these gaps. The study intends to answer the following research questions;

1. What are the gaps between acquired and required graduate attributes of the pass outs of MA Education/M.Ed. programs included in the sample?

2. To what extent does the accreditation of MA Education/M.Ed. programs influence on the gaps between acquired and required attributes of the graduates included in the sample?

Method

Research Design

A quantitative research design was used in this study to find the answers to the two research questions. The data were collected through survey, and a closed ended questionnaire that was administered to gather the required information. Graduates of the Master of Arts in Education and Master in Education programs who passed their final examinations in 2015 and 2016 were invited to participate in the inquiry.

Research Sample

Participants in this study included 160 graduates of the MA Education and M.Ed. programs offered in 24 public and private higher education institutions in Punjab and Islamabad capital territory. The researchers categorized the participants based on the accreditation status of their programs. Half of the participants (80) were awarded their degree from accredited MA/M.Ed. programs, while the remaining 80 participants were graduates of non-accredited MA/M.Ed. programs. Of the graduates of accredited MA/M.Ed. programs, 40% were males and 60% were females. While, 45% of the participants were graduates in 2015, while the remaining 55% were graduates in the year of 2016. At the time of data collection, 23.8% of the graduates of accredited programs were still enrolled in institutes of higher education, 38.7% were employees, and 37.5% were unemployed. From the non-accredited MA/M.Ed. programs, 20% of the graduates were male and 80% were female. Whereas, 53.8% of the participants were graduates in the year of 2015, the remaining 46.2% were graduates in the year of 2016. At the time of data collection, 26.2% of the graduates of non-accredited programs were still enrolled in institutes of higher education, 28.8% were employees, and 45% were unemployed.

Research Instruments and Procedures

In 2009, the government of Pakistan implemented a program called "Strengthening Teacher Education in Pakistan" (STEP), with the technical support of the United Nations Educational Scientific and Cultural Organization (UNESCO), and the financial support of the United States Agency for International Development (USAID). Under the umbrella of STEP, the National Professional Standards for Teachers (NPSTs) were formulated to empower the beginning and secondary level teachers (Government of Pakistan, 2009). These standards include mastery in subject knowledge, human growth and development, knowledge of Islamic values/social life skills, instructional planning and strategies, assessment, learning environment, effective communication and proficient use of information communication technologies, collaboration and partnerships, continued professional development and code of conduct, and teaching of English as a second/foreign language. These

standards were developed to define competencies, skills, and attributes of educators. Quality assurance agencies critically analyze all the factors that contribute to the quality of education and evaluate its impact on students' learning (National Professional Standards for Teachers in Pakistan, 2009). The researchers in this study have accepted these standards as required graduate attributes in teacher education.

Validity and Reliability

Experts in the field of education were requested to check the construct validity of NPSTs as graduate attributes. An average congruency percentage (ACP) approach by Popham (1978) was followed to measure the face and content validity of the self-reported questionnaire developed on the basis of NPSTs in Pakistan. The ACP > 90% of the questionnaire was understood as validated (Najafi-Sharjabad, Hajivandi, & Rayani, 2014; Popham, 1978). On the recommendations of the experts, the researchers separated the code of conduct attribute from continues professional development, separated effective communication from proficient use of information communication technology (ICT), and replaced the knowledge of Islamic ethical values/social life skills with ethical values. Twelve attributes with 36 statements under three divisions headed as knowledge, skill, and disposition were finalized. The questionnaire was validated at the 100% value of ACP. The internal consistency of the questionnaire was ensured at Cronbach's α > .7.

Data Analysis

Inferential statistics were applied on collected data to analyze the collected data. T-tests and mean difference techniques with the help of statistical Package of Social Sciences (SPSS) version 21.0 were applied in this process. Findings of the data analysis are presented in tabulation and figures. The data analysis facilitated in finding the answers of research questions.

Results

It was found that there are significant ($p < .001$) statistical mean differences between the acquisition of graduate attributes at a teacher education institution and their perceived requirement at the work place. The graduate attributes including mastery in subject knowledge, human growth and development (HGD), ethical values, instructional planning and strategies, assessment, learning environment, effective communication, information communication technology (ICT), collaboration and partnerships, life-long learning, code of conduct, and language were assessed in the study (Table 1). The requirement for the role teacher for these graduate attributes is higher than that of their acquisition to perform satisfactorily in the field of teacher education. It can be argued that the graduates acquired attributes do not meet the employers' expectations from the master level teacher education programs.

There were significant mean differences between the graduates' acquisition and requirement with respect to the accredited and non-accredited MA Education/M.Ed. programs on every attribute (Figure 1). It was found that the gaps on the graduate

attributes of subject knowledge, HGD, ethical, instructional planning, assessment, communication, ICT, and language were higher for the non-accredited MA Education/M.Ed. programs than for the accredited MA/M.Ed. programs.

The gaps on four of the graduate attributes learning environment, collaboration, lifelong learning, and code of conduct were greater for the accredited MA Education/M.Ed. programs in comparison to the non-accredited MA Education/M.Ed. programs. The cumulative gap between the acquired and required graduate attributes for the accredited MA Education/M.Ed. programs (MD = 1.70) was lower than the non-accredited MA Education/M.Ed. programs (MD = 1.78). It was discovered that the accreditation of MA Education/M.Ed. programs influenced the acquisition of graduate attributes (eight out of twelve) more positively than for the non-accredited MA Education/M.Ed. programs.

Table 1

Comparison Between Acquired and Required Attributes of MA/M.Ed. Graduates.

Graduate attributes	Acquired		Required			
	M	SD	M	SD	MD	t (df=159)
Subject knowledge	2.91	.619	4.03	.184	-1.12	-21.84*
HGD	2.98	.820	4.60	.331	-1.62	-22.62*
Ethical values	2.10	.549	4.68	.319	-2.57	-50.76*
Planning	2.43	.850	4.73	.288	-2.30	-31.56*
Assessment	2.76	.863	4.36	.255	-1.60	-25.12*
Environment	3.08	.997	4.76	.354	-1.69	-20.08*
Communication	2.61	.007	4.47	.400	-1.86	-25.17*
Use of ICT	2.90	.991	4.42	.418	-1.53	-20.47*
Collaboration	2.44	.721	4.34	.382	-1.90	-32.56*
Life-long learning	2.53	1.00	4.34	.405	-1.81	-22.23*
Code of conduct	3.16	.838	4.26	.436	-1.10	-17.06*
Language	2.66	.784	4.61	.321	-1.95	-30.32*

Note: Environment = learning environment; Planning = instructional planning and strategies; Communication = effective communication; Collaboration = collaboration & partnerships; Code of conduct = professional code of conduct; Language = teaching of English as a second/foreign language; * = The value is significant at .05 level.
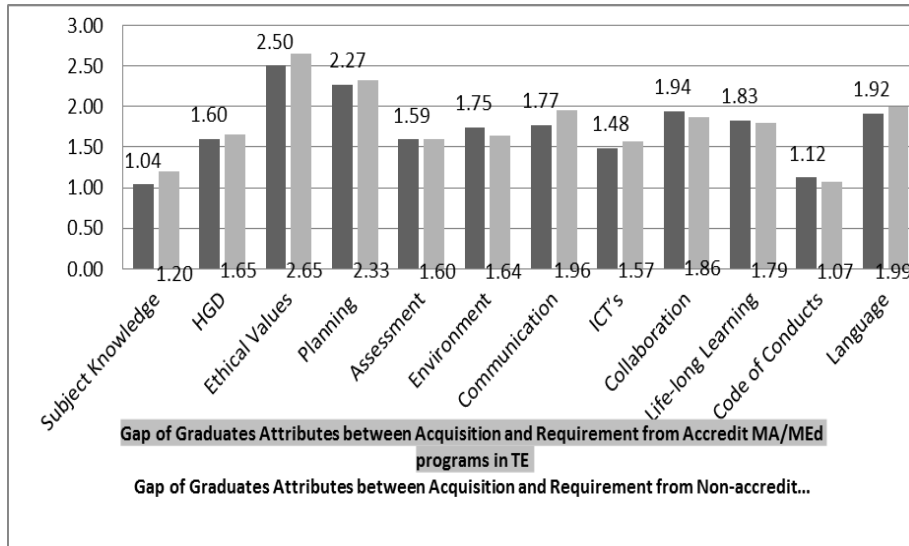
Figure 1. Comparison among the gaps of graduate attributes of the pass outs of accredited and non-accredited MA Education/M.Ed. programs

Discussion, Conclusion and Recommendations

The success in developing competent and expert teachers lies in the practices of their knowledge and skills on the job. Over the last twenty years, improving teacher education and the teaching profession have remained on the main agenda of market forces and overall development of the country (Czerniawski et al., 2018; McInerney et al., 2018). A wide range of graduates across the country enter the teaching profession. The employers need graduates who can perform well in the classroom, implementing their acquired knowledge, skills, and disposition appropriately. The current study focused on investigating the gaps between acquired and required graduate attributes through MA Education/M.Ed. programs, and the extent to which accreditation of MA Education/M.Ed. programs influence the gaps between acquired and required graduate attributes. The findings disclose varying patterns of gaps between acquired and required graduate attributes. The graduates' perceptions about their acquisition of relevant attributes reveal that their acquired attributes do not meet their perceived requirements on the working place. The findings of Hegji (2017) show that quality assurance organizations should concern themselves about students' acquired skills and learning outcomes because it enables the students to perform professionally with confidence. Li (2015), in her study, suggests that the requirements of graduates should be monitored through accreditation and that the substandard graduation of students should not be encouraged. Many employers both in out outside of Pakistan are reported to be less satisfied with the communication skills, leadership skills, self-motivation, work styles, quality, and analytical skills of graduates (Grant Thornton

Consulting, 2016; Shahbaz, 2016). This research demonstrates that the graduates themselves are less satisfied with their attributes acquired while enrolled in educational programs. Subject knowledge, professional skills, social responsibilities, ethical values, collaboration, and leadership skills are considered prime attributes for the overall development of teachers (Futerman, 2017; Oladokun & Olaleye, 2018; Osmani, Weerakkody, & Hindi, 2017; Shivoro, 2018). Communication skills and effective use of ICT enable the graduates to apply their knowledge effectively for solving teaching and learning problems more professionally (Czerniawski et al., 2018; Rowe, 2017). The current study also reveals differences between graduates' perception about their acquired and required attributes with reference to accredited and non-accredited teacher education programs. The reported gaps between acquired and required attributes of subject knowledge, human growth and development, ethical values, planning, communication, ICT, and English language, are lower among the graduates of accredited programs than the gaps of those who attended non-accredited programs. The findings of this study are consistent with the findings of Williams, Morton, Braun, Longo, and Baker (2017) who establish that accredited programs are better than non-accredited programs in terms of quality and outcomes. The gaps between acquisition and requirement of graduate attributes are higher on accredited programs in environment collaboration, lifelong learning and code of conduct in comparison to the non-accredited program. The minor differences are found between the accredited and non-accredited programs in terms of self-reported acquired and required graduate attribute in assessment.

It is explored in this study through the perceptions of participants that there are admittedly wide gaps between the acquired and required graduate attributes. While, these graduate attributes are identified as essential for the graduates to succeed in the teaching profession. Graduates' acquisition of the attributes of subject knowledge, human growth and development, ethical values, instructional planning and strategies, assessment, learning environment, effective communication, proficient use of information communication technology, collaboration and partnerships, life-long learning, professional code of conduct, and teaching of English as a second/foreign language do not meet employers' requirements to practice in the educational environment. The gaps between the all acquired and required graduate's attributes in both accredited and non-accredited master level teacher education programs were found. Furthermore, the gaps between the acquired and required graduate attributes in environment, collaboration, lifelong learning, and code of conduct are higher in accredited MA/M.Ed. programs.

The accreditation councils and higher education should assure that the graduates of accredited teacher education programs receive the desired attributes to perform well in their professional job place after graduation. It is important that the development of relevant and necessary graduate attributes should be considered a top priority by the accreditors, administrators, and policy makers. A vigorous policy for teaching faculty on continuous professional development, updated and improved teaching skills, stimuli for innovations in curriculum presentations, excellence in research, and appropriate professionalism is required to transform expected graduate

attributes in teacher education. The generalizability of the present study was limited to the sampled participants. It is suggested that a broader scale study in a new research setting should be conducted for a wider generalization of findings. Further research might also be conducted to reduce the gaps between acquired and required graduate attributes by strengthening the accreditation mechanism in the country.

References

Czerniawski, G., Gray, D., MacPhail, A., Bain, Y., Conway, P., & Guberman, A. (2018). The professional learning needs and priorities of higher-education-based teacher educators in England, Ireland and Scotland. *Journal of Education for Teaching*, *44*(2), 133–148. https://doi.org/10.1080/02607476.2017.1422590

Dodhy, M. (2016). *76% of employers not satisfied with the quality of Pakistani graduates, new survey*. Retrieved December 23, 2017, from https://www.techjuice.pk/76-of-employers-not-satisfied-with-the-quality-of-pakistani-graduates-new-survey/

Futerman, R. (2017). *"Design thinking" can prepare graduates for the real working world*. Retrieved August 8, 2017, from http://theconversation.com/design-thinking-can-prepare-graduates-for-the-real-working-world-72509

Glover, P., & Hope, H. (2015). Preparing for tomorrow's world of work. *Technology, Globalisation and the Future of Work in Europe: Essays on Employment in a Digitised Economy*, 42–47.

Government of Pakistan. (2009). *National Professional Standards for Teachers in Pakistan. Policy of Planning wining, Ministry of Education*. Retrieved from http://unesco.org.pk/education/teachereducation/files/National%20Professional%20Standards%20for%20Teachers.pdf

Grant Thornton Consulting (Pvt.) Limited. (2016). *Final Report: Employers' Perception Survey* (pp. 30–36). Islamabad: HEC. Retrieved from http://hec.gov.pk/english/universities/projects/TESP/Documents/Employers%20Perception%20Survey_May%202016.pdf

Hegji, A. (2017). *An Overview of Accreditation of Higher Education in the United States* (No. R43826). Congressional Research Service. Retrieved from https://fas.org/sgp/crs/misc/R43826.pdf

Kaushal, U. (2016). Empowering engineering students through employability skills. *Higher Learning Research Communications*, *6*(4), 1.

Khattak, K. (2016). *Mapping Higher Education in Pakistan*. Retrieved August 31, 2016, from http://www.technologyreview.pk/mapping-higher-education-in-pakistan/

Li, T. (2015). Required graduate attributes in Hong Kong and their relative importance. *European Scientific Journal, ESJ*, *11*(34). Retrieved from http://eujournal.org/index.php/esj/article/view/6747

Marsh, J. D., & De Luca, R. J. (2016). Transformative insights: A synthesis of threshold concept theory and graduate attributes. *Waikato Journal of Education*, *21*(1). Retrieved from http://www.wje.org.nz/index.php/WJE/article/view/208

McInerney, D. M., Korpershoek, H., Wang, H., & Morin, A. J. S. (2018). Teachers' occupational attributes and their psychological wellbeing, job satisfaction, occupational self-concept and quitting intentions. *Teaching and Teacher Education*, *71*, 145–158. https://doi.org/10.1016/j.tate.2017.12.020

Najafi-Sharjabad, F., Hajivandi, A., & Rayani, M. (2014). Knowledge, Attitude, and Practice about Emergency Contraception among Health Staff in Bushehr State, South of Iran. *Global Journal of Health Science*, *6*(1), 52. https://doi.org/10.5539/gjhs.v6n1p52

Nassef, I. (2016). An Investigation into the improvement of graduate attributes within the Egyptian university sector. *International E-Journal of Advances in Education, 2*(4). Retrieved from http://eprints.bournemouth.ac.uk/22560/

National Professional Standards for Teachers in Pakistan. (2009). Islamabad, Government of Pakistan. Retrieved from http://unesco.org.pk/education/teachereducation/files/National%20Prof essional%20Standards%20for%20Teachers.pdf

Naveed, H. (2017). *Quality, not quantity*. Retrieved December 23, 2017, from https://www.dawn.com/news/1334644

Oladokun, T. T., & Olaleye, A. (2018). Bridging skill gap in real estate education in Nigeria. *Pacific Rim Property Research Journal*, *24*(1), 17–34. https://doi.org/10.1080/14445921.2017.1409153

Oliver, B., & Jorre de St Jorre, T. (2018). Graduate attributes for 2020 and beyond: recommendations for Australian higher education providers. *Higher Education Research & Development*, *37*(4), 821–836. https://doi.org/10.1080/07294360.2018.1446415

Osmani, M., Weerakkody, V., & Hindi, N. (2017). Graduate attributes in higher education: Examining academics' perception in the Middle East. *Journal of Education for Business*, *92*(2), 53–64. https://doi.org/10.1080/08832323.2016.1274713

Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, N.J.: Prentice-Hall.

Révai, N. (2018). What Difference Do Standards Make to Educating Teachers? *Organisation for Economic Co-Operation and Development*, 70.

Rowe, A. D. (2017). *Developing graduate employability skills and attributes*: Curriculum enhancement through work-integrated learning, 18.

Shahbaz, M. M. (2016). *Academia-Industry Linkages Gap Analysis* Survey, 2–10. Retrieved from https://propakistani.pk/wp-content/uploads/2016/08/Survey_Report.pdf

Shivoro, R. S. (2018). *A framework for the integration of graduate employability attributes in the curricula of management sciences at higher education institutions in Namibia*. The University of Namibia, Namibia.

Sin, C., Tavares, O., & Amaral, A. (2017). The impact of programme accreditation on Portuguese higher education provision. *Assessment & Evaluation in Higher Education*, *42*(6), 860–871. https://doi.org/10.1080/02602938.2016.1203860

Sonnenschein, K., Barker, M., & Hibbins, R. (2017). Chinese international students' perceptions of and reflections on graduate attributes needed in entry-level positions in the Chinese hotel industry. *Journal of Hospitality and Tourism Management*, *30*, 39–46. https://doi.org/10.1016/j.jhtm.2017.01.008

The Quality Assurance Agency for Higher Education. (2016). *Thematic Report on Enhancement-led Institutional Review (ELIR) Reports 2013-16: Employability and Graduate Attributes* (No. QAA1578) (p. 28). England. Retrieved from http://www.qaa.ac.uk/en/Publications/Documents/Thematic-Report-ELIR-Employability-and-Graduate-Attributes-15.pdf

Williams, S. C., Morton, D. J., Braun, B. I., Longo, B. A., & Baker, D. W. (2017). Comparing Public Quality Ratings for Accredited and Nonaccredited Nursing Homes. *Journal of the American Medical Directors Association*, *18*(1), 24–29. https://doi.org/10.1016/j.jamda.2016.07.025

Evaluation of the Problems Encountered in Public Education Centers*

Taha YAZAR[1], Ozlem LALA[2]

ARTICLE INFO	ABSTRACT
	Purpose: Public Education Centers are important educational institutions where the knowledge and skills that the adults need are gained. The purpose of this research is to determine the problems encountered in Public Education Centers. **Research Methods**: A qualitative research method has been used in this research. The working group of the study is composed of 20 teachers, 46 qualified instructors and 45 trainees who attended courses in Diyarbakır Public Education Centers during the 2016–2017 academic year. The data of the study were collected by a semi-structured interview technique among the interview technique types. Descriptive and content analysis techniques were used in the analysis of the data obtained from the interview.

Findings: According to research findings, the teachers, the qualified instructors and the trainees stated that they had problems due to a "material shortage" in Public Education Centers in general. It was observed that the teachers faced problems related to trainees who registered for the courses but did not come to class regularly and a lack of guidance services. The fact that the qualified instructors are required to work for additional tuition fees without a permanent staff position and that they have problems in finding trainees for courses are among the problems mentioned in the research findings. **Implications for Research and Practice**: The findings of the research show that there is a material shortage in Public Education Centers. The budget allocated to the Public Education Centers needs to be increased to meet the materials needed in the courses.

Introduction

Developments in science and technology, and therefore changes in the lives of individuals, make ongoing education necessary for individuals of all ages. Educational institutions have an important place in bringing society's values, as well as its knowledge and culture into the future. The education of lifelong learners and of productive and creative individuals who question their learning is made possible through the education provided by educational institutions.

Education began with the birth of mankind. In primitive societies, people were engaged in fishing and hunting to maintain their lives on the one hand; and on the other hand, they transferred their knowledge and experiences from their parents and other adults in their surroundings (Gunes, 1996). In this sense, education is a process in which individuals integrate with their societies and learn to embrace their local culture, its values and rules (Guven, 2015). Today, rapid developments in science and technology cause changes in the economic, cultural, political and social structure of society. In order to adapt to these changes, it is compulsory for individuals to gain the knowledge, skills and values that such change requires. The knowledge and skills demanded by these changes are transferred to individuals only until a certain age through formal education (Celep, 2003).

Initially, public education closed the gaps in knowledge among people who were deprived of formal education. More recently, it has been charged with the task of introducing specific knowledge and certain skills to people of all ages (Geray, 1978). Public education aims to provide new knowledge and skills that will enable adults to make changes in their behaviours, thoughts and actions (Bilir, 2004). In Turkey, public education takes the form of a non-formal education entity to reach wider geographic areas and populations (Kurt, 2014).

The first public education unit within the Ministry of National Education (MNE) was established in 1926 as the "Public Decorum Branch". After the adoption of the Latin script alphabet in 1928, "Nation Schools" were opened and in 1932 "Public Houses" were established. The Ministry of National Education was asked to give a new direction to public education in 1951, and reports were prepared with the help of specialists from abroad. In the report prepared by Watson Dickerman, it was suggested that public education activities should be handled in a scientific and systematic manner and that a ministry-affiliated organization should be established. In conformity with the proposal, Turkey's first Public Education Centers were opened in 1956 (Kurt, 2014). The General Directorate of Public Education was established in 1960. In 1977, the courses that provided non-formal education services were collected under the same roof and were named the General Directorate of Non-formal Education (Tasdemirci, 2010).

The concept of adult education, which is expressed in terms of complementary concepts such as public education, non-formal education and continuing education, has emerged as a field of discipline that makes its importance and necessity more and more pressing in the world and Turkey (Kaya, 2015). Adult education programs have the potential to improve local communities and regional economies and

contribute to the global economy (Rith, 2008). Millions of adults all over the world are playing an active role in improving the living conditions of their communities. This participation process is a learning process for them at the same time (Miser, 2013). The fact that adult learning is involved in change is an existing phenomenon in learning theories (Cercone, 2008). In lifelong learning, the goals, processes, organizational structures and pedagogy in institutionalized education at all levels are becoming more and more diverse (Edwards & Usher, 2001). According to lifelong development theories, certain changes occur in a predictable way in an adult's life, and these changes often trigger learning needs (Knowles, Holton & Swanson, 2015). Learning resources are demanded in every environment. The primary task of lifelong learning is to identify these resources and effectively connect them with learners (Knowles, 1996).

Public Education Centers provide important services in the field of adult education (Okcabol, 2006). Adult education includes counselling and support to assist adults in accessing educational services, overcoming learning disabilities, and building both educational and career plans (Smith, 2010). Adult-oriented courses in Public Education Centers consist of two parts, general and vocational/technical departments. These sections are prepared in a mutually supportive manner (MNE, 2010).

Public education is considered one of the important factors in development. However, it can be seen that there is no production-oriented, economy- and development-oriented function in Turkey because public education there cannot be separated from literacy teaching. When the history of public education in Turkey is examined, it can be seen that social, political and economic support for public education was not provided (Turkoglu & Uca, 2011). Determining to what extent the Public Education Centers for adult education have been able to reach their intended targets, which principles they are in compliance with, what causes they pursue, and what impediments prevent them from reaching their targets will lead all institutions serving adult learners to pursue more appropriate targets for future goals and to embrace more appropriate adult education principles (Ural, 2013).

In examining the studies of Public Education Centers in Turkey, it was found that Alabay (2003) determined that teachers and qualified instructors working in Public Education Centers were required to complete the necessary training, that they were weak in the educational dimension, and that their problematic sides were directly attributed to Public Education Center and that they were uncomfortable with this matter. Tanir (2006) found that managers and teachers did not consult with trainees when determining the meeting time of courses, weekly course hours and topics. Elustu (2007) concluded in his research that the awareness level of the adults was low regarding public education and the work done by public education centers, the types of courses they offer, their participation requirements, etc.

When the literature is examined, it is obvious that there are various studies on Public Education Centers, but no research has yet evaluated the teachers, qualified instructors and trainees' opinions together. In this context, it is important that the

Public Education Centers, which have an important role in meeting the training needs of adult individuals in the context of lifelong education, should be evaluated in line with the opinions of teachers, qualified instructors and trainees in the institution. It is thought that the findings of such research will contribute to the development of programs and practices related to Public Education Centers. This research study attempted to determine what problems are encountered in Public Education Centers.

Purpose of the Research

The purpose of this research is to evaluate the problems that teachers, qualified instructors and trainees encounter in Public Education Centers. Depending on the purpose of the research, the following questions were posed:

1. What problems do teachers encounter in Public Education Centers?

2. What problems do qualified instructors encounter in Public Education Centers?

3. What problems do trainees encounter in Public Education Centers?

4. What suggestions can be made to solve the problems encountered by teachers, qualified instructors and trainees in Public Education Centers?

