Eurasian Journal of Educational Research 75 (2018) 37-58

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

As a Potential Source of Error, Measuring the Tendency of University Students to Copy the Answers: A Scale Development Study

Ergul DEMIR [1]

ARTICLE INFO

Article History:
Received: 07 Jan. 2018
Received in revised form: 22 Mar. 2018
Accepted: 28 Apr. 2018
DOI: 10.14689/ejer.2018.75.3

Keywords
answer-copying
dishonesty
scale development
copy-detection.

ABSTRACT

Purpose: The answer-copying tendency has the potential to detect suspicious answer patterns for prior distributions of statistical detection techniques. The aim of this study is to develop a valid and reliable measurement tool as a scale in order to observe the tendency of university students' copying of answers. Also, it is aimed to provide evidence with more comprehensive validity and reliability studies than the previously available researches. **Research Methods**: This is a scale development study. The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed under the "Classical Test Theory". Other theories were also considered, especially "Item Response Theory". After preliminary studies and item writing, a trial application with 711 students and main applications with 909 students was conducted. Structural validity, item and test descriptive statistics, item discriminations, inconsistency and test–retest reliability, classification accuracy, and item bias with differential item functioning were examined. **Findings**: The ACTS composed of 2 factor and 20 items. Total scores and item scores distributions are normal. Item discriminations are very high and over 0.40. α inconsistency coefficients are over 0.88 and test–retest reliability coefficient is 0.804. It provides highly correct classifications according to the students' answer-copying positions. There is no significant and serious DIF on items. **Implications for Research and Practice**: Unlike similar examples, it was studied on the large groups and used more comprehensive techniques to obtain evidences. Results show that the validity and reliability levels of the ACTS are very high. The ACTS can be used to understand the nature of the answer-copying. Also, and more importantly, it is thought that the ACTS can be used to detect suspicious answer patterns for prior distributions.

[1] Ankara University, Educational Science Faculty, TURKEY, e-mail:erguldemir@ankara.edu.tr, ORCID: https://orcid.org/0000-0002-3708-8013

Introduction

Academic integrity, dishonesty, cheating, or, in a more limited use, answer-copying has become an increasingly important problem in schools. Academic integrity is a more comprehensive definition including answer-copying, having proxies, plagiarism, academic misconduct, falsifying, etc. (Mullens, 2000). Some research shows that answer-copying has been observed at very serious proportions (Burke, 1997; Genereux & McLeod, 1995; May & Loyd, 1993; McCabe, 1993). Dwyer and Hecht (1994) stated that such dishonest behaviors by students have been increasing with increasing class sizes and reduced instructional resources. Similarly, McCabe and Trevino (1996) reported that the percentage of cheating behaviors among college students rose slightly. The inability to take control of this tendency leads to the injured reputation of institutions and hurts honest students, their families, and all of society (Aaron & Georgia, 1994).

Statistically, answer-copying is a source of systematic error and bias on items or tests. It may lead to unfair results. Unfortunately, it is very hard to detect answer-copying because initial response patterns are not easy to determine accurately for the statistical models. Improving the statistical techniques to detect answer-copying has been studied since the 1970s (Angoff, 1974; Frary, Tideman, & Watts, 1977; Holland, 1996; Maeda & Zhang, 2016; van der Linden & Sotaridona, 2006; Wesolowsky, 2000; Wollack, 1997). According to Dwyer and Hact (1994), some probabilistic techniques to detect cheaters have been used in American higher education since the 1920s. Generally, these techniques depend on matching responses between the copier and the source with a complex standardization process. Thus, the copiers and the sources have to be defined accurately. Also, there are some other challenges to using these techniques. As stated by van der Linden & Sotaridona (2006), these common techniques are limited by the null distribution of the set of items on which the statistics is defined. It is based on population-based statistics, and it is possible to arrive at unfair results. Alternatively, performances of IRT-based techniques also depend on the accurate estimation of copier ability, and it is obvious that the responses of copiers may be contaminated. For these techniques, especially, suspicious answer patterns should be defined accurately for prior and posterior distributions. It is understood that although there are some statistical techniques to detect answer-copying, these are limited or show doubt, a potential source of error, or just a probability.