Method

Research Design

A qualitative research method was used in this research. Qualitative data provides first-hand and effective descriptions of the truth (Miles & Huberman, 1994). The data required for qualitative research is usually obtained from field work. During field work, the researcher spends time in the work environment (Patton, 2002). In qualitative research, situations are assessed from the participants' points of view. (Robson, 2015). One qualitative data collection techniques, the semi-structured interview technique, was used in this study. In the semi-structured interview technique, the investigator prepares the questions in advance, but allows for the rearrangement and discussion of the questions by providing partial flexibility to the person surveyed during the interview (Ekiz, 2009).

Research Sample

The working group of the study is composed of 20 teachers, 46 qualified instructors and 45 trainees who attend the courses in Diyarbakır Public Education Centers during the 2016–2017 academic year. The socio-demographic data of teachers, qualified instructors and trainees participating in the research are given below.

Table 1

Surveyed Teachers, Qualified Instructors and Trainees Socio-Demographic Information

	Gender	f	%	Period of service	f	%	Education level	F	%
	Female	9	45	1–5 year	10	50	Bachelor's degree	16	80
	Male	11	55	6–10 year	5	25	Master's degree	4	20
				11–15 year	3	15			
Teacher				16 year and over	2	10			
	Total	20							
	Gender	f	%	Period of service	f	%	Education level	F	%
	Female	37	80	1–5 year	24	52	High school	15	32
	Male	9	20	6–10 year	13	28	Associate degree	22	48
Qualified Instructors				11–15 year	5	11	Bachelor's degree	9	20
				16 year and over	4	9			
	Total	46							
	Gender	f	%	Age	f	%	Marital status	F	%
	Female	32	71	18–25 age	20	44	Married	22	49
Trainees	Male	13	29	26–33 age	9	20	Single	23	51
				34–41 age	8	18			
				42 age and over	8	18			
	Total	45							

Nine of the teachers who participated in the research were female, 11 were male; 37 of the qualified instructors were female, 9 were male; 32 of the trainees were female, 13 were male. Sixteen of the teachers had a bachelor's degree and 4 had a master's degree; 15 of the qualified instructors were high school graduates, 22 had an associate degree and 9 had a bachelor's degree. Teachers and qualified instructors have the highest seniority period of 1–5 years.

In this study, the criterion sampling method was used among the purposeful sampling methods. Purposeful sampling was selected because the selected situations/cases for the study were more informative and enlightening (Christensen, Johnson & Turner, 2015). The basic understanding of the criterion sampling method is the study of all situations that meet a set of predetermined criteria. The criterion or criteria mentioned here can be created by the researcher, or a previously prepared list of criteria can be used (Yildirim & Simsek, 2013). Within the scope of the research, the following criteria were taken in the selection of the interviewees: teachers and qualified instructors should be working only in Public Education Centers, and trainees should be attending the courses regularly and voluntarily participate in the interview.

Research Instruments and Procedures

The data of this study were collected using the semi-structured interview technique. After the study of the literature for the research, an interview form consisting of two open-ended questions was prepared by the researchers. Separate interview forms were organized for teachers, qualified instructors and trainees. The same questions were asked in all three groups so that the problems encountered in Public Education Centers can be obtained and evaluated from a holistic perspective. The interviews at the Public Education Centers lasted thirty minutes. The research data were obtained from the participants' own handwritten responses to the questions on the interview form as a result of the face-to-face interviews the researchers had with voluntarily participating teachers, qualified instructors and trainees.

The following questions were asked to the teachers, qualified instructors and trainees in the interview:

- What problems do you encounter in the public education center?
- What solutions do you suggest for the problems you encounter in the public education center?

Validity and Reliability

In order to ensure the validity of the study, interview questions were presented to two specialists in the field of educational sciences, and the questions were finalized for implementation under the direction of specialists' opinions and recommendations. In order to ensure the reliability of the survey, the specialists examined the interview forms and determined the codes of "consensus" and "dissidence". The reliability of constructed codes was calculated using Miles and Huberman's (1994) reliability formula (Reliability = consensus/consensus+dissidencex×100). In qualitative research, the fit between specialist and researcher evaluations is expected to be 90% and above. As a result of the calculations, the inter-encoder reliability was calculated as 94.

Data Analysis

Qualitative data analysis consists of exporting the meaning of the data: merging, reducing and interpreting what people say, as well as what the researcher sees and reads (Merriam, 2013). Content analysis and descriptive analysis techniques were used in the analysis of the data obtained from the interview. Content analysis is used especially in the analysis of data obtained from observations and interviews (Buyukozturk, Cakmak, Akgun, Karadeniz, & Demirel, 2009). Content analysis is used to identify the presence of words, concepts, themes, idioms, characters or phrases in one or more texts (Kiziltepe, 2015). In descriptive analysis, the aim is to present the findings to the reader in an organized and interpreted way. The data obtained for this purpose are first described systematically and explicitly. Later on, these depictions are explained and interpreted, the cause-effect relations are examined, and some results are reached (Yildirim & Simsek, 2013). In this direction, the data were examined by the researchers, and the meaningful data were coded separately for the teachers, qualified instructors and trainees participating in the

research. For each category created by way of the participants' opinions, direct quotations were drawn from the statements on the interview form. While making quotations reflect the views of the participants, they were coded in this way: the teachers (T8), the qualified instructors (Q5) and the trainees (TR16).

Results

In this section, the findings obtained from the answers of teachers, qualified instructors and trainees, and direct quotations from them are included.

Findings Related to the Problems of Teachers in Public Education Centers

The answers given by the teachers in response to the question, "What problems do you encounter in the public education center?" are presented in Table 2.

Table 2

The Problems Teachers Encounter in Public Education Centers

Themes and Sub-Themes	Frequency (f)
Material Shortages	
Resource inadequacy	7
Copying machine	2
Interactive board	2
Trainee-Induced Problems	
Absenteeism of trainees	4
Discipline	2
Indifference	2
Classroom Problems	
Number of Classrooms	2
Custom language class	1
Multipurpose hall	1
Lack of Guidance Service	
Guidance service	3
Insufficient Course Hours	
Class hours	2

When Table 2 is examined, it appears that the opinions expressed by the teachers about the problems they encounter in Public Education Centers are collected into five themes. These are as follows: "Material Shortages" (11), " Trainee-Induced Problems" (8), "Classroom Problems" (4), "Lack of Guidance Service" (3) and "Insufficient Course Hours" (2).

Some examples of teachers' opinions on "Material Shortages" are as follows:

Trainees should be supported by resources (textbooks) for the examination. We want to test for the reinforcement of the information we have learned during the course, but there is no photocopier." (T10). "I would like to show the shapes and diagrams on a smart board, but I do not have a smart board. The subject does not have narrative books, question banks or trial exams. (T2).

The following expressions are related to "Trainee-Induced Problems":

Trainees show absenteeism problems after they started the course. We cannot bring the trainee in on a regular basis, and the incoming trainees sign in for each other. Although we warn the trainees, they are interested in their phones during the course. (T6).

Some teachers described the "Classroom Problems" as follows:

As an English teacher, I think the lack of a custom language class at the Public Education Center is a major shortcoming because the trainees have very little background in English. For this reason, I need a language class in which I can do activities better. (T8).

Findings Related to the Problems of Qualified Instructors in Public Education Centers

The answers given by the qualified instructors in response to the question, "What problems do you encounter in the public education center?" are presented in Table 3.

Table 3

The Problems Qualified Instructors Encounter in Public Education Centers

Themes and Sub-Themes	Frequency (f)
Material Shortages	
Material Inadequacy	13
Needs of trainees	6
Issues of Permanent Staff	
Absence of permanent staff recruitment	13
Wage Problem	
Inadequate salary	7
Low tuition fees	4
Future Concerns	
Uncertainty about the future	4
Fear of unemployment	2
Unclear date of starting work	2
Module Problem	
Inadequate time to implement the module	6
Insurance Problems	
Unpaid insurance during summer vacation	4
Insurance in return for additional courses	2
Paperwork Problem	
Unnecessary paperwork	4
Change of class books per module	1
Appreciation problems	
Communication problems with administrators	2
Absence of equal treatment	2
Lack of appreciation	1
The Problem of Employee Rights	
Inadequate employee rights	2
Elimination of severance pay	1
Education-training compensation	1
Development Problems	
Development in the field	3
Innovation by module	1
Trainee Finding Problem	
Finding trainees	2

As shown in Table 3, the qualified instructors have more problems in public education centers due to "Material Shortages" (19). In addition, they emphasized "Issues of Permanent Staff" (13) "Wage Problem" (11) and "Future Concerns" (8) as problems they encountered in public education centers.

Here are some example statements by the qualified instructors who gave an opinion on "Material Shortages":

"The most important problem we face in the public education center is the lack of material in the vocational education classes. The situation of qualified instructors to supply materials with their own means causes us to have financial problems. "(Q1).

"There are problems with the material. When there is no material, trainees do not come and we have trouble producing products."(Q39).

The following expressions are related to the "Issue of Permanent Staff":

"We, qualified instructors, have been working for years, but we do not receive compensation for our labour. We do not have any qualification as instructors." (Q38).

"Before all else, we have a permanent staff problem. We have been working hard for years, but we are not staffed permanently."(Q11).

The statements regarding the "Wage Problem" are as follows:

"We do all the work required by the course with the trainees, but I do not think I get paid enough." (Q8).

"It is not clear how much we will get each month as we work for additional tuition fees. Additional tuition fees are not paid, especially on official holidays. "(Q43).

Some qualified instructors described their "Future Concerns" as follows:

"In the public education centers, the uncertainty of the length of the summer holidays (the start date of work after the holiday is not clear), and the gap between the courses we give causes us to worry about the continuity of the work." (Q2).

"Qualified instructors are concerned about the future. At the beginning of the semester, we are uncertain about the starting date, and we are afraid of dismissal at the end of the course. "(Q6).

Findings Related to the Problems of the Trainees in Public Education Centers

The answers given by the trainees in response to the question, "What problems do you encounter in the public education center?" are presented in Table 4.

Table 4

The Problems Trainees Encounter in Public Education Centers

Themes and Sub-Themes	Frequency(f)
Material Shortages	
Computer shortage	13
Clothing material	6
Instruments	4
Brush and paint	3
Problems with Canteen	
Insufficient seating area	4
Food Prices	3
Need for Education in Different Fields	
New and different fields	3
Cleaning Issue	
Hygiene problem	3
Certificate Fee	
Giving the certificate for the fee	3
Classroom Area	
Small classrooms	3

When Table 4 is examined, the trainees stated that they had problems due to "Material Shortages" (26) in the public education centers. The other problems in the public education centers that the trainees noted are listed as "Problems with Canteen", "Need for Education in Different Fields ", "Cleaning Problem", "Certificate Fee" and "Classroom Area".

Here are some example quotations from the trainees on "Material Shortages":

I have difficulty obtaining the materials required for the course. I have no good financial situation, and I have children studying at school. I cannot even get a paintbrush. But I need materials to be able to practice and have a profession."(TR41). "We have difficulties in supplying materials. There are deficiencies in the clothing department. We need to wait for each other in order to be able to practice. (TR1).

Some statements by the trainees describe the "Problems with Canteen" as follows:

"The canteen is small, we can hardly find a seating area between classes. There is no room for us to have a seat and drink tea with our course friends." (TR9).

"I find the prices of food and drinks in the canteen high. The courses offered in public education are free, but canteen prices are expensive." (TR30).

A trainee explained the "Need for Education in Different Fields" :

There are no courses in different fields at the public education center. We need new courses for changing conditions and needs. As for the language course, only English comes to mind. However, I need a course in Arabic; therefore, I think there is a very little variety of courses.(TR37).

Findings of Teachers' Suggestions for Solving the Problems Encountered in Public Education Centers

The responses of the teachers to the question, "What solutions do you suggest for the problems you encounter at the Public Education Center?" are given in Table 5.

Table 5

Teachers' Suggestions for Solving the Problems in the Public Education Centers

Themes and Sub-Themes	Frequency(f)
Solutions to Material Shortages	
Communication with MNE	7
Adequate lending to courses	4
Solutions to Trainee-Induced Problems	
Guidance for trainees	5
Determination and application of disciplinary rules	3
Solutions to the Classroom Problem	
Increasing classroom numbers	3
Buildings prepared by expert opinion	1
Solution to Guidance Service Deficit	
Presence of guidance counsellors	3
Solution to Insufficient Course Time	
Increasing class hours	2

When the teachers' suggestions are examined, it can be said that the teachers considered the MNE (Ministry of National Education) to be effective in solving the problems they encountered in the Public Education Centers, such as allocating sufficient funds to the courses, increasing the number of classrooms and keeping guidance counsellors in the institution.

Some examples of teachers' suggestions about solutions to "Material Shortages" are as follows:

"Communication with MNE may meet the required materials." (T1).

"The problem of materials (tests, trial exams, copying machine) can be solved by allocating more funds to the courses." (T10).

The following expressions are related to the solutions to "Trainee-Induced Problems":

" Guidance can be provided for the trainees in the institution, and the absenteeism problem can be abolished." (T20).

"Trainees should not take mobile phones into the classroom . Disciplinary rules in schools can also be applied in Public Education Centers." (T6).

The statements regarding the solutions to the "Classroom Problem" are as follows:

"Buildings and classrooms should be built with a more professional point of view to give continuous service." (T5).

"The number of classrooms in public education centers should be increased." (T19).

Findings of Qualified Instructors' Suggestions for Solving the Problems Encountered in Public Education Centers

The suggestions by the qualified instructors in response to the question "What solutions do you suggest to the problems you encounter at the Public Education Center?" are given in Table 6.

Table 6

Qualified Instructors' Suggestions for Solving the Problems Encountered in Public Education Centers

Themes and Sub-Themes	Frequency(f)
Solutions to Material Shortages	
Authority work	13
Cooperation with institutions	6
Solution of Permanent Staff Issue	
Effective studies of unions	8
Addressing of the problem by authorities	5
Solution to Wage Problem	
Reform of the relevant regulation	7
Improvement of tuition fees	4
Solution to Future Concerns	
Contract work	5
Certain star date of course	3
Solution to Module Problem	
Extension of course module durations	7
Solution to Insurance Problem	
Annual contract	6
Solution to Paperwork Issues	
Extension of the period of book control	4
Solutions to Appreciation Problem	
Fairness	3
Effective communication	2
Solution to Employee Rights	
Regulation of employee rights	4
Solutions to Development Problem	
Giving in-service courses	3
Awarding successful qualified instructors	1
Solution to Trainee Finding Problem	
Effective publicity of courses	2

In Table 6, it is obvious that the qualified instructors made suggestions for the ministry, the authorities and unions to do their part to solve the problems in Public Education Centers.

Here is a suggestion from a qualified instructor for a solution to the "Material Shortage":

"Authorities should come together for the provision of materials and product support for the trainees and cooperate with the necessary institutions in this regard." (Q12)

An opinion on the solution to the "Permanent Staff Issue" is as follows:

"It is possible to solve this problem if the unionists work effectively on the issue of permanent staff and be more determined about the permanent recruitment of qualified instructors that can provide certain criteria (service year, etc.)." (Q11).

The following expressions point out solutions to the "Wage Problem":

"A step may be taken to change the regulation to increase tuition fees."(Q24).

"An arrangement can be made on tuition fees."(Q15).

The following quotation describes a solution to the "Future Concerns":

"A periodic or annual contract can be made with qualified instructors. This will ensure that they are guaranteed, so they will not feel temporary. "(Q6).

Findings of Trainees' Suggestions for Solving the Problems Encountered in Public Education Centers

The suggestions of the trainees for the question "What solutions do you suggest for the problems you encounter at the Public Education Center?" are given in Table 7.

Table 7

Trainees' Suggestions for Solving the Problems Encountered in Public Education Centers

Themes and Sub-Themes	Frequency(f)
Solutions to Material Shortages	
Supply of materials	17
Getting new computers	12
Solutions to the Problems with Canteen	
Creation of seating area	4
Reduction of food prices	3
Solution to Need for Education in Different Fields	
Taking the views of trainees	3
Solution to the Cleaning Issue	
More frequent cleaning	3
Solution to the Certificate Fee Problem	
Removing the certificate fee	3
Solution to Classroom Area Problem	
Expansion of the classroom area	3

For the solution to the various problems, the trainees emphasized that opinions of trainees should be asked in order to supply the materials they need, to improve the physical conditions of the school, and to determine the courses to be offered.

Here are some example statements by the trainees on solutions to "Material Shortages":

"The necessary materials for the courses should be provided at the public education centers to provide the trainees with convenience." (TR26).

"Computers must be updated with better and higher-processor versions."(TR4).

An opinion on the solution to "Problems with Canteen" is as follows:

"A seating area for the canteen should be created, and the food prices should be discounted considering the financial situation of the trainees." (TR10).

Regarding the problem of "Need for Education in Different Fields", one trainee stated:

"The opinions of trainees can be taken, and they can be asked whether they need courses for different requirements." (TR37).

Discussion, Conclusion and Recommendations

In the research findings, teachers, qualified instructors and trainees stated that they have problems due to "Material Shortages" in Public Education Centers in general. In order to increase the effectiveness of the education given in Public Education Centers and to be able to practise, more tools and equipment are needed. The use of materials provides an important contribution to achieving the goal of teaching, but the point that needs to be emphasized here is the necessity of meeting the material requirements of the Public Education Centers. Because, in the case of research findings, the trainees cited that it is difficult to obtain the necessary materials for the courses. In the study conducted by Asir (2011), it was concluded that the course costs affected the program choice, and the money that was to be spent on the course materials prevented the high cost courses from being taken. According to the research findings of Tezcan (2012), it was determined that the leading situation in which purpose-oriented adult learners had negative considerations about the courses involved the material usage in the courses. In the research by Temiz (2009), it was explained that the Public Education Centers had problems with the personnel, equipment and tools. These findings support the findings of the research.

In the present research findings, it has become apparent that the teachers had problems due to trainees who registered the courses but did not come regularly, who were irrelevant to the courses, who showed irregular behaviours, and lack of guidance service in the course. The services provided by the guidance counsellors in the Public Education Centers are identified as ensuring that trainees are informed about the methods needed to be successful in their chosen program, providing motivation for continuing and succeeding in the courses and programs they are enrolled in, and informing them about the acquisition they have documented (MNE, 2011). In this direction, it is possible to say that the problems caused by the trainees are being experienced due to the inadequacy of the guidance service offered to the

trainees in Public Education Centers. In the study conducted by Alabay (2003), the teachers in the Public Education Center stated that the quality of education should be increased by making the necessary arrangements by the administrators' to prevent the trainees' absenteeism. This finding is parallel to the result of the present research.

That qualified instructors work without permanent staff positions and for additional tuition fees is one of the leading problems mentioned in the research findings. In the guideline issued by the MNE (2011) to implement public education activities, it is explained that the need will be covered by paid specialists and qualified instructors who will work for the additional tuition fee if there are not enough teachers or permanent qualified instructors in Public Education Centers. Qualified instructors who work in a similar way alongside their permanent colleagues in the implementation of the course programs, and who contribute to the economy by educating the required qualifying members, expressed that they should be given a permanent position. In the research conducted by Alabay (2003), qualified instructors expressed that they wanted the permanent position and assignment to be done according to years of service. It can be argued that this finding overlaps with the result of our research.

Another finding that is obtained in the research is that qualified instructors are not appreciated in Public Education Centers and they feel unimportant. The courses taught by qualified instructors do not have continuity as they last for only a certain period of time. For this reason, it can be considered that there is incomplete communication between qualified instructors and institution managers. In his study, Alabay (2003) concluded that qualified instructors believed that they were not valued and recognised. This finding is parallel to the results of the present research.

In the research findings, it becomes evident that qualified instructors have problems improving themselves, and in this sense, they need in-service training. It is necessary for qualified instructors to follow the developments in their professions, to acquire new knowledge and technologies, and to train with in-service courses. In the study conducted by Yayla (2002), it was concluded that qualified instructors needed training to develop competences. This finding supports the results of our study.

According to our research findings, qualified instructors stated that they had problems in finding trainees for the courses in Public Education Centers. Elustu (2007) found that the adults viewed the Public Education Centers as places where literacy training was given. Yet there are courses in many fields (social, vocational and technical) to meet the diverse needs of adults in the Public Education Centers. In this respect, the Public Education Centers can provide publicity about the courses given to the adults in their regions, and develop awareness in adults, which can contribute to the solution to the qualified instructors' problem of enrolling trainees.

In the research findings, it has been seen that the trainees have problems with the physical conditions and the certificate fee taken at the end of the course. According to Turkoglu and Uca (2011), in public education activities, which are considered free in Turkey, the demand for remuneration from people causes the adults to move away from public education activities; training taking place in unhealthy places frightens

away adults who are in need of education and prevents the public education courses from reaching their goals. In Ata's (2016) survey, trainees who participated in vocational courses at the Public Education Center suggested that the physical conditions of the courses should primarily be improved and the certificate fees should not be taken because of its deterring effect. It can be argued that this finding overlaps with the results of our research.

In the present research findings, the trainees stated that new courses are not available in different fields to meet their needs. Today's developments in social, economic and technological life show that the adults' need for education changes according to the present circumstances. In this direction, the content of the programs given to the adults in the Public Education Centers should also be rearranged according to the needs of the age to meet the needs of individuals and society. The research findings of Elustu (2007) show that the education needs of the adults are various, but the Public Education Centers concentrate on some programs and are not oriented towards opening new education programs. Tanir (2006) found that while the courses in the Public Education Centers were being planned, offering courses according to the teachers and the existing branches was preferred to creating courses that people needed. Trainees, in the survey conducted by Yancar (2014), implied that more courses should be opened. These findings are parallel to the results of the present study.

Under the direction of the findings obtained in this research, the budget allocated to the Public Education Centers should be increased to meet the materials needed in the courses. Problems related to the permanent staff position, wages and employee rights of qualified instructors should also be addressed. Necessary work can be done to improve the working conditions of Public Education Centers. Finally, programs in the Public Education Centers should be diversified as required by today's conditions.

References

Alabay, N. (2003). *Kirikkale ilinde halk egitim merkezinde gorev yapan ogretmen ve usta ogreticilerin mezun olduklari kaynaklari ve halk egitim merkezi icerisindeki sorunlari*. [Graduation sources of teachers and qualified instructors serving in public education center in kirikkale and their problems in public education center] (Unpublished master thesis). Gazi University. Ankara.

Asir, S.K. (2011). *Mamak halk egitim merkezinde acilan kurslara katilan kadinlarin egitim egitime katilma ve terk nedenleri*. [The reasons for the participation and abandonment of the female trainees in the courses opened at mamak public education center] (Unpublished doctoral thesis). Ankara University. Ankara.

Ata, E. (2016). *Baskent halk egitim merkezindeki mesleki kurslara yetiskinlerin katilim oruntuleri*. [Adult participation patterns in vocational courses in the baskent public education center]. (Unpublished master thesis). Ankara University. Ankara.

Bilir, M. (2004). *Kentlesme sürecinde yetiskinlerin egitim gereksinimi: Ankara mamak ilcesi, tuzlucayır semti, ege mahallesi ornegi.* [Training needs of adults in urbanization process: Ankara mamak district, tuzlucayır district, ege neighborhood sample]. *Ankara Universitesi Egitim Bilimleri Fakültesi Dergisi, 37*(1), 98-121.

Buyukozturk, S., Cakmak, E K.., Akgun, O.E., Karadeniz, S., & Demirel, F. (2009). *Bilimsel arastirma yontemleri.* [Scientific research methods]. Ankara: Pegem Yayincilik.

Celep, C. (2003). *Halk egitimi.* [Public education]. Ankara: Ani Yayincilik.

Cercone, K. (2008). Characteristics of adult learners with implications for online learning Design. *AACE Journal, 16*(2), 137-159.

Christensen, L.B., Johnson, R.B., & Turner, L.A. (2015). *Research methods design and analysis.* Ahmet Aypay (Translated). Ankara: Ani Yayincilik.

Edwars, R. & Usher, R. (2001).Lifelong learning: A postmodern condition of education?*American Association for Adult and Continuing Education*, 51(4), 273-287.

Ekiz, D. (2009). *Bilimsel arastirma yöntemleri.* [Scientific research methods]. Ankara: Ani Yayincilik.

Elustu, A. (2007). *Yetiskinlerin egitim ihtiyaclari ve halk egitimi ile ilgili farkindalik duzeylerinin belirlenmesi.* [Determination of education needs and awareness level of the adults for public education]. (Unpublished master thesis). Marmara University. Istanbul.

Geray, C. (1978). *Halk egitimi.* [Public education]. Ankara: Ankara Universitesi.

Gunes, F. (1996). *Yetiskin egitimi* (Halk egitimi). [Adult education (Public education)].Ankara: Ocak Yayinlari.

Guven, I. (2015). *Turk egitim tarihi.* [History of Turkish education]. Ankara: Pegem Yayincilik.

Kaya, E. (2015). Turkiye'de halk egitimi merkezleri. [Public education centers in Turkey]. *International Journal of Science Culture and Sport (IntJSCS)*, 3, 268-277.

Kiziltepe, Z. (2015). *Icerik analizi. Nitel arastırma yontem, teknik, analiz ve yaklasimlari* [Content analysis. Qualitative research method, technique, analysis and approaches.] F. N. Seggie & Y.Bayyurt (Ed.). Ankara: Ani Yayincilik.