Another way to understand the nature of copying is to consider related factors (Gerdeman, 2000; Hughes & McCabe, 2006). Some research reported that dishonest behaviors, mostly including answer-copying, were related to individual characteristics like GPA, age, gender, etc. (Crown & Spiller, 1998; McCabe & Trevino, 1997; Selçuk, 1995; Whitley, 1998). Students with lower GPAs, younger students, and males are most likely to cheat. Also, these groups show more tolerance for cheating behaviors.

Although they are less likely to engage in dishonesty or copying, dishonest behaviors seem to be related to particular educators and institutional policies (Aaron, 1992; Genereux & McLeod, 1995; McCabe & Trevino, 1996). Other important factors are attitudes, perceptions, and tendencies toward copying (Genereux & McLeod, 1995; Hughes & McCabe, 2006; McCabe & Trevino, 1997; Öztürk & Yeşilyaprak, 1997). The students with high work ethic, self-esteem, and lower test anxiety are less likely to cheat. On the other hand, the prevalence of cheating and the perception of cheating as acceptable increases cheating behaviors.

In many studies, questionnaires or self-reporting have been used to observe students' perceptions or tendencies on dishonest behavior such as copying (Bolin, 2004; McCabe & Travino, 1997; Selçuk, 1995). There are a few measurement tools available developed in these contexts (Eminoğlu & Nartgün, 2009; Gardner & Miller, 1988; LaGrange, 1992). Unfortunately, these tools are very limited in use and mostly unavailable. Generally, these tools were developed on a small and limited group and have comprehensive context. Most of them provide less proof about validity and reliability. Most of them were developed a long time ago. So, although cheating has been studied for a long time, it is obvious that valid and reliable tools are still needed.

Research Objectives

The aim of this study is to develop a valid and reliable measurement tool as a scale in order to observe the tendency of university students to copy answers. Also, it is aimed to provide evidence with comprehensive validity and reliability studies. With this aim, psychometric studies have been carried out for the "Answer-Copying Tendency Scale in University Students (ACTS)": (1) Structural validity, (2) item and test descriptive statistics, (3) item discriminations, (4) inconsistency and test–retest reliability, (5) classification accuracy, and (6) item bias with differential item functioning.

Statistically, the detection of answer-copying is a challenge. On the other hand, research shows that attitudes, perceptions, and tendencies are related to answer-copying behavior. In order to understand the nature of answer-copying and then detect it, we need a valid, reliable, and, most importantly, available tool to be used in this context. Also, the answer-copying tendency has the potential to make suspicious answer patterns detectable for prior/initial distributions of statistical detection techniques.

Method

This is a scale development study. The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed under the "True Score Theory" or "Classical Test Theory". In this process, other theories were also considered, especially

"Item Response Theory". The scale development steps were followed by considering DeVellis (2003).

Preliminary Studies and Item Writing

Developing the ACTS began with the preliminary studies. First, the trait and the aim were defined. The aim of the scale was defined as to measure the tendency of university students to copy answers. After that, the observable behaviors of students' copying tendencies was tried to be linked with the related literature. Simultaneously, around 80 university students were asked to write an essay on their perceptions and views about answer-copying. These documents were analyzed, and a total of 123 draft expressions were prepared for review by experts. The prepared draft form was sent to six specialists and academics working in the field of educational sciences. According to their opinions, 35 items were cancelled, 14 items were reorganized, and 79 items were accepted as they were. Of the 93 items, there were 34 items with negative direction. At the same time, the experts were asked about which type of scale would be used most effectively. Some experts recommended the Likert-type scale with five categories, while others recommended more than five categories. In response, and also by considering the related literature, it was decided that two differently numbered scoring categories would be used in the trial application. The first would be the classic Likert-type scale with five categories, and the second would be to score between 0 and 10 points. At the end of the preliminary studies, the trial application form was organized with 93 items.

Trial and Main Application

Trial application was executed with 711 undergraduate students from 16 universities and 18 faculties in May and June of 2017 in Turkey as a paper-pencil test. The main application was executed with 909 undergraduate students from 29 universities and 30 faculties in November and December of 2017 in Turkey as a web based application. Some demographic characteristics of the students whom participated in trial and main applications are given at Table 1.