Knowles, M. (1996). *Adult earners: A neglected species.* (S .Ayhan Translated). Ankara: Ankara University.

Knowles, M., Holton, E., & Swanson, R. (2015). *Adult education.* Okhan Gunduz (Translated). Istanbul: Kaknus Yayinlari.

Kurt, I. (2014). *Yetiskin egitimi.* [Adult education]. Ankara: Akçag Yayinlari.

MNE (2010). Yaygin egitim kurumlari yonetmeligi. [Regulation on non - formal education institutions]. *Official Newspaper*.

MNE (2011). Halk egitimi faaliyetlerinin uygulanmasina dair yonerge (2645 Sayili). [Guideline on the implementation of public education activities (No. 2645)]. *Magazine of papers*, 74.

Merriam, S.B. (2013). *A guide for qualitative research design and implementation* (S. Turan Translated). Ankara: Nobel Yayincilik.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis. Thousand oaks,CA:* Sage.

Miser, R. (2013). Training of local communities. In A. Yıldız & M. Uysal (Ed.), A*dult education* (pp. 273-285). Istanbul: Kalkedon Yayinlari.

Okcabol, R. (2006). *Halk egitimi*. [Public education]. Ankara: Utopya Yayinlari.

Patton, M.Q.(2002). *Qualitative research & evaluation methods.* Thousand oaks, calif: Sage.

Robson, C. (2015). *Real world research.* (S. Cinkir & N. D. Kasimoglu Translated). Ankara: Ani Yayincilik.

Rith, E. (2008). Redefining tradition: adult learners and higher education. *Adult Learning*, 19(1-2), 12-16.

Smith, C.S. (2010). The great dilemma of improving teacher quality in adult learning and literacy. *Adult Basic Education and Literacy*. 4(2), 67.

Tanir, M. (2006). *Turkiye'de halk egitim hizmetlerinin yonetimi ve halk egitim merkezlerinin etkiligi uzerine alan arastirmasi*. [A survey on the management of public education services and the impact of public education centers in Turkey]. (Unpublished master thesis). Gazi Universitesi. Ankara.

Tasdemirci, E. (2010). *Türk egitim tarihi*. [History of turkish education]. Ankara: Gunduz Yayincilik.

Temiz, M. (2009). *Halk egitim merkezlerinde gorevli mesleki - teknik ve sosyal-kulturelkurslari ogretmenlerinin oz yeterlilik duzeylerinin incelenmesi.* [Examination of self-sufficiency levels of teachers at vocational-technical and social-cultural courses in public training centers]. (Unpublished master thesis). Marmara Universitesi. Istanbul.

Tezcan, F. (2012). *Halk egitimi merkezi kurslarina katilan yetiksin ogrenenlerin guduselyonelimleri (Mugla ili merkez ilce halk egitimi merkezi kurslar ornegi).* [Motivational orientations of adult learners participating in public education center courses (Mugla provincial center for public education courses)]. (Unpublished doctoral thesis). Ankara Universitesi. Ankara.

Turkoglu, A., & Uca, S. (2011). Turkiye'de halk egitimi: tarihsel gelisimi, sorunlari ve cozum onerileri. [Public education in Turkey: historical development, problems and solution suggestions]. *Adnan Menderes Universitesi Egitim Bilimleri Dergisi,* 2, 48-62.

Ural, O. (2013). Yetiskin egitiminde program alanlari. In A. Yildiz & M.Uysal (Ed.), A*dult education* (pp. 177-198)]. Istanbul: Kalkedon Yayinlari.

Yancar, M. (2014). *Halk egitim merkezlerinde duzenlenen mesleki ve teknik kurslara devam eden yetiskinlerin bu kurslari tercih etme nedenleri ve kursa iliskin beklentileri (Menemen ornegi).* [Reasons of adults for choosing vocational and technical courses in public education centers and their expectations of the courses (Menemen's case)]. (Unpublished master thesis). Okan University. Istanbul.

Yayla, D. (2002). *Ucretli usta ogreticilerin yetiskin egitimi acısından yeterlik duzyleri.* [Qualifying levels of paid qualified instructors in terms of adult education]. (Unpublished master thesis). Ankara Universitesi. Ankara.

Yildirim, A., & Simsek, H. (2013). *Sosyal bilimlerde nitel arastirma yontemleri.* [Qualitative research methods in the social sciences]. Ankara: Seckin Yayincilik.

Halk Eğitim Merkezlerinde Karşılaşılan Sorunların Değerlendirilmesi

Atıf:

Özet

Problem Durumu: Halk Eğitim Merkezleri bireylerin kişisel, mesleki ve sosyal gelişimlerine yardımcı olan, öğrenmeye, üretmeye, kendisine ve çevresine faydalı birer birey olmalarına katkı sağlayan çok önemli yaygın eğitim kurumlarıdır. İlk başta okuma yazma kursları olarak faaliyet gösteren ancak değişen sosyal, ekonomik ve teknolojik yaşamla birlikte ortaya çıkan farklı ihtiyaçları karşılamak üzere çeşitli alanlarda eğitimler veren bir kurum haline gelmiştir. Halk Eğitim Merkezlerinde verilen kurslar aracığı ile birçok kişi meslek sahibi olmakta ve ekonomiye katkı sağlamaktadır. Yetişkinlere yönelik önemli çalışmalar yapan Halk Eğitim Merkezlerinin amaçlarını daha iyi gerçekleştirmesi için sorunların belirlenmesi ve bu doğrultuda çözüm önerilerinin geliştirilmesi gerekmektedir.

Araştırmanın Amacı: Bu araştırmanın amacı; öğretmen, usta öğretici ve kursiyerlerin Halk Eğitim Merkezlerinde karşılaştıkları sorunları değerlendirmektir. Araştırmanın amacına bağlı olarak şu sorulara yanıt aranmıştır:

1. Öğretmenlerin Halk Eğitim Merkezlerinde karşılaştıkları sorunlar nelerdir?

2. Usta öğreticilerin Halk Eğitim Merkezlerinde karşılaştıkları sorunlar nelerdir?

3. Kursiyerlerin Halk Eğitim Merkezlerinde karşılaştıkları sorunlar nelerdir?

4. Öğretmen, usta öğretici ve kursiyerlerin Halk Eğitim Merkezlerinde karşılaştıkları sorunların çözümüne yönelik önerileri nelerdir?

Araştırmanın Yöntemi: Bu araştırmada nitel araştırma yönteminden yararlanılmıştır. Araştırmanın çalışma grubunu 2016-2017 eğitim-öğretim yılında Diyarbakır Halk Eğitim Merkezlerinde görevli 20 öğretmen, 46 usta öğretici ve bu kurslara devam eden 45 kursiyer oluşturmaktadır. Bu araştırmada amaçlı örnekleme yöntemlerinden ölçüt örnekleme yöntemi kullanılmıştır. Araştırma kapsamında görüşme yapılan öğretmen ve usta öğreticilerin belirlenmesinde sadece Halk Eğitim Merkezlerinde çalışıyor olmaları; kursiyerlerin belirlenmesinde ise katıldıkları kursa düzenli devam etmeleri ve bu araştırma için yapılan görüşmeye gönüllü olarak katılmaları ölçüt olarak alınmıştır. Araştırmanın verileri, görüşme tekniği türlerinden yarı-yapılandırılmış görüşme tekniği ile toplanmıştır. Araştırmacılar tarafından açık uçlu iki sorudan oluşan bir görüşme formu hazırlanmıştır. Araştırmanın geçerliliğini sağlamak için eğitim bilimleri alanında iki uzmana bu sorular sunulmuş, uzmanların görüş ve önerileri doğrultusunda sorular uygulama için son şeklini almıştır. Araştırma verileri, araştırmaya gönüllü olarak katılan öğretmen, usta öğretici ve kursiyerlerle, araştırmacıların yüz yüze gerçekleştirdiği görüşmeler sonucu, katılımcıların kendi el yazılarıyla görüşme formundaki soruları cevaplandırmalarıyla elde edilmiştir. Görüşme sonucunda elde edilen verilerin analizinde betimsel ve içerik analiz tekniklerinden yararlanılmıştır. Bu doğrultuda veriler araştırmacılar tarafından incelenmiş, anlamlı veriler araştırmaya katılan öğretmen, usta öğretici ve kursiyerler için ayrı ayrı kodlanarak kategoriler oluşturulmuştur. Katılımcıların görüşlerinden yola çıkılarak oluşturulan her kategori için, görüşme formundaki ifadelerden doğrudan alıntılar yapılmıştır. Katılımcıların görüşlerini yansıtmak için alıntılar yapılırken, öğretmenler için (Ö8); usta öğreticiler için (UÖ5) ve kursiyerler için (K16) şeklinde kodlama yapılmıştır. Araştırmanın geçerliliğini sağlamak için eğitim bilimleri alanında iki uzmana görüşme soruları sunulmuş, uzmanların görüş ve önerileri doğrultusunda sorular uygulama için son şeklini almıştır.

Araştırmanın Bulguları: Öğretmenlerin, Halk Eğitim Merkezlerinde gerekli materyali bulma, öğrencilerin kurslara düzenli gelmemeleri, rehberlik hizmeti eksikliği ve yetersiz ders saati nedeniyle sorun yaşadıkları görülmüştür. Usta öğreticiler, Halk Eğitim Merkezlerinde gerekli materyali bulma, kadrosuz çalışma ve çalışmaları karşılığında düşük ücret alma, işle ilgili gelecek kaygısı taşıma, kurs modül sürelerinin azlığı, tatillerde kesilen sigorta, gereksiz evrak işleri, kurumda değer görmeme, özlük haklarının yetersiz oluşu, kendi alanında gelişim göstermeme ve kurslara yeterli öğrenci bulamama nedeniyle sorun yaşadıklarını belirtmişlerdir. Kursiyerler, Halk Eğitim Merkezlerinde ihtiyaç duydukları materyalin bulunmadığını, kantinin küçük olması nedeniyle oturma için yer sıkıntısı çekildiğini, farklı alanlara yönelik kursların bulunmadığını, temizliğe yeterince önem verilmediğini, kurs bitiminde ücret karşılığında sertifikaların alındığını ve derslik alanının yetersiz olduğunu belirtmişlerdir.

Araştırmanın Sonuçları ve Önerileri: Araştırma sonucunda Halk Eğitim Merkezlerinde öğretmenler, usta öğreticiler ve kursiyerler için yeterli materyalin olmadığı görülmüştür. Öğretmen ve usta öğreticilerin derslerini etkili bir biçimde sunabilmeleri ve öğrenmeyi kolaylaştırabilmeleri için gerekli materyale ihtiyaçları vardır. Materyal gereksinimini kursiyerlerin karşılayamadığı ve bu nedenle sorun yaşadıkları görülmüştür. Bu doğrultuda materyal sorununun ortadan kalkması için Halk Eğitim Merkezlerine ayrılan ödeneğin arttırılması ve ilgili kurumlarla işbirliği yapılması yoluna gidilebilir. Öğretmenlerin, kursa kayıt yaptıran ancak düzenli gelmeyen, ders esnasında ilgisiz olan ve tam olarak hangi alanda eğitim ihtiyacı olduğunu bilmeyen öğrenciler nedeniyle sorun yaşadıkları belirlenmiştir. Halk Eğitim Merkezlerinde rehberlik hizmetinin bulunması, öğrencileri ilgi ve ihtiyaçları doğrultusunda yönlendirmede etkili olacağı ve öğrenciden kaynaklı diğer sorunları çözümünde de etkili olacağı söylenebilir. Usta öğreticiler Halk Eğitim Merkezlerinde kadrolu olarak çalışmak istediklerini belirtmişlerdir. Araştırma bulguları incelendiğinde usta öğreticilerin ücret, gelecek kaygısı, değer görmeme ve sigorta sorununun temelinde Halk Eğitim Merkezlerinde kadrosuz olarak çalışmalarının neden olduğu söylenebilir. Bu bağlamda kadrosuz usta öğreticilerin kadroya alınması konusunda yetkililerin bir çalışma yapması sağlanabilir. Araştırma bulgularında kursiyerlerin ders aralarında oturmaları için kantin alanının yetersiz olduğu ve derslik alanının küçük olduğu görülmüştür. Yetişkinlere yönelik eğitim veren Halk Eğitim Merkezlerinin binaları planlanırken, yetişkinlerin fiziki ihtiyaçlarının göz önünde bulundurulması gerekmektedir. Değişen sosyal, kültürel, ekonomik ve toplumsal yaşamla birlikte kursiyerlerin farklı alanlarda eğitim ihtiyacı ortaya çıktığı araştırma bulgularında görülmüştür. Bu nedenle çağın ve toplumun gereksinimlerine uygun farklı alanlarda kursların verilmesinin faydalı olacağı söylenebilir.

Anahtar Kelimeler: Halk eğitimi, yaşam boyu eğitim, yetişkin eğitimi, yaygın eğitim.

The Effect of Digital Texts on Primary Students' Comprehension, Fluency, and Attitude*

Safak KAMAN[1], Ihsan Seyit ERTEM[2]

ARTICLE INFO	ABSTRACT
	Purpose: In today's world, the rapid spread of technology affects educational life. It has become common place in many countries to read digital texts on tablet computers, preferred for their portability, and long charging times; projects are done to improve educational quality in many countries. However, research is limited about how reading digital texts affects reading skills of primary students. The aim of this study is to investigate the effect of digital text readings on reading comprehension, reading fluency, and reading attitude of 4th grade students.

Research Methods: The mixed method, in which quantitative and qualitative research methods are used together, has been employed in the research. The reading levels of 75 students attending 4th grade in four primary schools were determined by pre-test, and the 30 students with the lowest scores were selected. An experimental and a control group were formed by randomly assigning the students to one of the two groups. Each group contained 15 students. **Findings:** The quantitative findings showed that the use of digital texts had influence on improving fluency and reducing reading mistakes, while it did not affect students' reading attitudes and its effect on comprehension lasted a short time. According to qualitative findings, students were eager and excited. **Implication for Research and Practice:** If educational environments are inevitably influenced by technological devices, it becomes crucial to use technological devices in a planned and appropriate manner. We suggest that teachers also apply reading activities with digital texts to improve the reading achievement of primary students.

Introduction

New technologies are increasingly prevalent in all parts of life throughout the world. Mobile devices in particular are used for learning purposes both in and out of school, and they have been quite beneficial (Wagner, Castillo, Murphy, Crofton, & Zahra, 2014). Tablet computers especially have a significant role in increasing the effectiveness and efficiency of educational activities (Gungoren, Bektas, Ozturk & Horzum, 2014). In this regard, education and the use of technology in education have been two concepts that are not considered separately from each other (McCannon & Crews, 2000; Komis, Ergazakia & Zogzaa, 2007). Using technology has brought many benefits to educational environments, such as improving students' critical thinking (Jonassen, 1999), making their preferred learning models more effective (Quintina & Donovan, 2013), and removing the limitation of only written words and images in traditional texts (Larson, 2010). These advantages should be made available to students.

Students in advanced societies are interacting with information and communication technologies. While technology and the internet have continued to change the learning and development of students, educators can use various materials in reading education, from digital texts to the internets. (Thoermer & Williams, 2012). Students are increasingly taking advantage of the technological materials in learning to read and write (Walsh & Simpson, 2013), both outside of school (Jahson, Adams & Cummins, 2012) and within the classroom (Akbar, Taqi, Dashti & Sadeq, 2015). If students have a choice, they prefer to do these activities on the internet or on the phone instead of listening to the teacher or reading the written materials in the classroom (Tolani, McCormac & Zimmermann, 2009). In this regard, technology is becoming an indispensable part of education.

As mentioned above, researches support this situation. When the related studies were examined, Leu, O'Byrne, Zawilinski, McVerry and Everett-Cacopardo (2009) found that students improved their reading skills if they were constantly using web-based literacy activities; Sackstein, Spark and Jenkins (2015) ascertained that most of the participants read faster with the iPad as a result of their work and VanAken (2014) also found that reading e-books or printed texts influenced reading attitudes and reading level, but not reading skills. In the lights of these findings, it is clear that technology does not only make it possible to reach out to the rest of the world, but has become necessary as a tool in education and training (Luke, 2012). Furthermore, motivation and reading fluency of learners are improving in educational environments where digital texts are used (Thoermer & Williams, 2012). It can be seen that technological tools positively affect reading skills.

When we also review the relevant literature in our country, we see that no research reveals the effect of digital texts on reading fluency, reading comprehension and reading attitudes of elementary school students. In this regard, A time when reading comprehension skills are so important for students, it is thought that determining level of reading comprehension skills and reading fluency skills of students who regularly perform reading activities on tablet computers will contribute to the field.

The main purpose of the study is to answer the question "What is the effect of digital texts on the reading comprehension, reading fluency and reading attitude of 4th grade primary school students?" In accordance with this purpose, this study aims to answer the following research questions:

1. Is there a significant difference between reading comprehension scores of the experimental group of students using digital texts and the control group students using printed texts?

2. Is there a significant difference between fluent reading skills scores of the experimental group of students using digital texts and the control group of students using the printed texts?

3. Is there a significant difference between the reading attitude scale scores of the experimental group students using digital texts and the control group students using the printed texts?

4. Is there a significant difference in correcting reading errors between the experimental group of students using digital texts and the control group of students using printed texts?

5. What are the opinions of students in the experimental group on using digital texts in the reading process?

   • The situation of facilitating reading and comprehension,

   • The situation of liking and resuming reading,

   • The situation of facilitating / complicating reading process.

Method

Research Design

This research uses the convergent parallel mixed method, in which quantitative and qualitative research methods are used together. Mixed methods contribute greatly to the researcher's ability to minimize the bias that can arise from the nature of the researcher or from the nature of the research in the research process. Furthermore, diversifying the data sources in the study allows the results to be more valid (Yildirim, 2010). The mixed method involves collecting and analyzing qualitative and quantitative data together during the research process (Creswell, 2003). The reason for choosing the mixed method is to make the results more understandable by supporting the quantitative data with qualitative data.

Research Sample

The study was conducted with seventy five-4th grade students, who studied in four different elementary schools selected among the eight primary schools and who were in the same socio-economic status in the Kaman county center of Kirsehir province in the fall semester of the 2016-2017 academic year. The reason for choosing these schools

includes the appropriateness of the physical possibilities of the schools. Moreover, student achievements in these schools are lower than in other schools in the county.

Fluent reading scores were determined of all 75 students who regularly attended the 4th grade in the designated schools. The texts were selected from the Turkish lesson theme section on the "Morpa Campus" education site, which was also used during the study period when fluent reading scores were determined. Media scale and vocalizing scale scores were determined with the selected text "Icimizdeki Guzellikler", and the comprehension scores were determined with the selected text "Ah Su Gorgu Kurallari". These scores are also pretest scores of the study group. As a result of the scanning, the 75 students' fluent reading scores were ranked from the highest to the lowest, and the 30 students with the lowest scores were selected for the study. The student with code Y17 was excluded from the study, because the fluent reading score (110) of the student was too low and the reading was too slow; and the next lowest student was included in the study instead of this student. The names of these selected students were written down and put into a bag and 15 students were assigned to the experimental group and 15 students to the control group by drawing names from the bag.

Table 1

Distribution of Students Participating in the Survey According to Schools

Schools	Students being scanned		Participants	
	f	%	f	%
C. Primary School	25	33.3	10	33.3
G. Primary School	20	26.6	8	13.3
V. Primary School	13	17.3	2	6 .6
Y. Primary School	17	22.6	9	30
Total	75	100	30	100

Research Instrument and Procedure

The study used, Error Analysis Inventory, Reading Comprehension Scale and Reading Attitude Scale of the convergent mixed model to obtain quantitative data. The Semi-structured Student Interview Form was used to collect qualitative data. The names of the schools in which the survey conducted were coded with letters. For the four schools; - School-C, School-G, School-Y, and School-V- codes were used. All the schools have full-time schooling, and the course starts at 08.30 and ends at 14.30. There is one hour break for lunch from 12.00 and 13.00. The school library in School-C on Monday, the teacher's room in School-G on Tuesday, a vacant class not used in School-Y on Wednesday and the school library in School-V on Thursday were used with the study group for 14 weeks. The work was carried out at noon or during the appropriate class hours. The research was carried out by the researcher in the fall semester of the 2016-2017 academic year. Table 2 presents the number of experimental and control group students according to the schools.

Table 2

Distribution of Students by Schools

Groups	Schools			
	School-C	School-G	School -Y	School-V
Experimental Group	4	4	5	2
Control Group	6	4	5	-
Total	10	8	10	2

Table 2 shows the distribution of experimental and control group students across the schools. It is seen that School-C and School-Y had the highest number of participating students, whereas school-V had the minimum number of participants in the study. While there are no control group students in School-V, this school has at least experimental group students.

Before practice, the class teachers were given detailed information about the study and teachers were reminded that any reading activity should not be done with the students. The students of the experimental and control group were kept on their education within the curriculum framework. Digital texts were selected from three types of Turkish lesson themes, including storytelling, informative and poetic texts from Morpa Campus education site. Experimental group students read these digital texts. A tablet with a 9.7 inch screen was selected. In this way, students could have the opportunity to read on a wide screen. The control group students read the same texts in print. Experimental and control group pupils could access the texts used in the study with Morpa Campus membership; however, this membership has a subscription fee, therefore none of the students have a membership. Teachers can use this site free of charge with teacher login. Teachers were asked to not use the texts in the classroom during the course and the teachers complied with request.

In the application process, the experimental and control group students read 14 texts, one text every week. The title of these reading texts were "Ataturk'un Insan Sevgisi", "Cay Agaci", "Cocuklar Buyudugu Zaman", "Evliya Celebi", "Fiyonk Makarna", "Gelin Geldi Golu", "Hasta", "Ilk Hava Sehidimiz", "Karli Yollarda", "Kuslar ve Yemek Arkadaslari", "Kutuphanedeki Kedicik", "Mustafa'nin Dogusu", "Mutlu Adam" and "Okula Donus".

The texts used for the measurements in the evaluation process are as follows. Texts read for the pre-test were "Icimizdeki Guzellikler" for vocalizing and, "Ah Su Gorgu Kuralları" for comprehension questions; texts read for the mid-test were "Atatürk ve Koylu" for vocalizing and "Durak Usta" for comprehension questions; texts read for the post-test were "Bebek Takimi" for vocalizing and "Mustafa Kemal'in Anadolu Yolculugu" for comprehension questions. The audio recorder was used and the students were recorded while they read the texts. The students were asked 3 non-text questions and 3 in-text questions after they had read the comprehension texts. In order to collect the data in the study, the readings of the students were recorded by voice recorder. Voice recordings of the participants were listened to by the investigator and processed into the wrong analysis inventory.

In the study, the "False Analysis Inventory" adapted from Ekwall and Shanker (1988) by Akyol (2008: 237) was used to determine the level of comprehension and reading fluency. Scale consists of environment, vocalizing and question scale. The scales are scored within themselves, and the scores of the three scales are summed up as fluent reading scores.

In the determination process of reading attitudes, the "Adolescent Reading Attitude Scale", which was developed by McKenna, Conradi, Lawrence, Jang and Meyer (2012) and adapted into Turkish by Bastug and Keskin (2013) was used. The reading attitude scale consists of 7 items and with a 6-point Likert type measuring instrument. The answers range from 'very good' to 'very poor' and are scored by choosing one of 6 values. Number of errors made during reading were indicated with Error Ratio = Error Number Mean / (Number of words × Average of Reading Time) formula.

Data Analysis

The Mann Whitney U-Test was used to analyze the pre-test and post-test scores of the students in the experimental and control groups. The Wilcoxon Signs Rank test was used when the experimental group and the control group students were evaluated within themselves. The significance level in the comparison was taken as 05. Descriptive analysis was used while analyzing the qualitative data obtained as a result of the research. The data obtained according to this approach are summarized and interpreted below.

Results

Quantitative data related to pretest, mid-test, and post-test scores of the experimental and control group students and qualitative data and interpretations are given below.

Table 3

Comparison of Reading Comprehension, Reading Fluency and Reading Attitude Pre-test Scores of the Students in the Experimental and Control Group

	Groups	n	Mean Rank	Total Rank	U	*p*
Reading Comprehension	Experimental	15	16.80	252.00	93.00	.406
	Control	15	14.20	213.00		
Reading Fluency	Experimental	15	17.60	264.00	81.00	.190
	Control	15	13.40	201.00		
Reading Attitude	Experimental	15	15.67	235.00	110.00	.917
	Control	15	15.33	230.00		

p>.05

The results of the Mann Whitney U-Test on reading comprehension, fluent reading and reading attitude pre-test scores of the participants according to the experimental and

control groups are given in Table 3. The analysis results show that reading comprehension, fluent reading, and reading attitudes pre-test scores do not differ significantly between the experimental and control groups. According to the results, reading comprehension is (U = 93.00, p>.05), fluent reading is (U = 81.00, p>.05), and reading attitude is (U = 93.00, p>.05). This finding shows that experimental and control groups' reading comprehension, fluent reading, and reading attitude pre-test scores were equal.