Data Analysis

Defining the factor structure, exploratory and confirmatory studies were executed. "Principal component analysis" and "multidimensional scaling with ALSCAL method" was used with the trial application data. "Confirmatory factor analysis" was used with the main application data. As the test and item statistics, it was analyzed the distribution of the scores and calculated the descriptive statistics. As the item discrimination, item-total score correlations were calculated with Pearson' Product Moments correlation coefficient. For reliability, α inconsistency coefficient and test-retest reliability were calculated. For classification accuracy, discriminant function

analysis was used with four classification model. Finally, item bias was examined with the "lordif" techniques based on Item Response Theory. Before all analyses, main assumptions were checked in detail. The softwares SPSS 24.0 and Lisrel 8.7 and R 3.4.4 were used for these analyses.

Table 1

Students' Demographic Characteristics for Trial and Main Application Samples

Trial Application		*n*	%	*Main Application*		*n*	%
Gender	Female	525	73.8	Gender	Female	674	74.1
	Male	158	22.2		Male	235	25.9
	Missing	28	3.9		Total	909	100.0
	Total	711	100.0	University	Ankara University	325	35.8
University	Ankara University	447	62.9		Trakya University	115	12.7
	Hacettepe University	161	22.6		Gazi University	77	8.5
	TED University	29	4.1		Hacettepe University	69	7.6
	Konya Selçuk University	23	3.2		Other	321	35.3
	Other	22	2.9		Missing	2	0.2
	Missing	29	4.1		Total	909	100.0
	Total	711	100.0	Faculty	Education	271	29.8
Faculty	Educational Sciences	263	37.0		Educational Sciences	204	22.4
	Education	173	24.3		Science & Literature	145	16.0
	Theology	55	7.7		Language & History-Geography	88	9.7
	Science	47	6.6		Others	201	22.1
	Other	145	20.2		Total	909	100.0
	Missing	28	3.9	Class	Preparation and 1	212	23.3
	Total	711	100.0		2	235	25.9
Class	1	152	21.4		3	194	21.3
	2	206	29.0		4	218	24.0
	3	132	18.6		5 and Graduate	50	5.6
	4	158	22.2		Total	909	100.0
	Graduate	34	4.8				
	Missing	29	4.1				
	Total	711	100.0				

Results

Structural Validity: Exploratory and Confirmatory Studies

Exploratory studies for the ACTS were executed on the trial application data, and confirmatory analysis was executed on the main application data. As an exploratory analysis, "principal component analysis (PCA)" and "multidimensional scaling with ALSCAL model" were used respectively. PCA was executed with both the classic five categories of Likert-type scores and the scores between 0 and 10 separately. Before analysis, some items were recoded in order to equalize the way of all items. Missing values were checked, and there was no serious missing data problem. Nor were there any outliers (max. Mahalanobis distance < χ^2_{kritik}=25).

For the five-category Likert-type scale data, sampling is adequate (KMO=0.974), and multiple correlations among variables are statistically significant (for Bartlett's test of sphericity, χ^2=8895.179, df=190 and p<0.001). After the data reduction, 20 items with one factor could be identified as a structure. Communalities of each item are between 0.412 and 0.725. The eigenvalue of one factor is 11.463, and the total variance explained is 57%. Factor loadings are between 0.649 and 0.852.

Similarly, for scoring 0 to 10, sampling is adequate (KMO=0.959), and multiple correlations among variables are statistically significant (for Bartlett's test of sphericity, χ^2=6585.353, df=190 and p<0.001). After the data reduction with varimax and oblique rotation, 20 items with two factors could be identified as a structure. The communalities of each item are between 0.485 and 0.744. The eigenvalue of the first factor is 9.684 and second is 2.659. Total variance explained for the first factor is 48.42%, the second is 13.29%, and the total is 61.71%. The correlation between factors is statistically significant an, shows a negative and moderate relationship (r=-0.49 and p<0.05). These items and factor loadings are shown in the Table 2.

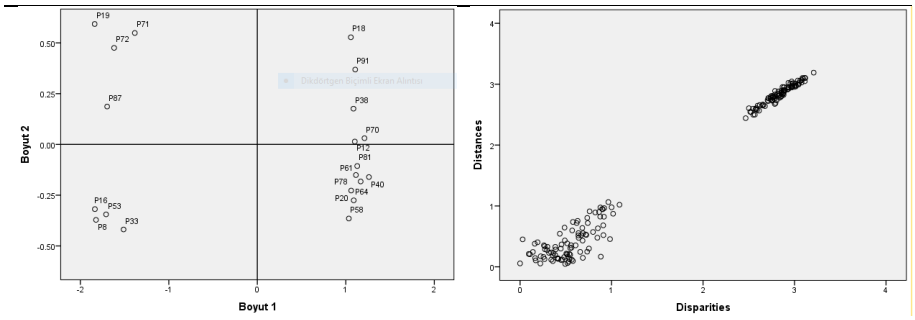
As seen at Table 2, 12 of 20 items are in the first factor and eight items are in the second factor. These eight items are the items with negative directions. The first factor was named "negative perception of exam and grade (NEGALGI)", and the second factor was named "ethical value (ETIK)". These two factors have negative correlations to each other (r=-0.49 and p<0.01). According to these, after standardization, students who have high ethical values also have lower levels of negative perception about exams and grades.