Table 4

Comparison of Reading Comprehension and Reading Fluency Mid-test Scores of the Students in the Experimental and Control Groups

	Groups	n	Mean Rank	Total Rank	U	p
Reading Comprehension	Experimental	15	19.17	287.50	57.500	.019
	Control	15	11.83	177.50		
Reading Fluency	Experimental	15	18.23	273.50	71.500	.088
	Control	15	12.77	191.50		

p>.05

The results of the Mann Whitney U-Test on reading comprehension, fluent reading, and reading attitude mid-test scores of the participants according to the experimental and control groups are given in Table 4. The results of the analysis show a significant difference in reading comprehension mid-test scores (U = 57.500, p <.05) while the reading fluency mid-test scores (U = 71.500, p>.05) did not differ significantly between the experimental and control groups. According to these results, it can be said that reading the digital texts has no significant effect on the fluent reading of experimental group students, while it positively affects reading comprehension.

Table 5

Comparison of Reading Comprehension, Reading Fluency and Reading Attitude Post-test Scores of the Students in the Experimental and Control Groups

	Groups	n	Mean Rank	Total Rank	U	p
Reading Comprehension	Experimental	15	18.03	270.50	74.500	.107
	Control	15	12.97	194.50		
Reading Fluency	Experimental	15	18.70	280.50	64.500	.046
	Control	15	12.30	184.50		
Reading Attitude	Experimental	15	16.30	244.50	100.500	.618
	Control	15	14.70	220.50		

p>.05

The results of the Mann Whitney U-Test on reading comprehension, fluent reading, and reading attitude post-test scores of the participants according to the experimental and control group are given in Table 5. Analysis results indicate that the reading fluency scores differ significantly while the reading comprehension and reading attitude scores of the post-test do not change significantly between the experimental and control groups. According to the results, reading comprehension is

(U = 74.500, *p*>.05), reading fluency is (U = 64.500, *p*<.05), and reading attitude is (U = 100.500, *p*>.05). This finding shows that digital reading has a significant effect on reading fluency, while it has no significant effect on the reading comprehension or reading attitude of the experimental group. In other words, it can be said that reading digital texts has an impact on reading, but it does not influence reading comprehension or reading attitude.

Table 6

Experimental Group Students' Reading Comprehension Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	*p*
Pre-Test Mid-Test	Negative Rank	2	3.50	7.00	-2.103	.035
	Positive Rank	8	6.00	48.00		
	Equal	5				
Mid-Test Post-Test	Negative Rank	3	5.33	16.00	-1.181	.238
	Positive Rank	7	5.57	39.00		
	Equal	5				
Pre-Test Post-Test	Negative Rank	0	.00	.00	-3.076	.002
	Positive Rank	12	6.50	78.00		
	Equal	3				

p>.05

Table 6 shows the results of the Wilcoxon sign test, which indicates whether there is a significant difference between pre-test, mid-test and post-test scores of the experimental group students' reading comprehension. The results of the analysis show that a significant difference between the pre-test and mid-test scores of the students in the experimental group (z = -2.103, *p*<.05); there is also a significant difference between pre-test and post-test (z = -3.076, *p*<.05). There is no significant difference between mid-test and post-test (z = -1.181, *p*>.05). These findings show that the digital text reading activity applied to the students improves their reading comprehension. It can be said that the longer the activity lasts, the lower the overall effect of reading digital text on the students' reading comprehension.

Table 7

Experimental Group Students' Reading Comprehension Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	*p*
Pretest-Midtest	Negative Rank	2	6.00	12.00	-1.587	.112
	Positive Rank	8	5.38	43.00		
	Equal	5				
Midtest Posttest	Negative Rank	1	10.00	10.00	-2.284	.022
	Positive Rank	11	6.18	68.00		
	Equal	3				
Pretest-Posttest	Negative Rank	2	4.75	9.50	-2.875	.004
	Positive Rank	13	8.50	110.50		
	Equal	0				

p>.05

Table 7 indicates the results of the Wilcoxon sign test, which reveal whether pre- test, mid-test, and post-test scores of the control group students differed significantly in their comprehension. The results of the analysis show no significant difference ($z = -1.587$, $p>.05$) between the pre-test and the mid-test scores of the control group students participating in the study and, but a significant difference between the mid-test and post-test ($z=-2.284$, $p<.05$), and a significant difference between pre-test and post-test ($z=-2.875$, $p>.05$). These findings show that the activity of reading printed text, which was applied to the control group, has a significant effect on reading comprehension skills when the process is extended, while this activity has not positive effect in the short term.

Table 8

Experimental Group Students' Reading Fluency Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	p
Pretest-Midtest	Negative Rank	2	5.50	11.00	-2.784	.005
	Positive Rank	13	8.38	109.00		
	Equal	0				
Midtest Posttest	Negative Rank	3	5.33	16.00	-2.503	.012
	Positive Rank	12	8.67	104.00		
	Equal	0				
Pretest-Posttest	Negative Rank	0	.00	.00	-3.415	.001
	Positive Rank	15	8.00	120.00		
	Equal	0				

p>.05

The results of the Wilcoxon Signs test in Table 8 show whether there is a significant difference between pre-test, mid-test, and post-test reading scores of the experimental group students. The results of the analysis show a significant difference between pre-test and mid-test scores ($z = -2.784$, $p <.05$), mid-test and post-test scores ($z = -2.503$, $p<.05$), and pretest and post test scores ($z = -3.415$, $p < 05$) of experimental group students. If these findings are taken into consideration, it can be said that the activity of reading digital text, which was applied to the experimental group, has an important effect on improving students' reading fluency skills.

Table 9

Control Group Students' Reading Fluency Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	p
Pretest-Midtest	Negative Rank	3	7.17	21.50	-2.189	.029
	Positive Rank	12	8.21	98.50		
	Equal	0				
Midtest Posttest	Negative Rank	3	5.17	15.50	-2.530	.011
	Positive Rank	12	8.71	104.50		
	Equal	0				
Pretest-Posttest	Negative Rank	1	1.00	1.00	-3.352	.001
	Positive Rank	14	8.50	119.00		
	Equal	0				

p>.05

The results of the Wilcoxon Signs test shown in Table 9 indicate whether there is a significant difference between pre-test, mid-test and post-test reading scores of the control group students. The results of the analysis reveal a significant difference between pre-test and mid-test scores (z = -2.189, *p* <.05), mid-test and post-test scores (z=-2.530, *p*<.05), and pre-test and post-test scores (z=-3.352, *p*<.05) of control group students. When these findings are taken into consideration, it can be said that the activity of reading printed text, which was applied to the control group, has an improving effect on the reading fluency skills of the students.

Table 10

Experimental Group Students' Reading Attitude Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	*p*
Pretest-Posttest	Negative Rank	9	5.39	48.50	-.655	.513
	Positive Rank	6	11.92	71.50		
	Equal	0				

p>.05

The results of the Wilcoxon Signs test shown in Table 10 reveal whether there is a significant difference between experimental group students' reading attitude scores obtained before and after application. The results of the analysis show no significant difference between the pre-test and post-test scores (z = -655, *p*>.05) of the experimental group students. It can be thus observed that the digital text reading activity did not change students' reading attitudes in these findings.

Table 11

Control Group Students' Reading Attitude Scores Wilcoxon Signed Ranks Test Results

		n	Mean Rank	Sum of Rank	z	*p*
Pretest-Posttest	Negative Rank	7	5.50	38.50	-.039	.969
	Positive Rank	5	7.90	39.50		
	Equal	3				

p>.05

Wilcoxon signed rank test results are given in Table 11, indicating whether there is a significant difference between pre-test and post-test reading attitude scores of control group students. The results of the analysis show no significant difference between the pre-test and post-test scores (z = -039, *p*>.05) of the control group students. It can be understood that the printed text-reading activity does not change the reading attitudes of the students, according to these findings.

Table 12

Average Results of Experimental and Control Group Students' Reading Errors

	Experimental Group			Control Group		
	Pre-test	Mid-test	Post-test	Pre-test	Mid-test	Post-test
Number of Word	196	238	216	196	238	216
Average Reading Time	3.0	2.98	2.50	3.89	3.67	3.23
Average Number of Error	20.2	14.2	10.9	25.3	25.4	18.0
Error Rate	0.0343	0.0200	0.0201	0.0331	0.0290	0.0258

When Table 12 is examined, it can be seen that error rates are determined by taking the average of reading errors in the texts, which are found in the pre-test, mid-test, and post-test measurements of the experimental and control group students. It can be seen that the pre-test error averages of the experimental group are close to the averages of the control group. In the implementation process, it can be seen that the error rates of the experimental group decreased significantly as compared to the control group. In other words, reading digital text is more effective than reading printed text in reducing the number of words read incorrectly.

Opinions of the Experimental Group Students on Using Digital Texts in the Reading Process

The opinions of the experimental group students were taken in accordance with the fifth sub problem, expressed as "What are the opinions of students in the experimental group on using digital texts in the reading process?". Thus, quantitative data collected during the implementation process can be supported by qualitative data. In this context, the following questions were answered in interviews with the experimental group students.

- Facilitating reading and comprehension,

- Liking and resuming reading,

- Facilitating and enforcing the reading process.

Questions were asked to the students (n = 15) in the experimental group, in order to get their views on the application. When asked "How do you like the idea of reading digital text on a tablet?" before the activity, the students (n=12) said that they would be excited, pleased, well and happy. Student C16, G10, and C23 expressed their thoughts by saying, "I like it a little, I did not like it once", "Tablet is not good, book is better", and "I like to read the book."

When the question "How was it to read digital texts on the tablet?" was asked to the students (n = 12) after the application, they said it was good, they liked it, it was fun and it was nice to read. C16 said that "I had read on a small screen tablet before and did not like it very much but now I like to read and read comfortably because the screen was large". G10 expressed that his idea had not changed and that reading printed text was better than tablet, and C23 said that he liked it and his negative idea had changed.

When the students were asked, "Let's say you will read a story; you have a tablet and a story book on your table, and both of them have the story you will read. Which one do you prefer to read? Why?", students (n = 10) said they would choose the tablet because it was fun and the pictures were more colorful, students (n = 2) expressed that they would choose both the paper and the tablet because both of them were nice, and students (n = 3) said that they would prefer paper because they would read more easily.

When asked that "Is there a physical fatigue such as eye pain, headache, back pain during reading?", student (n=12) expressed that, in general, the high color made it easier for them to read as opposed to giving them eye pain. G10 said that electronic devices hurt his eyes and he had headache when he (n=12) looked at the screen too much; V4 said that the screen strained his eyes a little, G7 expressed that he had a little eye pain.

When asked "Does reading digital texts through the tablet contribute to your comprehension of the text?", students (n = 10) said they remembered more easily when they read with the tablet, while students (n = 5) explained that there was no difference or if they read on paper that, they would understand better.

When asked "Is there any difficulty in reading with the tablet?, If so, what is it?", students (n=13) explained that, in general, it was easy and there was no difficulty. Y8 and C19 students said that they could not complete the reading process if the tablet battery was drained.

The experimental group students (n = 15) generally explained that they liked to read digital texts on tablets, that they found it entertaining and that it contributed to their comprehension of the texts. The period of application was limited to 14 weeks. During this time, no fuss was observed in the students. On the contrary, they said they waited enthusiastically for their weekly reading activity. It was observed that the reason of students' desire to not finish the reading hours was reading a different text on a different material, and this situation increased their desire to read.

Discussion and Conclusion

In this study, a significant difference was seen between the group reading digital text and the group reading printed text in 4th grade students. When the quantitative dimension of the research results is examined, it can be seen that the reading activities with digital texts developed fluent reading skills; the students' reading comprehension was positively influenced by the intermittent measurement results but the effect of the

application decreased with the prolonged application time. The activity had no significant effect on reading attitude, and the number of reading mistakes made by students decreased. As for the data related to the qualitative dimension of the study, the students generally stated that reading digital text on tablet was a positive experience, and that they read fondly.

These results are also supported by related studies. Although only few studies investigate the effect of digital texts on fluent reading skills, Akbar et al., (2015) conducted a study in which they investigated the effect of reading digital text on reading fluency. They found that effect of reading digital text on reading comprehension and reading attitudes was negative while it increased students' reading fluency. Schneps, Thomson, Sonnert, and Pomplun (2013) compared reading on printed material with reading on electronic devices in terms of levels of reading fluency and reading comprehension of students who had reading disabilities. According to the research results, reading on the device significantly influenced reading fluency and comprehension, although not for all of the group just a large part. In his study, Reichenberg (2014) examined the influence of second-grade students' reading on a tablet computer on reading comprehension. At the end of the study, no significant difference was found between the comprehension scores of the students in the experimental group and the control group. However, the test scores of the experimental group were higher than those of the control group.

Although our study has shown that reading on a tablet does not have a significant effect on reading attitude, studies are available in which this influence was indicated. Williams (2010) investigated the influence of tablet computer use on reading comprehension and attitudes towards the lesson. The students in the experimental group were applied a reading program using a tablet computer, and at the end of the study, a decrease was found in the students' reading comprehension tests scores and an increase was found in the attitudes towards the lessons.

Students' reading digital texts on tablets is a situation that students willingly encounter in their daily life or in our education system. Young age groups in particular seem to enjoy reading. It is also possible to see that tablet computers and digital texts have different effects on students. In his work with second grade students, Larson (2010) found that using digital reading devices developed new literacy practices and increased interaction between readers and reading. In addition, electronic devices and their features make it necessary to move these reading texts by hand. Larson has provided a distinctive contribution to the field by stating that this situation has increase the interaction between the reader and the reading material.

This study differs from other research in that it reveals the influence of digital texts on reading fluency skills. The lack of such studies in the related literature makes this study important. Even though technological products are used in the education system in our country, it can still be not used at the desired level in educational environments. One of the most important reasons for this situation is the lack of software programs suitable for our education system. It is possible to enrich the educational environment

by using facilities provided by the technology in a planned way, where these problems can be solved.

References

Akbar, R. S., Taqi, H. A., Dashti, A. A., & Sadeq, T. M. (2015). Does e-reading enhance reading fluency? *English Language Teaching*, *8*(5), 195-207.

Akyol, H. (2008). *Turkce ogretim yontemleri* [Turkish teaching methods]. Ankara: Kok Yayincilik.

Bastug, M. & Keskin, H. K. (2013). Ergenlik donemi okuma tutum olceginin turkceye uyarlanmasi [Adaptation of adolescence period reading attitude scale to turkish]. *Turkısh Studies, 8*(4), 295-311.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* London: Sage Publications.

Gungoren, O. C., Bektas, M., Ozturk, E., Horzum, M. B. (2014). Tablet bilgisayar kabul olcegi- gecerlik ve guvenirlik calismasi [Tablet computer acceptance scale-validity and reliability study]. *Egitim ve Bilim*, *39*(176), 69-79.

Johnson, L., Adams, S., & Cummins, M. (2012). Join the conversation. join the community. Join the NMC. (Horizon Report). Retrieved from https://www.nmc.org/pdf/2012-horizon-report-K12.pdf

Jonassen, D. H. (1999). *Computers as mindtools for schools: Engaging critical thinking*. Englewood Cliffs, NJ: Prentice Hall.

Komis, V., Ergazakia, M., & Zogzaa, V. (2007). Comparing computer-supported dynamic modeling and 'paper & pencil' concept mapping technique in students' collaborative activity. *Computers & Education, 49*(4), 991-1017.

Larson, L. C. (2010). Digital readers: the next chapter in e-book reading and response. *The Reading Teacher*, *64*(1), 15-22.

Leu, D. J., O'Byrne, W. I., Zawilinski, L., McVerry, J. G., & Everett-Cacopardo, H. (2009). Expanding the new literacies conversation. *Educational Researcher*, *38*(4), 264-269.

Luke, A. (2012). Talking and Reading and writing the world, journal of adolescent & Adult Literacy. *International Reading Association*, *56*(1), 8-13.

McCannon, M., & Crews, T. B. (2000). Assessing the technology needs of elementary school teachers. *Journal of Technology and Teacher Education*, *8*(2), 111-121.

Quintana, B. B., & Donovan, L. (2013). Storytime using ipods: using technology to reach all learners. *TechTrends*, *57*(6), 49-56.

Reichenberg, J. L. (2014). *Theuse of iPads to facilitate growth in reading comprehension skills of second grade students.* Doctorial Dissertation, Liberty University. Lynchburg, VA.

Sackstein, S., Spark, L., & Jenkins, A. (2015). Are e-books effective tools for learners? Reading speed and comprehension iPad vs. paper. *South African Journal of Education*, *35*(4), 1-14.

Schneps, M. H., Thomson, J. M., Sonnert, G., Pomplun, M., Chen, C. & Heffner Wong, A. (2013). Shorter lines facilitate reading in those who struggle. *PloSome*, *8*(8), 1-16.

Thoermer, A., & Williams, L. (2012). Using digital texts to promote fluent reading. *The Reandin Teacher*, *66*(7), 441-445.

Tolani, B. N., McCormac, M., & Zimmermann, R. (2009). An analysis of the research and impact of ICT in education in developing country contexts. *Journal of Education for International Development*, *4*(2), 1-12.

VanAken, A. M. (2014). *Effect of EBooks on Reading Level, Reading Behaviours and Attitude of Second Grade Students* (Doctoral Dissertation). Liberty Universiy, Lynchburg, VA.

Walsh, M., & Simpson, A. (2013). Touching, tapping … thinking? Examining the dynamic materiality of touch pad devices for literacy learning. *Australian Journal of Language and Literacy, 36*(3), 148-157.

Wagner, D. A., Castillo, N. M., Murphy, K. M., Crofton, M., & Zahra, F. T. (2014). Mobiles for literacy in developing countries: An effectiveness framework. *Prospects*, *44*(1), 119-132.

Williams, M. A. (2010). *Examining the efficacy of using iPod touches to deliver reading comprehension strategy instruction and to provide electronic text support on the reading comprehension performance of sixth-grade students* (Doctorial Dissertation). University of Oregon.

Yildirim, A., & Simsek, H. (2006). *Sosyalbilimlerdenitelarastirmayontemleri* [Qualitative research methods in the social sciences]. Ankara: Seckin Yayincilik.

Dijital Metinlerin İlkokul Öğrencilerinin Okuduğunu Anlamasına, Akıcı Okumalarına ve Tutumuna Etkisi

Özet

Problem Durumu: Eğitim öğretim denilince ilk akla gelen kavramlardan birisi ve belki de en önemlisi okumadır. Okuma becerisi ile ilgili çocuklar çok erken yaşlarda tanışmakta ve okuma işi ilgilerini çekmektedir. Günümüzde çocuklar okul çağına gelmeden okuma ve yazma becerilerini kazanabilmektedirler. Bunun en önemli sebebi ailelerin ilgisi ve çevre uyarıcıların fazlalığıdır. Okuma becerisini kazanan çocuklar ileriki yıllarda da bu beceriyi kullanmakta ve sürekli geliştirme durumunda kalmaktadır. Günümüzde bilgiye ulaşmanın çok farklı yöntemleri bulunmaktadır. Bu yöntemlerin başında ise teknoloji kullanımı gelmektedir. Kendini sürekli yenileyen teknoloji, çağın gereklerine hizmet etmek için en kullanışlı yol haline gelmiştir. Teknolojideki bu durum eğitim öğretim veren okullarımıza da yansımış ve her kademede kullanılmaya başlanmıştır. Öğrenciler ve öğretmenler gerek okul içerisinde gerekse okul dışında öğretme ve öğrenme etkinliklerini daha hızlı ve kalıcı hale getirmek için teknolojiyi kullanmaktadırlar. Bu bağlamda gelişmekte olan yeni okuryazarlığı faydalı ve etkili kullanmak için bazı becerilerin toplumun bireylerinde olması gerekmektedir. Kuşkusuz burada ki en büyük pay teknolojinin olacaktır. Günümüzde bilgisayarın bile artık demode olduğu, yerini dokunmatik tablet bilgisayarların ve akıllı cep telefonların aldığı bir süreçten geçilmektedir. İnternet kullanımının okul öncesi dönemlerde başladığı düşünüldüğünde teknolojinin getirdiği yenilikler hayatımıza çok erken yaşlarda girmeye başlamaktadır. Ülkemizde bu hızlı değişim karşısında farklı projeler geliştirmekte ve uygulamaktadır. Bu projeler arasında en kapsamlı olan FATİH (Fırsatlar Artırma ve Teknolojiyi İyileştirme Hareketi) projesi dikkat çekmektedir. Bu proje ile ilkokul, ortaokul ve lise öğrencilerinin tamamına kademeli olarak tablet dağıtımı gerçekleştirilecektir. Sınıflardan akıllı tahta ile birlikte de çalışabilecek olan bu tabletler, derslere yönelik dijital metinler ve ek çalışma kaynaklarının da sunulması ve etkileşimli bir sınıf ortamında eğitim-öğretim hizmetlerinin sürdürülmesi amaçlanmaktadır.

Araştırmanın Amacı: Bu araştırmanın temel amacı, dijital metinlerin ilkokul 4. sınıf öğrencilerinin okuduğunu anlama, akıcı okuma ve okuma tutumu üzerindeki etkisini ortaya koymaktır. Bu amaç doğrultusunda aşağıdaki sorulara cevap aranacaktır.

1. Dijital metinlerin kullanıldığı deney grubu öğrencileri ile basılı metinlerin kullanıldığı kontrol grubu öğrencileri arasında okuduğunu anlama puanları arasında anlamlı bir farklılık var mıdır?

2. Dijital metinlerin kullanıldığı deney grubu öğrencileri ile basılı metinlerin kullanıldığı kontrol grubu öğrencileri arasında akıcı okuma becerileri puanları arasında anlamlı bir farklılık var mıdır?

3. Dijital metinlerin kullanıldığı deney grubu öğrencileri ile basılı metinlerin kullanıldığı kontrol grubu öğrencileri arasında okuma tutum ölçeğinden aldıkları puanlar arasında anlamlı bir farklılık var mıdır?

4. Dijital metinlerin kullanıldığı deney grubu öğrencileri ile basılı metinlerin kullanıldığı kontrol grubu öğrencileri arasında okuma hatalarını düzeltmede anlamlı bir farklılık var mıdır?

Araştırmanın nicel boyutuyla ilgili araştırma sorularına ilişkin ulaşılan sonuçları desteklemek, ulaşılan sonuçların uygulanan yöntemden mi yoksa farklı değişkenlerin de bu süreçte etkili olup olmadığını ortaya koymak, araştırma sürecine ilişkin daha zengin ve derinlemesine bilgilere ulaşmak amacı ile araştırmada ayrıca aşağıdaki sorulara cevap aranacaktır. Açıklayıcı sıralı karma yöntemin nitel boyutuna ilişkin sorular aşağıdaki gibi ifade edilmiştir:

5. Deney grubundaki öğrencilerin okuma sürecinde dijital metinleri kullanmaya ilişkin görüşleri nelerdir?
- Okuma ve anlamayı kolaylaştırma durumu,
- Okumayı sevme, devam ettirme durumu,
- Okuma sürecini kolaylaştırma/zorlaştırma durumu.

Araştırmanın Yöntemi: Araştırmada nicel ve nitel araştırma yöntemlerinin bir arada kullanıldığı yakınsayan (eş zamanlı) paralel karma yöntem kullanılmıştır. Karma yöntemin seçilme nedeni, elde edilen nicel verilerin nitel veriler ile desteklenerek sonuçların daha anlaşılır olmasını sağlamaktır. Çalışma grubu, 2016-2017 eğitim-öğretim yılı güz döneminde Kırşehir ili Kaman ilçe merkezinde bulunan sekiz ilkokul arasından sosyo-ekonomik düzeyi denk olan dört ilkokulun 4. sınıfında öğrenim gören toplam 75 öğrenci üzerinde gerçekleştirilmiştir. Yapılan tarama sonucunda 75 öğrencinin akıcı okuma puanları yüksek puandan başlanarak sıralanmış ve puanı en düşük 30 öğrenci çalışma için seçilmiştir. Seçilen bu öğrencilerin isimleri yazılarak bir torbaya atılmış ve kura çekilerek 15 öğrenci deney grubuna, 15 öğrencide kontrol grubuna atanmıştır. Deney ve kontrol grubu öğrencilerinin ön test- ara test ve son test puanların analizinde Mann Whitney U-Testi kullanılmıştır. Deney grubu ve kontrol grubu öğrencilerini kendi içerisinde değerlendirirken Wilcoxon işaretler sıralar testi kullanılmıştır. Karşılaştırmalarda anlamlılık düzeyi .05 olarak alınmıştır.

Araştırmanın Bulguları: Araştırma sonucunda dijital metin okumanın öğrencilerin akıcı okuma becerilerinde, okuduğunu anlama becerilerinde ve okuma sırasında yaptıkları okuma hatalarında anlamlı düzeyde etkisinin olduğu bulunmuştur ($p < .05$). Ayrıca okuma tutumu üzerinde anlamlı bir etkinin olmadığı görülmüştür.

Araştırmanın Sonuçları ve Önerileri: Araştırma sonuçlarının nicel boyutuna ilişkin veriler incelendiğinde, dijital metinler ile yapılan okuma çalışmalarının akıcı okuma becerilerini geliştirdiği, öğrencilerin okuduğunu anlama becerilerini de aralıklı ölçüm sonuçlarına göre olumlu yönde etkilediği fakat uygulama süresinin uzamasıyla

etkisinin azaldığı, okuma tutumu üzerinde anlamlı düzeyde bir etkinin olmadığı ve öğrencilerin okuma sırasında yaptıkları okuma hatalarının sayısında da azalma olduğu bulunmuştur. Araştırmanın nitel boyutuna ilişkin veriler incelendiğinde öğrenciler genel olarak tablet ile dijital metin okumanın olumlu olduğunu, severek okuma yaptıklarını belirtmişlerdir. Araştırmanın diğer araştırmalardan farkı dijital metinlerin akıcı okuma becerileri üzerindeki etkisini ortaya koymasıdır. İlgili literatürde bu tür çalışmaların yok denecek kadar az oluşu bu çalışmayı sonuçları bakımından önemli kılmaktadır. Ülkemizde eğitim sistemimizde teknolojik ürünler kullanılmasına rağmen eğitim ortamlarında istenilen düzeyde istifade edilememektedir. Bu durumun en önemli sebeplerinden bir tanesi de eğitim sistemimize uygun yazılımsal programların eksikliğidir. Bu sorunun ortadan kalkması ile teknolojinin getirdiği imkânlar planlı bir şekilde kullanılarak eğitim ortamlarını zenginleştirmek mümkündür. Öğretmenlere öğrencilerin akıcı okumalarını geliştirmede ve okuduğunu anlamada dijital metinleri kullanmalarını öneririz.

Anahtar Kelimeler: Okuma becerileri, teknoloji, tekrarlı okuma.

Curriculum Development Study for Teacher Education Supporting Critical Thinking*

Sahika UNLU[1]

A R T I C L E I N F O	A B S T R A C T

Purpose: The existence of teachers who think critically and reflect this in their profession is an attribute that should not be abandoned in the process of generating a free, critical thinking, and democratic society. Teachers are the most important role models students have in shaping their behavior in all environments. Therefore, it is necessary for teachers to possess the skills of the new era, and the continuity and support of their professional development is fundamental. One of the foremost objectives of the education programs in Turkey is providing students with critical thinking ability. However, a teacher education program intended to support critical thinking in teachers, the executors of the program, is not available. The purpose of this study is to develop an educational program intended for primary education teachers that supports critical thinking and makes reflective evaluations. **Research Methods:** This research used the "Taba-Tyler Model," a curriculum development model. This research method was selected because a new education program is designed in this study. Findings: In the "Draft Supporting Critical Thinking Curriculum" that was formed as a result of the research, 12 objectives are presented under five themes (open-mindedness, questioning of the accuracy and reliability of information, reason-evidence seeking, openness, asking high-level questions). The application of the draft requires 15 hours. As a result of the reflective evaluation, the objectives and content of the draft program were found to be relevant regarding the learning experiences and the purposes of assessment and evaluation were consistent within themselves. **Implications for Research and Practice:** It is proposed that "Supporting Critical Thinking Curriculum" be developed for teachers working at different levels of education, because critical thinking education is a process that involves all levels of education.

Introduction

Thinking is the foremost trait of humankind that distinguishes people from other living creatures. However, this human trait is not sufficient in assuring that individuals will have good reasoning skills. In the ideal way of thinking, our minds use thought processes, such as information acquiring, understanding, comprehending, analyzing, correlating, interpreting, evaluating, making judgments, and making decisions when reaching a conclusion about what is "good or bad, right or wrong". Critical thinking is a form of high-order thinking that involves all of these thinking processes.

The emergence of critical thinking as a concept dates back to Socrates in ancient Greece in the 6th century BC. Socrates defined critical thinking as "evaluating something in terms of its good or bad aspects," and defined critical thinking as the *art of judgment*. Today, "Socratic inquiry method" remains to be the most widely known and used method in teaching critical thinking.

In the face of the intense flow of information in the 21st century and our exposure to it in every area, we must rely on using all the processes of critical thinking. This includes a willingness to acquire information, inquiring into the accuracy and reliability of the information, reading the data and indicators correctly based on evidence, and posing the correct questions. This is imperative in decision making. (Paul, Elder & Bartell 1997, p.1).

This inquiry process, on which critical thinking is based, has led to the intense interest of experts from different disciplines, especially psychology, philosophy and education, as well as an interest in creating it as a field of study. However, this interest has led to a serious confusion in defining the term itself. As a result, various definitions have been used by field experts.

In order to be able to make a comprehensive definition of critical thinking, the "deep/reflective thinking" definition of Dewey is important (Gundogdu, 2009). Dewey defines critical thinking as the highest level of awareness to which a person is conscious, not only through human senses but also through the mind. It distinguishes in-depth thinking from "random thinking, fictitious thinking and imitative thinking" (Dewey, 1910, p. 4). The distinguishing features of Dewey's definition can be summarized as "a careful evaluation as an active thinking process, a thinking that is aware of its evidences and consequences, the process of thinking itself, the open thinking towards change and self-betterment" (Fisher, 2001; Lipman, 2003; as cited by Gundogdu, 2009), that is to say "controlled thinking." Other examples of definitions that different researchers suggest for critical thinking are as follows:

- Reaching results based on observation and information (Paul 1991, p. 125)

- Logical reflective thinking that is focused on helping in making judgments or making decisions (Paul & Elder, 2008)

- Problem solving (Ritchhart & Perkins, 2005)

- Effective interpretation and evaluation of observation, communication, information, and discussions (Scriven, 2007)

- Adopting critical thinking attitudes and tendencies (Glaser, 1983)

- Analyzing the situation with a careful, objective, and lasting approach (Beyer, 1987)

- Providing a consistent internal motivation (Facione & Facione, 1998)

- Having the tendency and skill to act with a reflective skepticism (McPeck, 1981)

- Following a versatile process that involves many mental activities (Lipman 2003; Ritchhart and Perkin, 2005; Fisher, 2001 and Doganay, 2001)

- Developing a skill that is not inherent, that is acquired afterwards and develops in proportion to the level of life (Cikrikci, 1996)

In his definition intended for this concept, Halpern (2003) outlines the limits of critical thinking as follows: (as cited by Gundogdu, 2009, p. 58):

A thought, by functioning as a mental hygiene that prevents manipulation on human thought, that helps everyone in noticing the propaganda, not being cheated, analyzing the implicit and unvoiced assumptions underlying the claims and arguments asserted, immediately noticing a deception, evaluating the reliability of a source of information, approaching a problem or a decision as best as possible.

Examining the generally accepted definition used by Ennis (1985), critical thinking is reflective and rational thinking that is focused on deciding what to believe or what to do.

Given the fact that experts, after long discussions in the field, have agreed that the definition of critical thinking is composed of skills, attitudes and tendencies, lends itself to the emerging view that critical thinking can be taught. According to De Bono (2007), thinking is a skill that can be learned and developed from education because it is the basic competence of humankind. However, this process, which is expressed as abstract and superficial, requires intense work and determination and takes place in a very complex process (Sensekerci & Bilgin, 2008, p. 27). In addition, experimental studies in the literature reveal that critical thinking education at all levels of the learning process has a positive impact on students. So who will lead this process?

In theoretical literature, many researchers indicate that the foremost factor in teaching critical thinking is the teacher (Ennis, 1985, 1993, Newmann, 1991, Paul 1991, Rovegno 1992, Marzano 1988, Fisher 2001, Halpern 2004; Doganay and Sari, 2012; Nosich, 2012), and the teacher even plays a more important role than the teaching of thinking skills. Research findings (Kaloc, 2005; Aybek, 2006; Cekic, 2007; Ozcan, 2007) also support this view. Since it has been emphasized that the most important role is

that of the teacher, the question, "What kind of teacher?" has presented a problem, and researchers have concentrated intensely to determine the answer. The teacher is not only a person who transfers the information, but also sets an example with his/her attitudes and behaviors. Field experts remark that only teachers who can think broadly can educate individuals in this ability and become role models (Walsh and Paul, 1998). Critical thinking should be taught by well-trained, knowledgeable, and experienced teachers in this field (Marzano, 1988; Demirci, 2002, Has, 2012). Researchers clarify the role of the teacher in instructing critical thinking within the context of a "guide" concept in the most effective way. However, the question of how teachers will guide students in learning to think in the classroom has existed in the literature for many years and it is as a question/problem that challenges teachers today (Doganay, 2012). It has been said that teachers who adopt critical thinking consider the high-order cognitive skills, such as analysis, synthesis and evaluation, necessary for their students (Sahinel, 2007). They contribute to the cognitive development of their students as a result of designing the teaching process accordingly, and they positively affect the attitude toward critical thinking (Seferoglu & Akbiyik, 2006).

According to Paul (1989), the foremost point that is overlooked is that the primary task of teachers who apply critical thinking teaching must "realize that they should teach students how to learn." Teachers should search for meanings, ask for reasons and evidence, facilitates cooperation, prevent controversies that could lead to disorder, encourage others to listen, guide students toward efficient comparisons and oppositions, emphasize contradictions and inconsistencies, and know to ask questions that enlighten the inferences and consequences (as cited by Koc, 2007). The findings obtained in the study of Erdamar and Alban (2017 p. 793), in which teachers' opinions were obtained as to what the traits of critical thinking teachers should be, are as follows:

Teacher must:

- Respect student opinions

- Be a model to the students in critical thinking

- Create a democratic and secure learning environment

- Monitor innovations in the field

- Listen to students

- Be open to criticism

- Be unbiased

- Address individual differences of students by using different methods techniques

While the role of the teacher is crucial in teaching critical thinking, Ashton (1988) states that the biggest obstacle in the objective of raising critical thinking individuals

in educational programs is that the teachers are not equipped with critical thinking knowledge and skills. He supports this view in his research on enhancing teacher qualifications and emphasizes the importance of focusing on prospective teachers and teacher education programs to emphasize critical perspectives to teachers (Paul, 1991; Tinning, 2002; Aybek, 2006, Boonjeam, Tesaputa & Sri-ampai, 2017).

According to other researchers, it is imperative to support and develop the professional knowledge and skills of teachers, in addition to their standard education, in order for them to be able to possess critical thinking skills, attitudes, tendencies, and behaviors, which enable them to transfer this knowledge into their lives (Critical Thinking Skills and Teacher Education, 1988; Peterson, Kromrey, Borg, and Lewis, 1990; Senemoglu, 1996; Kokdemir, 2003; Turnuklu & Yesildere, 2005; Battal, 2008; MoNE, 2008; Yesilpinar, 2011).

In their study, Peterson, Kromrey, Borg, and Lewis (1990) also demonstrated that teachers educated in high-order thinking skills, such as problem-solving and critical thinking skills, performed better in teaching other relevant skills.

As long as teachers do not support critical thinking through their behaviors in the classroom, it is difficult, as a whole, to educate critical thinking individuals and thus bring about a transformation in teaching programs (Alkın-Sahin & Gozutok, 2013). Therefore, building a teacher education system that supports the critical thinking process of students in the classroom environment of teacher education institutions and that brings the practice into the forefront is also of great importance in teacher education. According to Wilks (2005), in order for education systems to teach students who question well, participate more, are more open to debate, determine the predictions and priorities, seek alternatives, and infer from various perspectives, it is primarily necessary to educate the teachers in such a way that they can create these qualifications in their students. Moreover, in the findings of the study of Cave (1993) titled "Teachers' Behavior and Disposition on Students Critical Thinking," teachers with high-order critical thinking skills often provide their students with a rich learning environment by changing their learning activities through various teaching methods, thus ensuring that their students think critically.

The previous discussions demonstrate that the presence of teachers who think critically and reflect this in their behaviors is the starting point in educating critical thinking individuals.

When critical thinking in Turkish education systems is considered, and given that the measurement of such skills is an important indicator in international exams, a series of program changes become necessary. This should start with the primary education level and be included among the basic skills in educating individuals with 21st century qualifications required by the new age (MONE, 2005, 2017b). The role of the teacher is to bring to life relevant changes and innovations in education programs. However, definitions of "facilitator, guide, intermediary" designated to teachers who will bring change to education programs, remains a theory (Gurkaynak, Ustel and Gulgoz, 2008). In practice, critical thinking as a whole in Turkish literature (Kurum, 1999; Sensekerci and Bilgin, 2008; Tanriverdi, Doganay and Sari, 2012; Alkin-Sahin,

2012; Ulusoy & Turan, 2012; as cited by Unlu, 2017) is aimed at determining the critical thinking levels of teachers and prospective teachers. These critical thinking levels are inadequate, there are very few teachers with these qualifications, and the vast majority of those who do have the skills are insufficient. International exam results show that the objective of higher-order achievements is not realized (OECD, 2013, TEDMEM, 2014; MoNE, 2015; OECD, 2016: as cited by Unlu, 2017). The necessity to support the professional development of teachers, who are the executors of the program to improve this situation, has been emphasized in the reports of different institutions (OECD, 2003; ERG, 2005, 2008; METU, 2017: as cited by Unlu, 2017).

According to Perry (2014), addressing teacher education strategies in critical thinking and supporting professional development of teachers should be emphasized to develop and implement a comprehensive program for teachers. The importance of critical thinking education and the education of teachers has been emphasized in policy decisions (MoNE 2006, 2017a; THS, 1999, 2004, 2006) in Turkey, and is understood from evaluations, the innovation movement did not achieve its goals. It has deficiencies, and teacher competence is still being discussed among the most significant obstacles in the implementation of the program. In improving this situation, the lack of a teacher education program, at all levels, supports the need for further research.

Teachers are the foremost role models in the critical thinking development of their students. Thus, designing an education program that will support critical thinking of teachers is a significant step both in supporting their development as critical thinkers and in the creation of a critical thinking society. Therefore, the purpose of this study is to prepare a draft education program intended for primary education teachers that supports critical thinking and to makes a reflective evaluation. In parallel with this purpose, the study aims to find the answer to the question, "Are the objectives, content, learning activities, and evaluation dimensions of the Supporting Critical Thinking Curriculum Draft (SCTCD) consistent with each other?"

Method

Research Design

The design of this research was shaped in accordance with one of the curriculum models, the "Taba-Tyler Model." The design of the program and its required elements are determined by and defined by Demirel. (2012). The stages of the model are listed below (White, 1988, p. 26): a) Identifying the needs, b) Specifying the overall objective and formulation of objectives, c) Selecting and arranging the content, d) selecting and organizing the learning experiences, e) Conducting an evaluation, and if the result is insufficient, then, returning to the stages after specifying the overall objectives.

The steps taken to design the program for SCTCD are as follows.

a) Identifying the Needs**:** In a study conducted by Unlu, (2017) in order to demonstrate the educational needs of teachers' behaviors that support critical

thinking, it was revealed that the teachers in the research group are inadequate in demonstrating behaviors that support critical thinking in classroom environments and that they have educational needs in relation to this. The needs for teacher education in the research group are in the areas of open-mindedness, questioning of the accuracy and reliability of information, seeking causes and evidence, and high-level questioning.

b) Specifying the Overall Objectives and Formulation of Objectives**:** The overall objective of the SCTCD is to bring knowledge and awareness about the development of behaviors supporting critical thinking of primary education teachers. The program objectives are written using open-mindedness, inquiring about the accuracy and reliability of information, reason-evidence seeking and asking high-order questions. These themes constitute the dimensions of teacher behavior that support critical thinking.

c) Selecting and Arranging the Content: Varis (1996) emphasizes that the following criteria should be taken into consideration in the selection of content in education programs: 1. Community benefit, 2. Individual benefits, 3. Learning and teaching criteria, and 4. The place that content occupies in information. Firstly, field screening is conducted according to the results of a needs analysis. The subject area criteria about community benefits, individual benefits, learning and teaching criteria, and teacher behavior that supports critical thinking are emphasized in the selection of content. The content concerns topics from life that are accurate from a scientific point of view, suitable for the needs of the participants, and will create awareness of teacher behavior that supports critical thinking. In arranging the content of SCTCD, adult education principles were taken into consideration and an inquiry-based program approach is used. Attention was paid to the fact that the content is about topics containing each dimension of the teacher behaviors supporting critical thinking.

d) Selection and Organization of Learning Experiences: The learning and teaching process is designed in accordance with the content. In teaching learning process, it is preferred to use methods that enable participants to participate effectively in the educational process, such as creative drama, narration, questions and answers, group work, opinion development, case studies, large group discussions, conversation rings, creative writing, t-chart, and Socratic questioning.

e) Conducting an Evaluation: This is aimed at determining whether participants have acquired the pre-selected objectives in the program. For this purpose, the evaluation of objectives was conducted at the beginning, during and at the end of the application of program. Critical objectives representing the themes in formative evaluation were selected. Participants were expected to adapt these objectives to the situations they might encounter in life and to their professions. They were expected to evaluate those objectives about the given situation within the framework of determined criteria and teachers' opinions about education. The most suitable assessment and evaluation methods for objectives, content, and learning-teaching process were identified and tools were developed.

Research Sample

A purposive sampling method was used in this research. Experts in the study group voluntarily participated in this study. There were six experts in the research group consisting of two experts in curriculum development, two in assessment and evaluation, one in Turkish teaching and one in philosophy.

Research Instruments and Procedures

A reflective evaluation form was used as a data collection tool in the research, which was prepared by the researcher. It aimed to make a reflective evaluation of the program through revealing whether SCTCD's aspects were compatible with each other. For the evaluation of each item, "comment" and "suggestion" sections were included. The ratings for this section were set to "Yes", "Partly" and "No".

The reflective evaluation form was emailed to eight experts along with the documents that they would analyze (needs analysis, curriculum draft). Six experts returned their comments about REF via email along with the documents that they analyzed.

Data Analysis

"Frequency" was used as a descriptive analysis tool in the evaluation of field experts' opinions.

Results

In this part of the study, findings and interpretations of expert opinions on objectives, content, learning experiences, and assessment and evaluation dimensions are considered. A reflective evaluation form is used as a data collection tool in the expert evaluations about program elements. Expert opinions on dimensions are shown in Table 1 on objectives, in Table 2 on content, in Table 3 on learning experiences and in Table 4 on dimensions of assessment and evaluation. The explanations and adjustments to objectives, content, learning experiences and assessment and evaluation dimensions of the program are shown under the tables. The following acronyms are used in the tables: PR (participation rate); CD (curriculum development experts); MA (measurement and evaluation experts); P (philosophy field expert); and TL (Turkish language field expert).

Table 1

Expert Opinion on the Dimension of Objectives of SCTCD

Objectives	PR	CD	MA	P	TL	Total (f)
They are suitable for the needs of participants.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They are suitable for the level of participants.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They are expressed clearly and understandably	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They are suitable for the subject field.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They support high-level thinking	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-

Expert opinions on the dimension of objectives of SCTCD are listed in Table 1, where it is shown that six experts opined that the objectives were suitable to the needs of the participants, the level of participants and the characteristics of the subject area. The table also shows that the objectives were expressed clearly and that they supported high-order thinking.

In addition to the above evaluations, a curriculum development expert and an assessment and evaluation expert stated that that the program was written in such a way that most of the objectives contained single verbs and a single judgment. Further, they followed a systematic format from simple to complex, which complied with the principles of outcome writing, such as using a clear and understandable expression. In addition to these opinions, they had criticisms and recommendations for the expressions in the first theme, outcome number two and theme four, outcome number two.

The adjustments for objectives as a result of these criticisms and recommendations are as follows:
- Theme 1, Outcome 2: "Accepts that views may change with a change in evidence or reasons."
   Theme 1, Outcome 2: "Accepts that the change in views with the change in evidence or reasons is a natural process."
- Theme 4, Outcome 2: "Understands the importance of evaluating a subject from different perspectives."

   Theme 4, Outcome 2: "Evaluates a subject from different perspectives."

Table 2

Expert Opinion on Content Dimension of SCTCD

Content		LP	CD	MA	P	TL	Total (f)
It is consistent with objectives.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
It is suitable for the needs of participants.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
It is scientifically correct.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
The given information is valid.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
The specified time is sufficient.	Yes	1	2	1	1		5
	Partly	1	-	-	-		1
	No	-	-	-	-		-

Expert opinions on the content of SCTCD are listed in Table 2. Six experts stated that the content of the program is consistent with the objectives, the content is suitable to the needs of the participants, it is scientifically correct, and that the information contained in the content is valid. Five experts agreed that the allocated time was also sufficient for content and one curriculum developer opined that the allocated time was partly sufficient. In addition to this, they indicated that the guidelines given in the program led the instructor and were detailed, eliminating the impact on the educator during education and provided guidelines for the methods and techniques used in the activities. They further concurred that the selected texts in the program for the activities were well-chosen. As a result of the recommendations, the allocated time of the activities in the program were reviewed and increased.

Table 3

Expert Opinion on Learning Experiences Dimension of SCTCD

Learning Experiences		LP	CD	MA	P	TL	Total (f)
The selected principles and methods of teaching are suited to objectives.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
Learning activities are consistent with the objectives.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
They are line with the features of information in the content.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-
Teaching techniques and methods used were suitable.	Yes	2	2	1	1		6
	Partly	-	-	-	-		-
	No	-	-	-	-		-

Table 3 Continue

They are suitable for the level of participants.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They are suitable for the characteristics of the participants (primary education teacher).	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
They are suitable for the active participation of participants.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
The suggested materials are suitable.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
The specified time is sufficient for learning activities.	Yes	1	2	1	1	5
	Partly	1	-	-	-	1
	No	-	-	-	-	-
Activities were expressed clearly and understandably.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-

The experts agreed that Table 3 showed that the teaching principles and methods selected in the learning experiences were suitable to the objectives, consistent with the activities, and suitable to the characteristics of the content information. They further indicated that the teaching principles and methods used met the level of the participants, their characteristics, and active participation, and that the materials recommended for their learning experiences were suitable and expressed in a clear and understandable way. One curriculum development expert mentioned that the allocated time for activities was partly sufficient. However, some experts (i.e., curriculum development, evaluation and assessment) opined that the application of the methods in the process of teaching-learning that enable the participants to participate efficiently in the educational process, such as creative drama, narration, questions-answers, group study, developing opinions, case studies, big group discussions, circles of discussion, creative writing, t-table, six thinking hats, and Socratic inquiry would enrich the application of program.

Table 4

Expert Opinion on Assessment and Evaluation Dimension of SCTCD

Measurement and evaluation;		LP	CD	MA	F	TL	Total (f)
Assessment activities are suited to attainment.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
There is an evaluation activity for each sub-dimension.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
Assessment activities are appropriate for the level of participants.	Yes	2	2	1	1	6
	Partly	-	-	-	-	-
	No	-	-	-	-	-
The methods and techniques used are true.	Yes	2	1	1	1	5
	Partly	-	1	-	-	1
	No	-	-	-	-	-

When viewing Table 4, it is seen that all experts participating in the evaluation opined that the evaluation activities of the program are suitable to the objectives, that evaluation activities for each dimension are part of the program, and that the evaluations are suitable to the level of the participants. While five experts indicated that techniques used for the evaluation of themes were suitable, one assessment and evaluation expert felt that the methods and techniques used were partly accurate.

In addition to these views, the opinions and criticisms of one curriculum development and one assessment and evaluation expert are provided below:

- Different assessment and evaluation techniques, such as banner preparation, slogan writing, and creative drama can be used in the evaluations at the end of the themes.
- In the evaluations at the end of the themes, not only preparing questions, but also developing an answer key may be useful to teachers in thinking about the question they are preparing.
- Criteria of "open-mindedness" and "inquiring into the accuracy and reliability of information" themes in the scoring key may be rearranged.
- It was mentioned that the use of participant diaries throughout the process would be a suitable method for collecting opinions of the participants about the program and would be effective in providing information to the researcher about the functioning of the program.

The adjustments made for assessment and evaluation dimension as a result of these criticisms and recommendations are as follows:

- The assessment and evaluation techniques of banner preparation and slogan writing were added to the assessment and evaluation activities in the first and fourth themes. Creative drama method was eliminated from the assessment and evaluation dimension because it is used in every warming-up activity in the program.
- The expression "*reasons are given*" was added to the sentence of "Examples are provided for the concepts of fact, data and opinions," in the criteria about inquiring the accuracy and reliability of information theme on the scoring key, which is worth two points.

Discussion, Conclusion and Recommendations

This study is intended for the development and reflective evaluation of an education program for the behaviors of primary education teachers in supporting critical thinking. It is found that the outcome, content, learning-teaching states and assessment and evaluation dimensions of SCTCD were consistent with each other and suitable for their purposes. When literature in teacher education in Turkey is examined, an education curriculum developed for the behaviors of teachers, which support critical thinking, is not available. In addition, the topic of teacher behaviors that support critical thinking are neither included in the education programs of the education institutions that educate teachers, nor in programs about in-service education held by MoNE. On the other hand, the lack of such study brings along a serious contradiction, despite the fact that there are studies in the current literature intended for supporting the development of critical thinking of students at different levels. As previously stated, teachers are the foremost role models responsible for the development process of critical thinking of students (Ennis, 1983, 1993; Newmann, 1991; Paul 1991; Rovegno, 1992; Marzano, 1988; Fisher, 2001; Halpern, 2004; Doganay & Sari, 2012; Nosich, 2012). In this sense, only teachers with critical thinking skills and those who reflect such skills in their behaviors can be expected to educate critical thinking individuals (Marzano, 1988). For this reason, educating teachers to have critical thinking skills and continuity in supporting their professional development will also contribute to their performing their duties as critical thinkers (Washington, 1987; Gonzales-Rubio, 1988; as cited by Boonjeam, Tesaputa and Sri-ampai, 2017). Developing an education program for the behavior of teachers, which supports critical thinking, may be considered as one of the steps taken in the education of critical thinking of students and in the process of forming a free and democratic environment, which will be created by the existence of critical thinking individuals in society.

By looking at the results of the program based on the results of the research, the outcome dimension of SCTCD is suitable to the needs of the participants, their levels and the characteristics of the subject area. Objectives are expressed in a clear and understandable manner, and high-order thinking skills that support critical thinking are mentioned. There are 12 objectives in the program. The content dimension of SCTCD is consistent with objectives, suitable to the needs of the participants, and scientifically accurate. The content is structured under five themes (dimensions of open-mindedness, inquiring about the accuracy of information, inquiring about the reliability of information, reason-evidence seeking, and asking high-order questions) in SCTCD.

In the teaching-learning experiences dimension of SCTCD, the activities are consistent with the objectives. The teaching principles and methods selected are suitable to the objectives and characteristics of the content information, the characteristics of the participants (teacher), their levels and the active participation of the participants. In addition to this, according to the experts, using interactive teaching principles in the designing of activities will contribute to the increase in interaction throughout education. According to Gozutok (2007), using interactive teaching methods and techniques, such as creative drama, narration, questions-answers, group study, developing opinions, case studies, big group discussions, circles of discussion, six thinking hats, and Socratic inquiry, may contribute to the development of critical thinking skills and improve success. When

looking in the studies conducted in Turkey, it is seen that activities designed using interactive teaching and methods intended for the critical thinking education of prospective teachers in different branches, have a positive impact on their learning process (Cave, 1993; Gultekin, 2016; Aybek, 2006; Narin 2009; as cited by Unlu, 2017). The time required for the application of draft program is 15 classroom hours. Attention should be paid to the use of time given for activities during the implementation of the activities.

The assessment activities in the assessment and evaluation dimension of SCTCD are suitable to the objectives and levels of participants. The methods and techniques used are accurate. According to expert opinions, the use of participant diaries during the evaluation phase of process improved the evaluation dimension of the program, because it includes the participant opinions about the evaluation. The reflective diaries written by participants are found to be functional in understanding what they have experienced in the research process, because such diaries help them in offering alternative solutions to the practical problems they encountered or may encounter. Reflective diaries enable monitoring and recording of the development of participant's learning in teacher education. They also contribute to developing the thinking skills of the individual, especially the critical thinking, the participant's active participation in the learning process by evaluating the data obtained, and thereby the quality of learning (Dixon, 2009; Abbas and Gilmer, 1997; Burbach, 2004; Ekiz, 2006; Langer, 2002; as cited by Unlu, 2017). When examining the literature, it is seen that reflective diaries are widely used in teacher education and are an effective method in evaluating the efficiency of teacher education (Ekiz, 2006; Isikoglu, 2007; Anilan & Anagun, 2007; Lee, 2008; Tok, 2008; Koc and Yildiz; 2012). This situation in literature supports the expert opinions obtained at the completion of the research. Because active participation of participants is important for the effectiveness of education, attention must be paid to ensure that active participation supports critical thinking. It is proposed that SCTCD be developed for teachers working at different levels of education, because critical thinking education is a process that involves all levels of education.

In summary, as a result of the study, it is apparent that the outcome, content, teaching-learning experiences, and assessment and evaluation dimensions of SCTCD are consistent with each other and applicable.

References

Alkın-Sahin, S., & Gozutok, F. (2013, Ekim). Elestirel dusunmeyi destekleyen ogretmen davranısları envanteri (EDDODE): Gelistirilmesi ve uygulanması. [Critical Thinking Supportive Teacher Behaviors Inventory (CTSTBI): Its Development and Application]. *Eğitim Bilimleri Arastırmaları Dergisi-Journal of Educational Science Reseaarch, 3*(2), 223-254.

Ashton, P. (1988). *Teaching higher-order thinking and content: An essential ingredient in teacher preparation.* Ashton, Gainesville, FL: University of Florida.

Aybek, B. (2006). *Konu ve beceri temelli elestirel dusunme ogretiminin ogretmen adaylarının elestirel dusunme eğilimi ve duzeyine etkisi* [Effect of subject and skill-based critical thinking teaching on the tendency and level of critical thinking of teacher candidates] (Doctoral dissertation). Cukurova Universitesi, Sosyal Bilimler

Enstitusu, Adana. Retrieved December 3, 2014 from, https://tez.yok.gov.tr/UlusalTezMerkezi/

Battal, N. (2008). *Elestirel dusunme* [Critical thinking]. Retrieved, May 14, 2014 from http://iys.inonu.edu.tr/webpanel/dosyalar/840/file/elestirel%20dusunme.doc.

Beyer, B. (1987). *Practical strategies for the teaching of thinking.* Boston: Allyn and Bacon, INC.

Boonjeam W., Tesaputa K., & Sri-ampai A (2017). Program development for primary school teachers' critical thinking. *International Education Studies*, 10(2), 131-138. Retrieved October 12, 2017 from http://www.ccsenet.org/journal/index.php/ies/article/view/62188/35760

Cave, L. M. (1993). *The relationship of teacher behaviors and characteristics to critical thinking skills among middle-level student* (Doctoral dissetation) Available from ProQuest Dissertations and Theses datebase. (UMI No. 1278952).

Critical Thinking Skills and Teacher Education. ERIC Digest 3-88 (1988). ERIC Clearinghouseon Teacher Education Washington D.C: ERIC Document Reproduction Service No: ED297003. Retrieved 10 March, 2017 from https://www.ericdigests.org/pre-929/critical.htm

Cekic, S. (2007). *Matematik ogretmenligi lisans ogrencilerinin elestirel dusunme gucu duzeylerinin bazı degiskenlere gore incelenmesi* [Examining the critical thinking power of undergraduate students of mathematics teachers according to some variables]. (Unpublished master's thesis). Dokuz Eylul Universitesi, Egitim Bilimleri Enstitusu, İzmir. Retrieved 4 June, 2014 *from* https://tez.yok.gov.tr/UlusalTezMerkezi/

Cıkrıkcı, N. (1996) Elestirel dusunme: Bir olçme aracı ve bir arastırma [Critical thinking: A measurement tool and a research] *3. Ulusal Psikolojik Danısma ve Rehberlik Kongresi.* Adana: Çukurova Universitesi. 208-216.

De Bono, E. (2007). *How to have creative ideas: 62 exercises to develop the mind.* Random House.

Demirci, C. (2002). *Elestirel dusunme* [Critical thinking]. Retrieved 28 Febuary, 2015 from http://www.epo.hacettepe.edu.tr/eleman/yayınlar/elestirel-dusunme.doc.

Demirel, O. (2012). *Kuramdan uygulamaya egitimde program gelistirme.* [Curriculum development from theory to practice]. Ankara: Pegem Akademi.

Dewey, J. (1910). *How we think: A restatement of the relation of reflective thinking to the educational process.* Lexington,Boston: MA: Heath.

Doganay, A. (2001). Yaratıcı ogrenme [Creative learning]. A. Simsek (Ed.), *Sınıfta demokrasi (Democracy in the class)* (2nd ed.) (pp. 50-89). Ankara: Eğitim Sen Yayınları.

Doganay, A., & Sarı, M. (2012). Dusunme Dostu Sınıf Olçegi (DDSO) gelistirme calısması [A study of developing the Thinking-Friendly Classroom Scale (TFCS)]. *İlkogretim Online, 11*, 1, 214-229.

Ennis, R. (1985). *Goals for critical thinking curriculum. Developing minds: A resource book for teaching thinking.* Alexandria, VA: Association for Supervision and Curriculum Development, 68-72.

Ennis, R. H. (1993). Critical thinking assessment. *Theory Into Practice, 3*(32), 179-186.

Erdamar, G. K., & Alpan, G. B. (2017). Elestirel dusunme algısı: Lise ogretmenleri uzerine bir arastırma [The perception on crıtıcal thınkıng: A study on high school teachers] *Elektronik Sosyal Bilimler Dergisi,* 16(62), 787-800.

Facione, P. A., & Facione, N. C. (1998). *The california critical thinking dispositions inventory.* California: Academic Press. Retrieved 2 May, 2015 from http://www.insightassessment

Fisher, A. (2001). Critical thinking: An introduction, Cambridge: Cambridge University Press.

Glaser, R. (1983). *Education and thinking: The rol of knowledge, learning research and development* . Pennsylvania: Center University of Pittsburgh.

Gundogdu, H. (2009). Elestirel dusunme ve elestirel dusunme ogretimine dair bazı yanılgılar [Critical thinking and some misperceptions on teaching critical thinking] . *Celal Bayar Universitesi Sosyal Bilimler Enstitusu Dergisi, 7*(1), 57-74.

Gurkaynak, İ., Ustel, F., & Gulgoz, S. (2008). Elestirel dusunme (Guncellenmis Baskı) [Critical Thinking (Updated Edition)].İstanbul: Eğitim Reformu Girisimi Yayınları.

Gozutok, F. D. (2007). *Ogretim ilke ve yontemleri* [*Teaching principles and methods*]. Ankara: Ekinoks Yayınları.

Has, E. (2012). *Gazi Universitesi Fransız Dili Egitimi Anabilim Dalı ogrencilerinin elestirel dusunme eğilimleri uzerine bir çalısma* [Gazi University French Language Education A study on the tendency of students of the Department of Education to think critically] (Unpublished master's thesis). Gazi Universitesi, Egitim Bilimleri Enstitusu. Ankara. Retrieved 12 October, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Halpern, D. F. (2004). *Thoughts and knowledge: An introduction to critical thinking.* New Jersey-London: Lawrence Erlbaum Associates.

Kaloc, R. (2005). *Orta oğretim kurumu ogrencilerinin elestirel dusunme becerileri ve elestirel dusunme becerilerini etkileyen etmenler* [Critical thinking skills of secondary school students and factors affecting critical thinking skills]. (Unpublished master's thesis). Gazi Universitesi Eğitim Bilimleri Enstitusu, Ankara. Retrieved 10 August, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Koc, C. (2007). *Aktif ogrenmenin okuduğunu anlama, elestirel dusunme ve sınıf içi etkilesim uzerindeki etkileri* [Understanding the active student's reading, critical thinking, and influences on inter-class interaction] (Doctoral dissetation). DEU Eğitim Bilimleri Enstitusu, İzmir. Retrieved 6 March, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Kokdemir, D. (2003). *Belirsizlik durumlarında karar verme ve problem çozme* [Decision making and problem solving in uncertainty situations] (Doctoral dissetation). Ankara Universitesi Sosyal Bilimler Enstitusu, Ankara. Retrieved 12 October, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Lewis, A., & Smith, D. (1993). Defining higher order thinking. *Theory into practice*, *32*(3), 131-137.

Lipman, M. (2003). *Thinking in education.* Cambridge: Cambridge University.

Marzano, R. J. (1988). *Dimensions of thinking: A framework for curriculum and instruction*.The Association for Supervision and Curriculum Development, 125 N.West St., Alexandria, VA 22314-2798.

McPeck, J. (1981). *Critical thinking and education* . New York.: St Martins.

Ministry of National Education. (2005). *Ilkogretim 1-5. sınıf program-ları tanıtım el kitabı* [Primary education 1-5. classroom programs introduction handbook]. Talim ve Terbiye Kurulu Baskanlığı. Ankara: Devlet Kitapları Mudurlugu Basımevi.

Ministry of National Education. (2006). TEDP-Temel egitime destek projesi: Ogretmen eğitimi bileseni. Ogretmenlik mesleği genel yeterlilikleri [BTSP-Basic training support project: Teacher training component. General qualifications of teaching profession] Ogretmen Yetistirme ve Egitimi Genel Mudurlugu. Ankara: MEB Yayınları.

Ministry of National Education. (2008). *64 soruda ortaoğretime geçis sistemi ve seviye belirleme sınavı* [64 questions secondary school transfer system and leveling examination]. Ankara: MEB Yayınları.

Ministry of National Education. (2017a). *Mufredatta yenileme ve degisiklik çalısmalarımız uzerine* [On the renewal and change studies in the curriculum]. Talim ve Terbiye Kurulu Baskanlıgı. Ankara: Devlet Kitapları Mudurlugu Basımevi.

Ministry of National Education. (2017b). *Ogretmenlik Meslegi Genel Yeterlilikleri* [*General Qualifications of Teaching Profession*] .Ogretmen Yetistirme Genel Mudurluğu. Ankara: Devlet Kitapları Mudurluğu Basımevi. https://oygm.meb.gov.tr/meb_iys_dosyalar/2017_12/06172441_Ygretmenlik_Meslegi_Genel_Yeterlikleri.pdf adresinden 12 Haziran 2017 tarihinde alınmıstır.

Newmann, F., M. (1991). Classroom thoughtfulness and students' higher order thinking: common ındicators and diverse social studies courses. *Theory & Research in Social Education*, 19(4), 410-433, DOI: 10.1080/00933104.1991.10505649

Nosich, G. M. (2012). *Elestirel dusunme rehberi* [Critical thinking guide]. (B. Aybek, Trans.) Ankara: Anı Yayıncılık.

Ozcan, G. (2007). *Problem çozme yonteminin elestirel dusunme ve erisiye etkisi.* [Critical thinking and access influence of problem solving method] (Doctoral dissetation). Abant İzzet Baysal Universitesi, Sosyal Bilimler Enstitusu, Bolu. Retrieved 12 December, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Paul, R. (1989). Center for critical thinking and moral. *Critical thinking handbook: 6. -9. Grades A Guide for Remodeling Lesson Plans in Language*. Sonoma State University. ERIC Document Reproduction Service No: ED325805.

Paul, R. (1991). *Staff development for critical thinking: Lesson plan remodelling as the strategy Developing Minds (A Reseource Book For Teaching)* (Volume1). (A.C., Ed.) Alexandria, Virginia:A.L.Costa.

Paul, W. R., Elder, L., & Bartell, T. (1997). California teacher preperation forinstruction in critical thinking: Research findings and policy recommendations. Commission on Teacher Credentialing. Sacramento, California. Retrieved June 12, 2016 from http://www.criticalthinking.org/resources/books/

Paul, R., & Elder, L. (2008). Minik elestirel dusunme kılavuzu: Kavramlar ve araçlar. (M. B. Fidan, Çev.) Retrieved May 6, 2015 from Foundation for critical thinking: www.criticalthinking.org/resources/international/Turkish.cfm

Perry, D. K. (2014). *Exploring critical thinking skills among undergraduate agriculture education and studies students*. Iowa State University. Avaliable from: http://lib.dr.iastate.edu/etd. 08 Mart 2017.

Peterson, D, Kromrey, J., Borg, J., & Lewis, A. (1990). Defining and establishing relationship between essential and higher order teaching skills. *Journal of Educational Research, 84*(1), 5-12.

Ritchhart, R., & Perkins, D. N. (2005). Learning to think: The challenges of teaching thinking. In *The Cambridge handbook of thinking and reasoning* (pp. 775-802). Retrieved

December 12, 2016 from http://arq510002.paginas.ufsc.br/files/2011/04/Book-Chapter-32_Learning-to_Think_The-Challenges-to-Teaching-Thinking_-livro-The-Cambridge-Handbook-of-Thinking-and-Reasoning-Ritchhart_Perkins-1.pdf

Rovengo, I. C. (1992). Learning to teach in a field- based methods course: The developmentof pedagogical content knowladge. *Teaching and Teacher Education*, 8(1), 69-82.

Seferoglu, S. S., & Akbıyık, C. (2006). Elestirel dusunme ve oğretimi. [Critical thinking and instruction] *Hacettepe Universitesi Egitim Fakultesi Dergisi, 30*(30), 62-75.

Senemoglu, N. (1996). Yaratıcılık ve ogretmen nitelikleri [Creativity and teacher qualifications] . *Yaratıcılık ve Egitim Paneli, Kara Harp Okulu*. Ankara. Retrieved October 3, 2015 from http://yunus.hacettepe.edu.tr/n.senem/index_tur.html

Scriven, M., & Paul, R. (2007). Defining critical thinking: A draft statement for thenational council for excellence in critical thinking. Retrieved December 5 2015, from [Online]: http://www.criticalthinking. org/aboutCT/define_critical_thinkng.cfm.

Sahinel, S. (2007). *Elestirel dusunme* [Critical thinking] (2nd ed.). Ankara: Pegem Akademi Yayıncılık.

Sensekerci, E., & Bilgin, A. (2008). Elestirel dusunme ve ogretimi [Critical thinking and ınstruction]. *Uludag Universitesi Fen-Edebiyat Fakultesi Sosyal Bilimler Dergisi*, 15-43.

TTK (Talim ve Terbiye Kurulu). (1999). *16. Milli Egitim Surası* [16th National Education Council]. Retrieved May 12, 2016 from https://ttkb.meb.gov.tr/meb_iys_dosyalar/02113618_16_sura.pdf .

TTK (Talim ve Terbiye Kurulu). (2006). *17. Milli Egitim Surası* [19th National Education Council]. Retrieved May 12, 2016 from http://ttkb.meb.gov.tr/meb_iys_dosyalar/202113631_17_sura.pdf

TTK (Talim ve Terbiye Kurulu). (2014). *19. Milli Egitim Surası* [19th National Education Council]. Retrieved May 12, 2016 from http://www.meb.talimveterbiyekurulu

Tinning, R. (2002). Toward a 'modest pedagogy': Reflections on the problematics of critical pedagogy. *Quest, 54*, 224-240. Retrieved 14 March from https://www.coursehero.com/file/p7cokl5/Tinning-R2002-Toward-a-modestpedagogy-Re%C2%B2ections-on-the-problematics-of/

Tokyurek, T. (2001). *Oğretmen tutumlarının ogrencilerin elestirel dusunme becerilerine etkisi* [Impact of teacher attitudes on students' critical thinking skills] (Unpublished master's thesis) Sakarya Universitesi Eğitim Bilimleri Enstitusu, Sakarya. Retrieved December 8, 2016 from https://tez.yok.gov.tr/UlusalTezMerkezi/

Turnuklu, E. B., & Yesildere, S. (2005). Turkiye'den bir profil: 11-13 yas grubu matematik oğretmen adaylarının elestirel dusunme eğilim ve becerileri[A Profile from Turkey: Critical Thinking Dispositions and Abilities of Pre-Service Mathematics Teachers of 11-13 Year]. *Ankara Universitesi Egitim Bilimleri Fakultesi Dergisi, 38*(2), 167-185.

Unlu, S. (2017). *Elestirel dusunmeyi destekleyen ogretmen egitimi programının gelistirilmesi* [Development of Teacher Training Program Supported by Critical Thinking] (Doctoral dissertation)Ankara Universitesi, Egitim Bilimleri Enstitusu, Ankara.

Varıs, F. (1996). *Egitimde program gelistirme "Teori ve teknikler" [Program development in education "Theory and techniques"].* Ankara: Alkım Yayınları.

Walsh, D., & Paul, R. (1998). The goal of critical thinking: From educational ideal to educational reality. *Amerikan Federation of Teachers.*

White, R. (1998). *The ELT curriculum: Design, ınnovation and mangement.* Newyork: Wiley-Blackwell.

Wilks, S. (2005). *Critical and creative thinking: Strategies for classroom inquiry.* Portsmouth: NH. Heinmann.

Eleştirel Düşünmeyi Destekleyen Öğretmen Eğitimi Program Geliştirme Çalışması

Özet

Problem Durumu: Düşünme insanoğluna ait ve onu diğer varlıklardan ayıran en önemli özelliktir. Ancak insanoğlunun bu özelliği iyi düşünen bireyler olmaları için yeterli değildir. İdeal olan düşünme biçiminde beynimiz iyi ya da kötü, doğru ya da yanlış olan hakkında bir sonuca ulaşırken birtakım düşünme süreçlerini kullanmaktadır. Eleştirel düşünme de karmaşık düşünme süreçlerini kapsayan bir üst düzey düşünme biçimidir. 21.yy'da her alanda maruz kaldığımız yoğun bilgi akışı karşısında bu yöntemin dayandığı doğru olan bilgiye ulaşmada isteklilik, bilginin doğruluğunun ve güvenirliğinin sorgulanması, verileri ve göstergeleri doğru okuma, kanıtlara dayandırma, doğru soruları sorma gibi eleştirel düşünmenin tüm göstergelerini kullanarak aklın rehberliğinde karar vermek için zorunlu hale gelmiştir. Eleştirel düşünmenin dayandığı bu sorgulama süreci basta Psikoloji, Felsefe ve Eğitim olmak üzere farklı disiplin alanı uzmanlarının yoğun ilgisini çekerek bir çalışma alanı haline gelmesine neden olmuştur. Alanda yapılan uzun tartışmalar sonucunda uzmanların eleştirel düşünme tanımının becerilerden, tutumlardan ve eğilimlerden oluştuğu konusunda hemfikir olmaları eleştirel düşünmenin öğretilebilir olduğu görüşünü de ortaya çıkarmıştır. Alanda yapılan deneysel çalışmalar da, her düzeydeki öğrenme surecinde eleştirel düşünme eğitimlerinin öğrenciler üzerinde olumlu etkisi olduğu sonuçlarını ortaya koyarak bu görüşü desteklemiştir. Peki bu süreci kim ya da kimler yönlendirecektir? Öğretmen öğretim sürecindeki en önemli aktörlerden biri olduğundan öğrenciler için birincil konumdaki rol modedir. En önemli rolün öğretmene ait olması ile araştırmacıların üzerinde düşündükleri meydan okuyucu bir diğer soru da "Nasıl bir öğretmen?" sorusu olmuştur. Öğretmen sadece bilgiyi aktaran değil; tavır ve davranışlarıyla öğrencilere örnek olandır. Eleştirel düşünme alan uzmanları ancak iyi düşünebilen öğretmenlerin örnek rol modeller olarak iyi düşünebilen bireyler yetiştirebileceğini, bunun da alanda iyi yetişmiş, bilgili ve deneyimli öğretmenler tarafından sağlanabileceğine dikkat çekmektedirler. Türkiye' de alınan politika

kararlarında; eleştirel düşünme, eleştirel düşünme eğitim-öğretimi ve eleştirel düşünen öğretmen yetiştirmenin önemine yer verilmiştir. Ancak uygulamada yeterli ve etkili girişimlerin olmadığı, alınan ulusal ve uluslararası sonuçlarda hâlâ öğretmen yeterliği konusunun tartışıldığı görülmektedir. Bu sebeple, öğretmenlerin eleştirel düşünmelerini destekleyecek bir eğitim programının tasarlanması öğretmenlerin hem birer eleştirel düşünür olarak gelişimlerinin desteklenmesinde hem de eleştirel düşünen bir toplum yaratılması sürecinde atılan önemli adımlardan biri olacaktır.

Araştırmanın Amacı: Uluslararası eğitim araştırmalarında eleştirel düşünen bireylerin varlığı modern eğitim sistemlerine sahip ülkelerin en önemli eğitim çıktıları arasında gösterilirken, okul öncesinden yükseköğretim kademesine kadar tüm branşlardaki öğretmenlerin nitelikleri sürekli sorgulanmaktadır. Özellikle temel becerilerin öğrencilere kazandırılmasında ilköğretim kademesi ön plana çıkmaktadır. Eleştirel düşünebilen bireylerin yetiştirilmesinde gerekli olan eleştirel düşünen ve sınıf ortamında davranışları ile eleştirel düşünmeyi destekleyen öğretmenlerin varlığıdır. Türkiye'de uygulamadaki öğretim programlarının önemli çıktılarından birisi eleştirel düşünebilen öğrencilerdir. Ancak, program uygulayıcısı olan öğretmenlerin çoğunluğu bu yeterliliğe sahip olmadıkları gibi eleştirel düşünmelerini desteklemek amacıyla bir öğretmen eğitim programı da mevcut değildir. Bu alandaki ihtiyacı gidermek amacıyla yapılan bu araştırmanın amacı, ilköğretim öğretmenlerinin eleştirel düşünmeyi destekleyen davranışlarına yönelik bir eğitim programı hazırlamak ve yansıtıcı değerlendirmesini yapmaktır.

Araştırmanın Yöntemi: Araştırma, bir program geliştirme modeli olan "Taba-Tyler Modeli" doğrultusunda şekillendirilmiştir. Bu modelin hedefleri, verimlilik ve uygulanabilirlik açısından en iyi duruma getirilmiş eğitsel tasarım ve programlar ortaya koymaktır. Bu araştırmada da alandaki ihtiyaca yönelik yeni bir eğitim programı tasarlandığından Taba-Tyler Model'inin kullanılması tercih edilmiştir. Araştırmada, amaçlı örnekleme yöntemi kullanılmıştır.

Araştırmanın Bulguları, Sonuçları ve Öneriler: Araştırma sonucunda ortaya konan "Eleştirel Düşünmeyi Destekleyen Öğretmen Eğitim Program Taslağı [EDDÖEPT]'da beş tema altında (açıklık, bilginin doğruluğunun güvenirliğinin sorgulanması, açık fikirlilik, neden-kanıt arama ve üst düzey soru sorma) 12 kazanım yer almaktadır. Taslağın uygulanması 15 ders saatini gerektirmektedir. Yansıtıcı değerlendirme sonucunda taslak programın kazanımları, içeriği, öğrenme yasantıları ve ölçme değerlendirme boyutları kendi içlerinde tutarlı ve programın genel amacına uygun bulunmuştur. Eleştirel düşünme eğitiminin tüm eğitim kademelerini kapsayan bir süreç olmasından dolayı EDDÖEP'nin farklı eğitim kademelerindeki öğretmenler için de geliştirilmesi önerilmektedir.

Anahtar Kelimeler: Eleştirel düşünme öğretimi, öğretmen davranışları, öğretmen eğitimi, yansıtıcı değerlendirme.

Appendix A
Table 5
Supporting Critical Thinking Curriculum

Subject Field	Objectives	Content	Learning Experiences	Evaluation	Materials	Time
General Aim: The overall objective of the Supporting Critical Thinking Curriculum is to gain knowledge and awareness about the development of primary school teachers' behaviors supporting critical thinking.						
Open Mindedness	Meeting whole group; Warm up	Throw a ball	Creative drama	-	Ball	15 min.
	1. Question the ideas accepted by the majority of society.	Interrogating Ideas (Employee in the Street Children)	Opinion Development	Worksheets (formative assessment) Summative assessment	Laptop, projector, slides, writing board, board pen, speech object, paper.	75 min.
	2. Realize that it is a natural process when the evidence or reasons change; change the views.	Our Views Can Change (Syrian Immigrants, Compactness)	Conversation Circle			
			Case Study Analysis	Summative assessment		
Asking High-Level Questions	Getting to know group members; Warm up	Secret Skills	Creative drama	-	Paper, pencil	10 min.
	1. Explain the proper question properties of the reason.	The Importance of Asking Questions	Whole Group DiscussionConversation Circle	Summative assessment	Laptop, projector, slides, writing board, board pen, speech object, paper, short film.	100 min.
	2. Ask questions that encourage high-level thinking in the teaching process.	Understanding Basic and Strong Concepts	Case Study Analysis Video tracing Small Group Study	Worksheets (formative assessment) Summative assessment		
Openness	Understanding disabled people; Warm up	Understanding Barriers	Creative Drama	-	-	10 min.
	1. Using different methods to explain unfamiliar concepts and definitions.	Teaching Unintelligible Concepts Understanding Basic and Strong Concepts	Role Play, Conversation Circle, Case Study Analysis, Individual Study, Socratic Inquiry, Creative Writing, Group Work	Worksheets (formative assessment) Summative assessment	Laptop, projector, slides, writing board, board pen, speech object, paper, colorful pen, short film, texts.	90 min.

Table 5 Continue

Subject Field	Objectives	Content	Learning Experiences	Evaluation	Materials	Time
Questioning of the Accuracy and Reliability of Information	Draw attention to the subject; Warm up	Why "Googling" Is Not Enough.	Text Reading	Summative assessment	Laptop, projector, slides, writing board, board pen, speech object, paper, colorful pen, short film, hats (white, black, yellow, blue, red), texts.	100 min.
	1. Know the scope of concepts that deliver reliable information	What are the Data, Facts and Findings? Sources of Information	Presentation, Questions-Answers Text Reading Conversation Circle			
	2. Organizing activities to inquire about the authenticity of information	Trustworthy Sources of Information	Conversation Circle Group Work			
	3. Applying activities to evaluate a topic at different points of view.	Evaluating Different Perspectives	Six Hat Thinking Reading Text			
	4. Give examples of events, events and findings	What are the differences between knowledge, cases, findings?	Group Work T- chart	Worksheets (formative assessment) Summative assessment		
Causes - Evidence Seeking	Questioning situations where participants do not like the environment they live in; Warm up	Complaint Box	Creative drama	-	Box, paper, pencil	10 min.
	1. Conduct activities to decide whether the reasons for the events/circumstances in the written and/or oral material are persuasive.	Presenting the Right Justification	Question-Answer Case Study Analysis	Summative assessment	Laptop, projector, slides, writing board, board pen, speech object, paper, letters, short film.	95 min.
	2. Conduct studies that explain the causes and evidence of ideas.	Why Do I Think So?	Conversation Circle			
	3. Organize incidents to question the cause-and-effect relationships of a controversial issue, such as discrimination, war, bloodshed, honor killings.	Cause - Result Relation	Creative writing	Summative assessment Worksheets (formative assessment)		

Analytical Weighting Scoring for Physics Multiple Correct Items to Improve the Accuracy of Students' Ability Assessment

Wasis[1] , Kumaidi[2], Bastari[3], Mundilarto[4], Atik Wintarti[5]

ARTICLE INFO	ABSTRACT
	Purpose: This is a developmental research study that aims to develop a model of polytomous scoring based-on weighting for multiple correct items in the subject of physics. Weighting was analytically applied based on question complexity and imposed penalties on wrong answers. **Research Methods:** Within the development model, Fenrich's development cycle, consisting of analysis, planning, design, development, implementation, evaluation, and revision, was selected throughout the cycle. The multiple correct items used have 3–4 options. The items were implemented to 140 upper secondary school students and 410 first-year undergraduate students. The students' physics ability was analyzed

using the Quest program, and the results of dichotomous and polytomous scoring were compared. **Findings:** The results of this study showed that the analytical weighting scoring based on a complexity and penalty system on the developed assessment items generated a higher number of scoring level categories (three to seven categories) than that of dichotomous scoring (only two categories), estimated students' physics abilities more accurately and in greater detail, had an approximate distribution closer to the normal distribution, and produced a standard deviation smaller than that of dichotomous scoring. Thus, the analytical weighting scoring for multiple correct items in this study was able to produce a more accurate estimation of physics ability than those using dichotomous scoring. **Implications for Research and Practice**: It is recommended that the assessment of physics ability using multiple-correct items on a large scale can apply the analytical weighting scoring based on the complexity of the content and a penalty system.

---

[1] Corresponding Author: Department of Physics, Universitas Negeri Surabaya, INDONESIA. e-mail: wasis@unesa.ac.id, ORCID: 0000-0002-4437-5141

[2] Departement of Psychology, Universitas Muhammadiyah Surakarta, INDONESIA. e-mail: kum231@ums.ac.id, ORCID: 0000-0002-2736-6139.

[3] Balitbang Kemendikbud, INDONESIA. e-mail: bastari@kemdikbud.go.id, ORCID: 0000-0003-3442-3063.

[4] Department of Physics, Universitas Negeri Yogyakarta, INDONESIA. e-mail: mundilarto@uny.ac.id, ORCID: 000-0003-2891-4317.

[5] Department of Mathematics, Universitas Negeri Surabaya, INDONESIA. e-mail: atikwintarti@unesa.ac.id, ORCID: 0000-0001-6514-1881.

Introduction

The study of scoring on selected-response items is still continuing in educational assessment in recent decades. Previous studies have examined this issue for the following purposes: to evaluate teachers' competency (Martin & Itter, 2014), to develop construction and scoring on selected-response items (Emaikwu, 2015), to compare the assessment using the question of multiple choice and constructed response (Hickson, Reed, & Sander, 2012; Stankous, 2016), to obtain immediate feedback assessment techniques (Merrel et al., 2015), and to propose a more comprehensive framework of writing a more sophisticated format of selected response items in hopes of reduced guesswork (Bush, 2015). However, all of those purposes are designed to improve the validity and reliability of the selected response items developed within those studies (Ali, Carr & Ruit, 2016)

Despite the various weaknesses of selected-response assessment items, including the issue of validity and reliability, facts show that this method of testing is still dominantly used, especially in large-scale tests with speedily delivered results (Oosterhof, 2003; Rodriguez, 2005; Merrel et al., 2015). This is because the execution of the test with selected-response items takes less time, requires a quick process, is easy to score, and has a high degree of objectivity (Wooten et al., 2014; Baghaei & Dourakhshan, 2016).

However, the use of selected-response items to measure, especially, the mastery of physics ability, has a number of weaknesses, particularly when the scoring is done using a dichotomous model. The topics of physics are known as a continuum, ranging from very simple to very complicated cognition, so the assessment instrument of physics ability, should include a representative scope of content and competence (Klein, et. al., 2017). Thus, when a person seeks to understand or master the subject of physics, his/her understanding or mastery can be in a position among all the possible positions in the continuum. The person's ability, position can be anywhere and is not limited to the lowest or highest position. Therefore, if a person's ability is assessed, then the result of his judgment cannot be forced to be dichotomous, which is to say high or low. Furthermore, the problem-solving in physics also has a number of stages (Adeyemo, 2010). Therefore, if a person's ability to solve a physics problem is assessed, the results of his or her judgment should not only be indicated in the final result but should demonstrate a knowledge of all stages leading up to find the outcome. The results of the assessment should be able to describe the student's achievements at each stage. Moreover, the phases of solving physics problems generally reflects various levels of complexity. When a person is able to complete a stage, then the assessment of his or her performance should also vary depending on the weight of its complexity. Based on the reasons listed above, the scoring of physics ability should be done analytically with regard to the results of each stage. According to Wiseman (2012) analytical scoring provides better assessment results than holistic scoring.

The use of selected-response items in assessing physics ability often causes students to answer the questions by guessing. To reduce the inaccuracies that may result from this practice, the incorrect answer needs to be punished under the

understanding that one should be cautious because penalties for incorrect answers could intimidate students (Holt, 2006). In analytical scoring, penalties can be integrated by weighting of each stage of solving physics problems. Therefore, another scoring model of selected response questions should be developed to test physics ability that considers completion stages, the weighting of each of those stages, and the responsibility to answer questions. It is hoped that with such scoring patterns that even though the assessment uses selected responses item, the judgment of the results will have many categories, or in other words, that they may be polytomous. Bond & Fox (2007) stated that adding a number of categories to the scoring will increase the reliability of the measurement. If the measurement of the results is reliable, then the assessment based on the results will be more accurate.

The current study aims to develop a polytomous scoring model based on weighting, focusing on complexity and assigning penalties for the incorrect answer of each stage of solving a given physics problem. In order for each stage to be analytically scored, an alternative multiple-choice question, or so-called as multiple-correct item, was selected. Questions in multiple-correct items are more efficient because they can generate better memory retention and provide more correct information (Bishara & Lanzo, 2015). Also, one question can be designed to measure multiple capabilities of the same dimensions (Haladyna et al., 2002), and it was also empirically proven to be more reliable than multiple choice items (Frisbie, 1992). Thus, the results of this study are expected to produce a model of polytomous scoring for multiple correct items so as to be able to estimate students' ability in physics more accurately.

Method

Research Design

This is a development research study following Fenrich's model (2004) (see Figure 1). The result of this development is a polytomous scoring model for multiple correct items. The model encompasses the phases of analysis, planning, design, development, and implementation; each of which provides an opportunity for possible evaluation and revision.
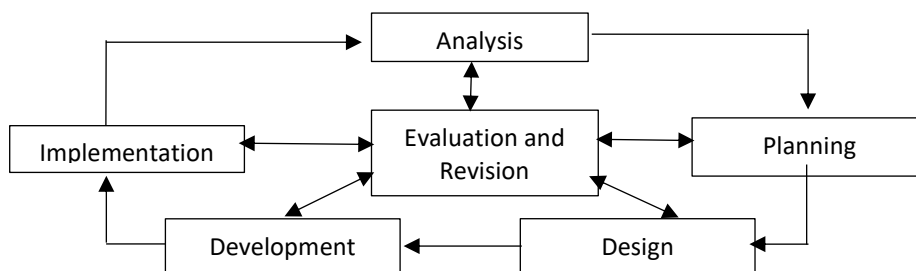

Figure 1. Fenrich model development cycle (Fenrich, 2004)

The analysis was carried out by focusing on the scoring method, which uses a partial-credit model instead of the dichotomous scoring model that only considers two options: Yes or No. The use of a partial-credit model on multiple choice questions has been investigated by several scholars (see e.g. Wright & Masters, 1982; Grunert et al., 2013 and King et al., 2004). However, the scores of each option are based on partial assumptions, for example 0.25, 0.50, 0.75, and 1. This study's analytical scoring is designed according to the weight of the complexity of each option and imposes penalties for incorrect answers. This relates to Ayedemo's (2010) statement suggesting that problems in physics typically require multi-stage problem-solving strategies. Therefore, the scoring needs to consider the weight of each stage. Also, this follows Holt's (2006) suggestion that the prudence of the punishment can reduce the carelessness of students in guessing as they answer multiple choice questions.

The scoring design described above can be applied to multiple-correct items, i.e. multiple-choice items with more than one correct option. This item format was selected because it can measure the number of stages of problem solving within physics. With multiple-correct item formats, it is possible that the stem of the problem in the items is in the form of a physics problem while the options are in the form of stages of problem solving. All options can be designated as being correct or incorrect. Based on previous research, it has been proven that multiple-correct items are more reliable and provide more advantages than traditional multiple-choice items (with only one correct option) (Frisbie, 1992; Haladyna et al., 2002; and Bishara & Lanzo, 2015).

The process of evaluation and revision were also carried out reciprocally throughout the development and test phases as shown by the two-way arrows in Figure 1. The results of the analysis emphasize the planning, design, and development of the problem along with the scoring model. The results of the four steps were evaluated by the experts, which in this case are physics lecturers through the validation mechanism. Validation results were used to revise the design of problems in the multiple correct items developed as well as their scoring. When the implementation phase was conducted, the activities of evaluation and revision were also conducted so that the representative respondents, either senior high school students (high, middle, and low) or undergradute students (physics, chemistry, and biology majors) were equally obtained.

Research Sample

The sample for this study were 140 high school students, coming from three different high schools in Surabaya, Indonesia and 410 first-year undergraduate students from four different majors within the Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Indonesia.

Research Instruments and Procedures

There were two types of multiple-correct item format developed in this study: items with three options and items with four options. The number of items for each

type is fifteen. The following is an example of a multiple-correct item with three options complete with its weighted analytical scoring model developed in this study.

A 200-gram bullet is fired vertically upward from the ground with a speed of 60 m/s. If the gravitational acceleration g = 10 m/s², then:

(a) The maximum height reached by the bullet is 180 meters
(b) At the highest point, the energy of the bullet is 360 joules
(c) At its highest point, the bullet has neither speed nor acceleration

Because this is in the form of a multiple-correct item, then each option is likely to be true and the participants are permitted to choose more than one option. The analytic scoring guide for each of the above responses is shown in Table 1.

Table 1

Analytical Scoring Guide

Completion stage	Analytic score for each stage	Total score for each option
The bullet reaches its highest point when $v_t = 0$	1	
The maximum level reached is given by $\frac{v_0^2}{2g}$	1	3
$h_{max} = \frac{60^2}{2.10} = 180$ meter		
➜ Proposition given in option (a) is correct	1	
At the highest level, the bullet will not be moving; thus, for a moment, $v_t = 0$, so $E_{K\ highest} = 0$.	1	
The bullet energy's at its highest level = potential energy = initial kinetic energy. $E_{highest} = E_p = mgh = (0.2)(10)(180) = 360$ joule	1	3
or $E_{highest} = E_{Ki} = \frac{1}{2}\,mv^2 = (\frac{1}{2})(0,2)(60)^2 = 360$ joule	1	
➜ Proposition given in option (b) is correct		
The object moves vertically up and then moves down to the earth. At its highest point, the object has no movement for a moment, so it has no speed.	1	
The object has a constant acceleration, including when it is at its highest point, that is, gravitation acceleration.	1	2
➜ The proposition given in option (c) is incorrect		

Based on the above scoring guidelines, if the test taker chooses option (a), he/she gets a score of 3; does not select option (a), gets a score (-3); selects option (b), gets a score of 3; does not select option (b) gets a score (-2); selects option (c), gets a score of

(-2); and does not select option (C), gets a score of 2 because the proposition in option (c) is incorrect. The following scoring categories are obtained for all possible response patterns to this question as follows:

Table 2

Polytomous Scoring Based on Analytical Weighting and Penalty System

Pattern of responses			Score for each option			Total score	Category
(a)	(b)	(c)	(a)	(b)	(c)		
✓			3	-3	2	2	4
	✓		-3	3	2	2	4
		✓	-3	-3	-2	-8	1
✓	✓		3	3	2	8	6
✓		✓	3	-3	-2	-2	3
	✓	✓	-3	3	-2	-2	3
✓	✓	✓	3	3	-2	4	5
			-3	-3	2	-4	2

Table 2 shows that in the polytomous scoring model developed in this study, the ability of test participants can be assigned to one of six categories, namely categories 1 through 6. Whereas if using dichotomous scoring, there are only two possible categories, namely score 1 for participants who choose options (a) and (b) but not (c) and score 0 for participants who choose any other combination.

The developed model of polytomous scoring, as illustrated in Table 1, has the following stages: (a) developing analytic scoring guidelines for each problem-solving step by considering the weight of its complexity, (b) identifying all possible response patterns, (c) scoring based on the response patterns and the assigned weight of each stage, (d) assigning a positive score for a correct answer and a negative score for an incorrect answer, (e) calculating the total score by adding the scores of each option/step, and (f) correlating the total score equivalent to a category or level.

Data Analysis

After the responses of the test participants were categorized according to the developed model, they were then analyzed by the Quest program (Adams & Khoo, 1996) in order to estimate the participants' abilities (θ) and to obtain the standard estimation error. The Quest program was selected because it is relatively simple due to its being limited to only one parameter of items, which is the Rasch model which applies an item response theory approach. The IRT approach and the Rasch model were chosen because this a study that assumes that guessing is a part of a student's ability and that all items that fit the model have equivalent discriminations so that items are only described by a single parameter (1 PL).

Results

Before analyzing the participants' aptitude and the difficulty of the items using Quest (Adams & Khoo, 1996), the authors first conducted the so-called 'fit item analysis' using an infit and outfit mean square and an infit and outfit t, as shown in Table 3.

Table 3

Fit Item Analysis

Scoring	Infit mean square	Outfit mean square	Infit t	Outfit T
Dichotomous	1.00	1.04	0.10	0.27
Polytomous	1.00	1.00	-0.02	-0.04

The data or responses are said to fit within the Rasch model when the infit and outfit mean's square value is close to 1 and the infit and outfit t are near 0 (Adams & Khoo, 1996). Table 3 shows that the responses of the test participants used in this study fit within the Rasch model, although the *fit* of dichotomous scores was lower than that of the polytomous scores.

The Quest program's estimations for a number of the upper secondary school student participants are shown in Table 4, and the summary of the mean is presented in Table 5.

Table 4

The of Physics-Aptitude Estimation of Several Participants

	Dichotomous Scoring					Polytomous Scoring			
Name	Score	Score Max	Estimate	Error	Name	Score	Score Max	Estimate	Error
14	3	15	-1.62	0.73	14	23	69	-0.46	0.22
27	3	15	-1.62	0.73	27	37	69	0.15	0.21
32	3	15	-1.62	0.73	32	36	69	0.11	0.21
33	3	15	-1.62	0.73	33	32	69	-0.06	0.21
68	3	15	-1.62	0.73	68	33	69	-0.02	0.21
76	3	15	-1.62	0.73	76	20	69	-0.60	0.23
78	3	15	-1.62	0.73	78	29	69	-0.19	0.21
81	3	15	-1.62	0.73	81	30	69	-0.15	0.21
82	3	15	-1.62	0.73	82	29	69	-0.19	0.21
83	3	15	-1.62	0.73	83	25	69	-0.37	0.21
103	3	15	-1.62	0.73	103	29	69	-0.19	0.21
105	3	15	-1.62	0.73	105	27	69	-0.28	0.21
111	3	15	-1.62	0.73	111	30	69	-0.15	0.21
112	3	15	-1.62	0.73	112	32	69	-0.06	0.21
118	3	15	-1.62	0.73	118	42	69	0.37	0.21
140	3	15	-1.62	0.73	140	27	69	-0.28	0.21

Based on the data in Table 4, it can be seen that the polytomous scoring method developed for this study yields a better capability to estimate the participants' physics ability than the estimations made using dichotomous scoring. This can be seen from the ability to estimate the ability of test participants in more detail. Of the 140 participants in the upper secondary school sample, sixteen students in the dichotomous scoring had the same estimated aptitude of -1.62 logits, but on the polytomous scoring of those sixteen students, the participants had different aptitude estimates ranging from -0.60 to 0.37 logits. On the other hand, of the 410 of undergraduate-student participants, as many as 151 participants are estimated to have the same aptitude, measured at -3.41 logits, the lowest ability. Meanwhile, when their abilities were scored using polytomous scoring, there was only 1 person who had the lowest aptitude, while the others had a higher estimation of their abilities measured by various scores.

Polytomous scoring provides a more accurate estimation of the participants' capabilities. This can be seen from the standard deviation (SD) values as shown in Table 5.

Table 5

Means of Estimation and Standard Deviation (SD) for the Participants' Ability

Participants	Dichotomous scoring			Polytomous Scoring		
	Number of categories	Mean of ability	SD	Number of categories	Mean of ability	SD
Upper secondary students	2	-2.47	0.56	3–7	-0.07	0.22
Undergraduate students	2	-2.76	0.81	5–10	-0.08	0.18

Table 5 shows that the standard deviation (SD) of ability in the polytomous scoring, which has more categories, is always smaller than the standard deviation resulting from dichotomous scoring. Those results were consistently obtained for both upper-secondary and undergraduate-student participants. Based on the above results, it can be concluded that the greater number of categories in a scoring system, the smaller the standard deviation of estimation is, meaning that the result of the assessment will contain fewer errors and be closer to the actual ability.

Quest analysis also maps the estimation of the students' ability and item difficulty for dichotomous and polytomous scoring, as shown in Figures 2 (a) and (b) for upper secondary school student and Figures 3 (a) and (b) for first-year undergraduate student participants.

Figure 2. The mapping of the estimation of the upper secondary school students' ability and the item difficulty for (a) dichotomous scoring and (b) polytomous scoring
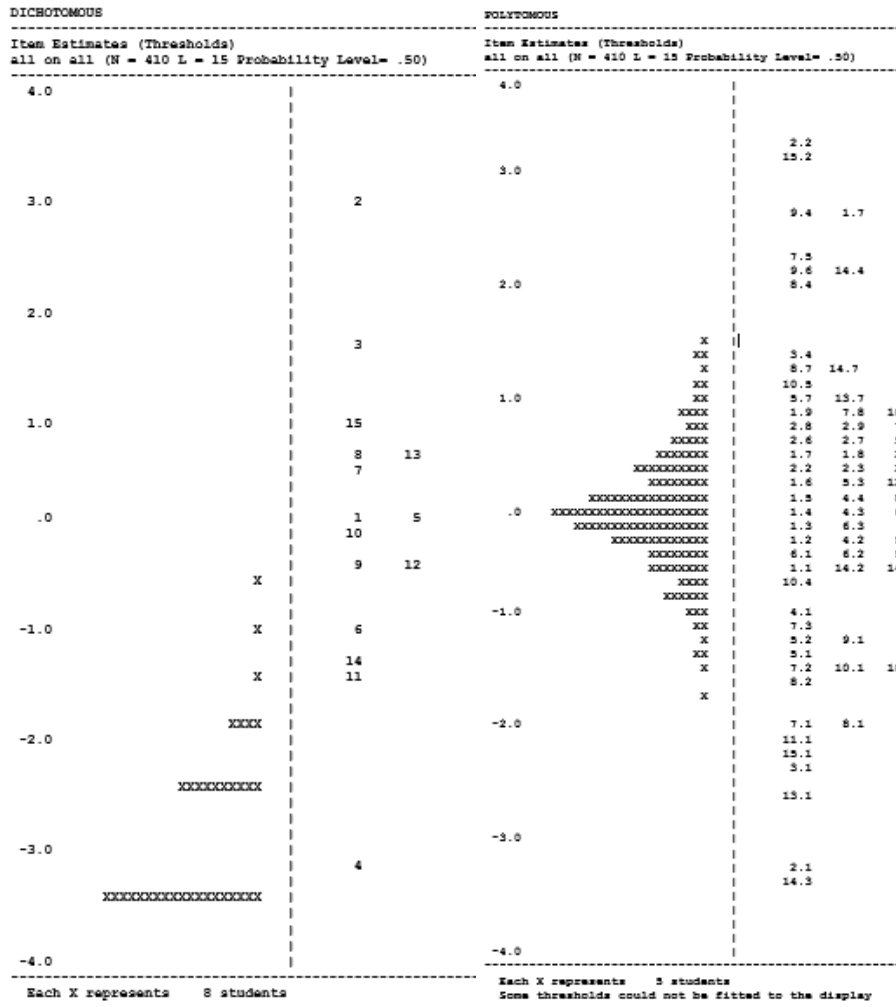
Figure 3. The mapping of the first-year undergraduate student s' ability and level of difficulty of the items based on the results of (a) dichotomous scoring and (b) polytomous scoring

Discussion, Conclusion, and Recommendations

Overall, the results of this analysis as indicated by Table 4 show that polytomous scoring's having more categories than dichotomous scoring makes the measurement results more accurate. This is analogous to measuring tools used in everyday life: i.e. the more strips or scale lines that a measuring device has, the more thorough the

measurement. On a ruler a distance of 1 mm is represented only by two lines, while on a slide rule, a distance of 1 mm is subdivided into 10 or 20 lines. Thus,the measurement resulting from the slide rule can become more accurate than that of the ruler. Bond & Fox (2007) stated that scoring within many categories can estimate the ability of test participants better than an assessment with fewer categories. Donoghue (2005), in an article on reading scores, reported that polytomous scoring yields an average of about two to three times the average function of dichotomous scoring information. The greater the information function, the more accurate the resulting estimate (Hambleton et al., 1991, Lin, 2008).

As mentioned above, the improvement of accurary in scoring physics ability assessments can also be claimed as the contribution of the analysis of the suitability of the developed scoring models with the items formats and characteristics of the physics content. Problem solving is a fundamental part of studying physics, and the students do not solve most problems at the desired level of proficiency (Redish et al., 2006). When the stages of problem-solving are formulated as self-selected options in multiple-correct items and each answer in the option is analytically assessed based on the degree of difficulty, the student's ability can be measured in more detail than if judged only by dichotomous scoring based on the final result of their problem solving. Haladyna et al. (2002), who reviewed various forms of multiple-choice questions, stated that in the development of choice questions, it is not only important to consider the content examined in the questions developed but also the methodological viewpoints used in the development process. If multiple-choice questions are methodically designed to analytically record the students' thinking processes and record such stages, then such questions will be able to assess students more accurately.

Figures 2 and 3 show that the estimation of the participants' aptitude under polytomous scoring tends to illustrate the real condition of a group of participants because it has a normal distribution with an average value close to 0, whereas the mean estimation of participants' ability under dichotomous scoring tends to tilt to the left. The polytomous item-person maps shown in Figures 2 and 3 demonstrate how the test is well matched to the sample (Bond & Fox, 2007). It is also known that the estimation of item difficulty by using polytomous scoring is more accurate and detailed than the dichotomous scoring results although both are relatively consistent in estimating the item difficulty of the most difficult and easiest items. Figures 2 and 3 also demonstrate that the estimation of the degree of difficulty of the problem within polytomous scoring stretches further and generates a more detailed report than that of dichotomous scoring. Like a ruler, the polytomous scoring system with the analytical weighting approach developed in this study produced a more detailed ruler on a smaller-scale as opposed to the dichotomous scoring's basic ruler with its wider-scale.

For the upper secondary school student participants, Figure 2 shows that both dichotomous and polytomous scoring indicated that item number 9 was the most difficult item. However, for the undergraduate participant, Figure 3 shows the most difficult problem was item number 2. For the easiest problem, both undergraduate student and upper secondary school student respondents showed a difference between the two scoring methods. Figure 2 shows that the most difficult item in

dichotomous scoring was item number 14, while the most difficult item in polytomous scoring was item number 10, especially in regard to reaching level 1 (10.1). For undergraduate students, the easiest item in dichotomous scoring was number 4, while in polytomous scoring, the easiest item was number 14, especially 14.3. The above results show that the scoring of responses of the difficult items, often coming from the respondents with high ability, are relatively consistent both dichotomously and polytomously. The consistent responses can be interpreted as the product of an actual thinking process among the respondents, instead of simply guessing. Meanwhile, the responses to the easy problem might come from the respondents with low ability, especially since they were empirically proven to be inconsistent, demonstrating that they might be the result of guessing.

The results of this study indicate that polytomous scoring is able to estimate more the abilities of the participants in more detail, yield a smaller standard deviation than that of dichotomous scoring, and cause the distribution of the estimates of participants' ability closer to a normal distribution. The above results serve as empirical evidence that weighted polytomous scoring is able to estimate students' physics ability more accurately than dichotomous scoring. The results of this study are in keeping with the results of Jiao et al. (2012) who applied computerized adaptive tests (CAT) on a large scale and found that polytomous scoring yielded a slightly more precise estimate of ability than dichotomous scoring. Also, the results are similar to the results of research of Grunert et al. (2013), which show that polytomous scoring on a chemical exam yields a higher percentage average score than dichotomous scoring. The results of this study along with the results of research by Jiao et al. (2012) and Grunert et al. (2013) have become empirical evidence of the statements of Baker et al. (2000), Tognolini & Davidson (2003), Wu (2003), and Bond & Fox (2007) that argue that multiple-category scoring (polytomous) can estimate the test-takers' aptitude better than that of dichotomous scoring.

The new finding of the development of polytomous scoring based on analytical weighting and a penalty system developed in this study as compared to previous polytomous scoring, especially in estimating students' physics ability, is its ability to appreciate respondents' thinking processes on a step-by-step basis when they are selecting response items. Because the appreciation for every step also considers the complexity of each option, then polytomous scoring system will not only be more accurate than that of dichotomous scoring but also becomes more equitable in estimating respondents' abilities and ensures that points are awarded for ability rather than luck (Wiseman, 2012; Holt, 2006). Thus, it can be emphasized that the polytomous-scoring system based in an analytical weighting and penalizing system using multiple-correct items is able to estimate physics ability more accurately than dichotomous scoring. This conclusion is based on the empirical facts that polytomous scoring is able to estimate more detailed capabilities with smaller average standard deviations and approximate distributions closer to a normal distribution than dichotomous scoring.

In summary, the analytical-weighting approach to scoring multiple-correct items in this study was able to produce a more accurate estimation of physics ability than

the dichotomous scoring approach. This is indicated by the findings that the scoring model estimated students' physics abilities in more detailed and with greater accuracy, had an approximate distribution closer to the normal distribution, and produced a standard deviation smaller than that of dichotomous scoring.

Based on the results of this research, it is recommanded that assessments of physics ability that use selected response items, especially multiple-correct items, on a large scale can apply the analytical weighting scoring based on the complexity of content and a penalty system. However, this scoring model might be difficult to apply manually on a large scale because it requires much more time than the standard multiple-choice item. Therefore, further research is needed to develop application software that supports this scoring model.

Acknowledgements

References

Adams, R.J., & Khoo, S.T. (1996). *Quest* (Computer software). *The interactive test analysis system*. Victoria: Acer.

Adeyemo, S. A. (2010). Students' ability level and their competence in problem-solving task in physics. *International Journal of Educational Research and Technology, 1*(2), 35 – 47.

Ali, S. H., Carr, P. A., & Ruit, K. G. (2016). Validity and reliability of scores obtained on multiple-choice questions: Why functioning distractors matter. *Journal of the Scholarship of Teaching and Learning*, *16* (1), 1-14.

Baghaei, P., & Dourakhshan, A. (2016). Properties of single-response and double-response multiple-choice grammar items. *International Journal of Language Testing, 6* (1), 33-49.

Baker, J.G., Rounds, J.B., & Zeron, M.A. (2000). A comparison of graded response and rasch partial credit models with subjective well-being. *Journal of Educational and Behavioral Statistic, 25*(3), 253-270.

Bishara, A. J., & Lanzo, L. A. (2015). All of the above: When multiple correct response options enhance the testing effect. *Journal Memory*, 23(7), 1013-1028.

Bush, M. (2015). Reducing the need for guesswork in multiple-choice tests. *Assessment & Evaluation in Higher Education*, *40*(2), 218-231.

Bond, T. G., & Fox, C. M. (2007). *Applying the rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah: Lawrence Erlbaum Associates, Publishers.

Donoghue, J. R. (2005). An empirical examination of the IRT information of polythomously scored reading items under the generalized PCM. *Journal of Educational Measurement*, *31*(4), 295-311.

Emaikwu, S. O. (2015). Recent issues in the construction and scoring of multiple-choice items in examinations. *The International Journal of Humanities & Social Studies*. *3*(6), 201-207.

Fenrich, P. (2004)**.** Instructional design tips for virtually teaching practical skills. In *Proceedings of the 2004 Informing Science and IT Education Joint Conference. Rockhampton, Australia June 25–28, 2004.*

Frisbie, D.A. (1992). The multiple true-false format: A status review. *Educational Measurement: Issues and Practice, 11*(4), 21-26

Grunert, M. L., Raker, J. R., Murphy, K. L., & Holme, T. A. (2013). Polytomous versus dichotomous scoring on multiple-choice examinations: development of a rubric for rating partial credit. *Journal of Chemical Education*, *90*(10), 1310-1315.

Haladyna, T.M., Downing, S.M., & Rodriguez, M.C. (2002). A review of multiple choice item-writing guidelines for classroom assessment. *Applied Measurement in Education, 15* (3), 309-334.

Hambleton, R.K., Swaminathan, H., & Rogers, H.J. (1991). *Fundamentals of item response theory.* London: Sage Publications.

Hickson, S., Reed, W. R., & Sander, N. (2012). Estimating the effect on grades of using multiple-choice versus constructive-response questions: Data from the classroom. *Educational Assessment*, *17*(4), 200-213.

Holt, A. (2006). An analysis of negative marking in multiple-choice assessment. *The 19th Annual Conference of the National Advisory Committee on Computing Qualifications* (NACCQ 2006), Wellington, New Zealand. Samuel Mann and Noel Bridgeman (Eds)

Jiao, H., Liu, J., Hainie, K., Woo, A., & Gorham, J. (2012). Comparison between dichotomous and polytomous scoring of innovative items in a large-scale computerized adaptive test. *Educational and Psychological Measurement*, *72*(3), 493–509.

King, K. V., Gardner, D. A., Sasha Zucker, S., & Jorgensen, M. A. (2004). The distractor rationale taxonomy: Enhancing multiple-choice items in reading and mathematics. *Assessment Report*, Pearson Education, Inc, July 2004.

Klein, P., Müller, A., & Kuhn, J. (2017). Assessment of representational competence in kinematics. *Physical Review Physics Education Research 13*, 1-18.

Lin, C.J. (2008). Comparisons between classical test theory and item response theory in automated assembly of parallel test form. *The Journal of Technology, Learning, and Assessment. 6*(8), 1-42.

Martin, D. L., & Itter, D. (2014).Valuing assessment in teacher education-multiple-choice competency testing. *Australian Journal of Teacher Education*, 39(7), 1-14.

Merrell, J. D., Cirillo, P. F., Schwartz, P. M., & Jeffrey A. Webb, J. A. (2015). Multiple-choice testing using immediate feedback—assessment technique (IF AT) forms: Second-chance guessing vs. second-chance learning. *Higher Education Studies*. *5*(5), 50-55.

Oosterhof, A. (2003). *Developing and using classroom assessments* (*3th ed*). Upper Saddle River: Merrill Prentice Hall.

Redish, E.F., Scherr, R.E. & Tuminaro, J. (2006). Reverse engineering the solution of a "simple" physics problem: Why learning physics is harder than it looks. *The Physics Teacher*, 44(5), 293-300.

Rodriguez, M.C. (2005). Three options are optimal for multiple-choice items: A meta-analysis of 80 years of research. *Educational Measurement: Issues and Practice*, Summer, 3-13.

Stankous, N. V. (2016). Constructive response vs. multiple-choice tests in math: American experience and discussion (Review). *European Scientific Journal.* May 2016, 308-3016.

Tognolini, J., & Davidson, M. (Juli 2003). *How do we operationalise what we value? Some technical chalenges in assessing higher order thinking skills.* Paper presented in the Natinaonal Roundtable on Assessment Conference. Darwin, Australia.

Wiseman, C. S. (2012). A comparison of the performance of analytic vs. holistic scoring rubrics to assess L2 writing, *Iranian Journal of Language Testing*, *2* (1), 59-92.

Wooten, M. M., Cool, A. M., Edward E. Prather, E. E., & Tanner, K. D. (2014). Comparison of performance on multiple-choice questions and open-ended questions in an introductory astronomy laboratory. *Physical Review Special Topics - Physics Education Research*, 10, 1-22.

Wright, B.D., & Masters, G.N. (1982). *Rating scale analysis.* Chicago: Mesa Press.

Wu, B.C. (2003). Scoring multiple true-false items: A comparison of sumed scores and response pattern scores at item and test level. *Research report*, Educational Resources International Center (ERIC) USA, 1-40.

Submission Checklist

Makale Sunumu Kontrol Çizelgesi

Indicate that this submission is ready to be considered by this journal by checking off the following.

Aday makalenin değerlendirilmeye hazır olduğunu aşağıdakilerin her biri ile karşılaştırarak kontrol ediniz.

Manuscript Agreement:		
I/We hereby accept that, the article after being accepted for publication in the Eurasian Journal of Educational Research (EJER), the author (s) as, all rights related to the article has been transferred to the ANI PUBLISHING under the laws of the, "copyright transfer".		
Makale Eurasian Journal of Educational Research (EJER) dergisinde basıma kabul edildikten sonra, yazar(lar) olarak; makale ile ilgili tüm hakları, "Telif Hakkı Devir" yasaları uyarınca, ANI YAYINCILIK'A devrettiğimizi kabul ediyoruz.		
Manuscript Submission:		
I have prepared my manuscript based on the criteria listed below and I accept all submission conditions.		
Makalemi aşağıda belirtilen kriterlere göre hazırladım ve makale gönderme koşullarının tamamını kabul ediyorum.		
Indicate that this submission is ready to be considered by this journal by checking off the following.		
Aday makalenin değerlendirilmeye hazır olduğunu aşağıdakilerin her biri ile karşılaştırarak kontrol ediniz.		
1	☒	*The manuscript is a report of original educational research or a discussion article on research topics in education.* Aday makale, eğitim alanıyla doğrudan ilgili bir araştırma ya da tartışma yazısıdır.
2	☒	*The submission has not been previously published, nor sent to another journal for consideration.* Sunulan çalışma daha önce herhangi bir yerde yayınlanmamıştır, başka bir derginin incelemesinde değildir.
3	☒	*Within a year, I have not submitted a manuscript to EJER as an author or co-author for review other than this manuscript.*

		Son bir yıl içerisinde, yazarı olduğum ya da yazarları arasında bulunduğum başka bir çalışma değerlendirilmek üzere EJER'e sunulmamıştır.
4	☒	*All authors of this manuscript are subscribers of the EJER and accept to pay 3**00 EURO** for the cost of proofreading and 10 loose copies of their articles if the manuscript is accepted.* Aday makalenin kabul edilip basılması halinde makalelerinin İngilizce son okuma (proofreading) ve 10 serbest kopya gönderimi nedeniyle oluşan **1275 TL** maliyeti Anı Yayıncılık'a ödemeyi kabul ederler.
5	☒	*The entire manuscript is written in English.* Aday makalenin bütününün yazım dili İngilizce'dir.
6	☒	*The original manuscript is typed on A4 paper. The margins are 2.5 cm.* Aday makale kenar boşlukları 2.5 cm olan A4 kağıda yazılmıştır.
7	☒	*Each paragraph is longer than two sentences.* Her bir paragraf en az üç cümle içermektedir.
8	☒	*The entire manuscript - including quotations, references, author note, content footnotes, figure captions, and all parts of tables – is double-spaced.* Aday makalenin tamamı, alıntılar, kaynakça, şekil ve tablo başlıkları da dâhil olmak üzere çift aralıklı yazılmıştır.
9	☒	*The submission file is in Microsoft Word document file format. 12-point Times New Roman font is used in entire manuscript.* Aday makale, tamamında 12 punto Times New Roman yazı tipi kullanılarak hazırlanmış bir Microsoft Word dokümanıdır.
10	☒	*The text has had the authors' names removed. If an author is cited, "Author" and year are used in the bibliography and footnotes, instead of author's name, paper title, etc. The author's name has also been removed from the attached document.* **Aday makale, yazar adları çıkarılarak sunulmuştur**. Eğer yazar kendisine atıfta bulunduysa yazarın adına ve çalışma başlığına yer verilmeyecek, sadece "Author" yazılarak çalışmanın yılı belirtilecektir. Eklenen dosyada yazar adı belirtilmeyecektir.
11	☒	*The title is 10 to 12 words.* Aday makalenin başlığı 10-12 sözcük uzunluğundadır.
12	☒	*The maximum length of the manuscript-including structured abstract in English, tables, and references is 6000 words. This limitation does not include Turkish extended abstract (750-1000 words) which is placed after the references section.* Aday makale, İngilizce abstract, tablolar ve kaynakça vb. tüm ögeler dâhil olmak üzere en fazla 6000 sözcüktür. Kaynakça'nın ardından yer verilen uzun Türkçe özet (750-1000 sözcük) bu sayıya dâhil değildir.

13	☒	*The article is preceded by **English Structured Abstract of not more than 250 words** and not less than 200 using five required headings:* **Purpose***: State the problem in field. Then explain the purpose of the study.* **Method:** *Specify the research design, sample, and research instrument and data analysis in brief.* **Findings***: Highlight the significant, interesting or surprising results.* **Implications for Research and Practice***.* (These headings may need some adaptation in the case of discussion papers: *Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions*). More information available from http://www.tandf.co.uk/journals/authors/rereabstracts.asp Yapılandırılmış İngilizce öz 200-250 sözcük uzunluğunda olup, aday makalenin başında yer almakta ve **Purpose** *(İlk önce alanda karşılaşılan sorunu belirtelim. Daha sonra araştırmanın amacını bir cümle ile veriniz),* **Method** *(Araştırma deseni, örneklem, very taoplama aracı ve verilerin analizinini kısaca açıklayınız),* **Findings** *(En önemli ve çarpıcı araştırma bulgularını verelim)* **Implications for Research and Practice,** *(Uygulama ve ileriye dönük araştırmalar için olası çıkarımlarınız)* başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: *Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar* şeklinde olabilir. Daha fazla bilgi için; http://www.tandf.co.uk/journals/authors/rereabstracts.asp adresine başvurunuz.
14	☒	*Following the structured abstract in English four to six keywords are included. They should **represent** the content of your manuscript and **be specific** to your field or sub-field. Avoid using keywords form the title of the paper.* Yapılandırılmış İngilizce özden sonra 4-6 anahtar sözcüğe yer verilmiştir. Anahtar kelimeler çalışmanızı temsil etmeli ve kendi alanınıza ya da alt alanlara özgü olmalıdır. **Makale adındaki kavramları anahtar kelime olarak seçmekten <u>kaçınınız</u>**.
15	☒	*An extended (750-1000 words) Turkish structured abstract is placed following the "References" section using five required headings: Problem Statement, Purpose of Study, Methods, Findings and Results, and Conclusions and Recommendations.* (These headings may need some adaptation in the case of discussion papers: *Background, Purpose of Study, Sources of Evidence, Main Argument, and Conclusions*). More information available from http://www.tandf.co.uk/journals/authors/rereabstracts.asp Kaynakça'dan sonra 750-1000 sözcükten oluşan Türkçe yapılandırılmış öze yer verilmiştir. Türkçe yapılandırılmış öz *Problem Durumu, Araştırmanın Amacı, Araştırmanın Yöntemi, Araştırmanın Bulguları, Araştırmanın Sonuçları ve Önerileri* başlıklarını içermektedir. Bu başlıklar tartışma yazıları için: *Çalışmanın Temeli, Çalışmanın Amacı, Kanıt Kaynakları, Ana Tartışma ve Sonuçlar* şeklinde olabilir. Daha fazla bilgi için; http://www.tandf.co.uk/journals/authors /rereabstracts.asp

16	☒	*Following the Turkish structured abstract, four to six keywords are included.* Uzun Türkçe özetten sonra 4-6 anahtar sözcüğe yer verilmelidir.
17	☒	*References are not cited in the structured abstracts in English and in Turkish.* İngilizce abstract ve Türkçe öz içerisinde atıfta bulunulmamıştır.
18	☒	*The format of headings, tables, figures, citations, references, and other details follow the **APA 6 style** as described in the Publication Manual of the American Psychological Association, 6th edition, available from http://www.apa.org* Aday makalenin başlıkları, tabloları, şekilleri, atıfları, kaynakçası ve diğer özellikleri tamamen APA altıncı baskıda belirtildiği şekildedir.
19	☒	*All illustrations, figures, and tables are placed within the text at the appropriate points, rather than at the end.* Aday makalenin şekilleri ve tabloları metin içerisinde bulunmaları gereken uygun yerlere yerleştirilmiştir. Makale sonunda sunulmamıştır.
20	☒	Citations in the text of the document include the author's surname, the year of publication, and, when there is a specific quote from a source used, a page number where the quote is located in the text. **Example:** Nothing seemed so certain as the results of the early studies (Tatt, 2001, p. 445). It was precisely this level of apparent certainty, however, which led to a number of subsequent challenges to the techniques used to process the data (Jones & Wayne, 2002, p. 879). There were a number of fairly obvious flaws in the data: consistencies and regularities that seemed most irregular, upon close scrutiny (Aarns, 2003; West, 2003, p. 457). With studies by two authors, always include both author names: (Anderson & Bjorn, 2003) As Anderson and Bjorn (2003) illustrated in their recent study As recently as 2003, a prominent study (Anderson & Bjorn) illustrated When a study has 3, 4, or 5 authors, include the names of all the authors the first time the work is cited: (Anderson, Myers, Wilkes, & Matthews, 2003) For all subsequent citations of this work, use "et al.": (Anderson et al., 2003) When a work has 6 or more authors, use et al.: (Bell et al., 2003) For unsigned works, include the title, enclosed in parentheses. Put quotation marks for short work titles, and italicize the titles of reports, books, and other significant works: ("Recent Developments," 2004) (Dictionary of Tetrathalocigistic Diseases, 2004) Metin içindeki atıfları üstte verilen örneklere uygundur.

21	☒	Three levels of headings are used: Level 1, Level 3 and Level 4. The headings are formatted as follows: Centered Uppercase and Lowercase Heading (Level 1)
		Flush Left, Italicized, Uppercase and Lowercase Side Heading (Level 3)
		Indented, italicized, lowercase paragraph heading ending with a period. Start writing after the period (Level 4).
		Aday makale içerisinde üç farklı düzey başlık kullanılmıştır. Düzey 1, Düzey 2, Düzey 3. Başlıklar bu düzeylere uygun olarak aşağıdaki şekilde biçimlendirilmiştir:
		Ortalı ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık (Düzey 1)
		Tam Sola Dayalı, İtalik ve Her Sözcüğün İlk Harfi Büyük Yazılmış Başlık (Düzey 3)
		İçeriden, italik, tamamı küçük harflerle yazılmış ve nokta ile biten başlık.
		Noktadan sonra normal metin yazımına devam edilmeli (Düzey 4).
22	☒	References are listed in alphabetical order. Each listed reference is cited in text, and each text citation is listed in the References. Basic formats are as follows:
		Haag, L., & Stern, E. (2003). In search of the benefits of learning Latin. *Journal of Educational Psychology, 95,* 174–178.
		Bollen, K. A. (1989). *Structural equations with latent variables.* New York: Wiley.
		Johnson, D. W., & Johnson, R. T. (1990). Cooperative learning and achievement. In S. Sharan (Ed.), *Cooperative learning: Theory and research* (pp. 173–202). New York: Praeger.
		Turkish References Only:
		Çınkır, Ş., & Çetin, S. K. (2010). Öğretmenlerin okullarda mesleki çalışma ilişkileri hakkındaki görüşleri [Teachers' opinions about the professional working relationships in schools]. *Kuram ve Uygulamada Eğitim Yönetimi, 1 6*(3), 353-371.
		Article in an Internet-only journal/Periodical, database
		Fredrickson, B. L. (2000, March 7). Cultivating positive emotions to optimize health and well being. *Prevention & Treatment, 3,* Article 0001a. Retrieved November 20, 2000, from
		http://journals.apa.org/prevention/volume3/pre0030001a.html
		More information is available from:
		http://citationonline.net/CitationHelp/csg04-manuscripts-apa.htm#references

		Kaynakça'nın yazımı üstte verilen örneklere uygundur.
23	☒	***Order of the main parts in the manuscript is as follows:*** *Main title in English (max. 12 words)* *Structured abstract (min. 200- max.250 words length)* *Keywords (in English, min. four-max. six)* *Main text* *References* *Main Title in Turkish (max. 12 words* *Extended structured abstract (min.750-max.1000 words length in Turkish)* *Keywords (in Turkish, min. four-max. six)* Aday makaleyi oluşturan ana öğeler aşağıdaki düzendedir: İngilizce Ana Başlık (En fazla 12 sözcük) Yapılandırılmış İngilizce Abstract (En az 200, en fazla 250 sözcük) Anahtar Sözcükler (İngilizce, en az dört, en fazla altı) Ana Metin Kaynakça Türkçe Ana Başlık (En fazla 12 sözcük) Yapılandırılmış Türkçe Öz (En az 750, en fazla 1000 sözcük) Anahtar Sözcükler (Türkçe, en az dört, en fazla altı)

24	☒	**Structure of the Manuscript should be as follows:**
		
		Manuscript Title
		
		English Abstract (200-250 words) Sub-headings:
		
		  Purpose:…
		
		  Purpose of Study:…
		
		  Method:….
		
		  Findings:….
		
		  Implications for Research and Parctice:…
		
		  Keywords:...
		
		Introduction Method (sub-headings):
		
		  Research Design
		
		  Research Sample

		♦Research Instrument and Procedure
		Data Analysis
		Results, Discussion and Conclusion, References
		Extended Turkish Abstract (750-1000 words) Sub-headings:
		Problem Durumu:…
		Araştırmanın Amacı:…
		Araştırmanın Yöntemi:…
		Araştırmanın Bulguları:…
		Araştırmanın Sonuçları ve Önerileri:…
25	☒	***EJER editor reserves the right to reject manuscripts and not to publish it that do not comply fully with the EJER Manuscript Submission Guidelines.***

♦ Reliability and the validity of the research instrument used or adapted in the work must be provided, and explained in detail.

SUBSCRIPTION/ABONELİK

Annually / Yıllık

Individual / Kişi Aboneliği .... : *80 TL*

Institution / Kurum Aboneliği .... : *250 TL*

Abroad / Yurtdışı Aboneliği: .... : *250 $*

Student / Öğrenci Aboneliği: .... : *70 TL*

For single issue / Tek sayılar için

Individual / Kişiler .... : *25 TL*

Institution / Kurumlar .... : *75 TL*

Payment / Ödeme

İş Bankası Kızılay Şubesi, Anı Yayıncılık 4214-0962561

Account number/Numaralı hesap