Table 2

Factor Loadings for ACTS Items

No.	Item	F1	F2
12	I'll copy answers when I get the chance.	0.791	
18	Having friends I know that get higher scores by copying answers, I also have the ambition to copy answers.	0.789	
20	I think about copying answers for the examinations that I have not prepared enough for.	0.836	
38	Copying answers can be enough to pass the exam.	0.801	
40	I feel compelled to copy answers myself in some lessons.	0.817	
58	It makes sense for me to copy answers in lessons that I will forget and not use in the future.	0.785	
61	Even if I do not do it, it gives me confidence to know that I can copy answers at the exam.	0.792	
64	It makes sense for me to copy answers to the questions I do not know.	0.821	
70	I will copy answers if I know I will not be punished.	0.808	
78	It makes more sense to copy answers when I memorize so much information that will not work in my own life.	0.752	
81	Anxiety about earning high grades pushes me to copy answers.	0.823	
91	If I am not afraid of getting caught, I'll copy answers.	0.811	
8	Copying answers makes a fool of a person.		-0.763
16	I see copying answers as an unfair advantage.		-0.761
19	I'm absolutely against copying answers.		-0.755
33	Copying answers is disrespectful to the teacher's endeavors.		-0.751
53	Copying answers is not my achievement but my deceit.		-0.744
71	Defending copying answers is completely nonsense.		-0.741
72	I would like to take real deterrent measures to prevent students from copying answers.		-0.694
87	I believe that the copying answers is immoral behavior.		-0.674
	Eigenvalue	9.684	2.659
	Total Variance Explained (%)	48.42	13.26

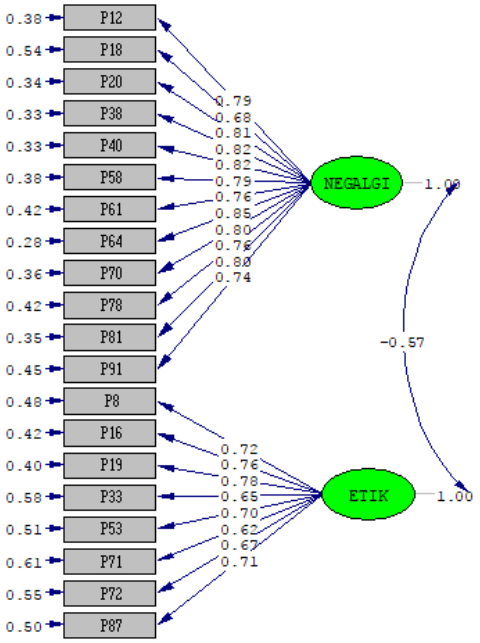
After PCA, the multidimensional scaling ALSCAL model provides visual and supportive evidence of the structure. According to the Euclidean distances, item locations are shown in Graphic 1.



Graphic 1. Item Distribution According to Multidimensional Scaling (ALSCAL) Euclidean Distances

As seen in Graphic 1, items are clearly separated by two factors. In the right-hand graphic, eight items are on one factor and twelve items are on the other factor. Furthermore, in the left-hand graphics, these two factors show a linear relationship. Also, outputs of MDA-ALSCAL show that the model provides good-fit with high variance accounted for (Stress=0.08623 and RSQ=0.98503).

After the exploratory studies, "confirmatory factor analysis (CFA)" was conducted on the main application data. Graphic 2 shows the standardized values for each path.



Graphic 2. Standardized Path Coefficients According to the Results of the CFA

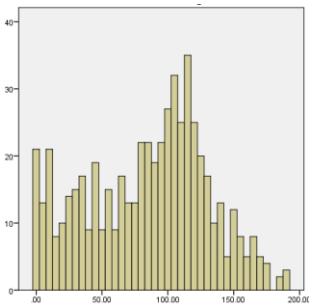
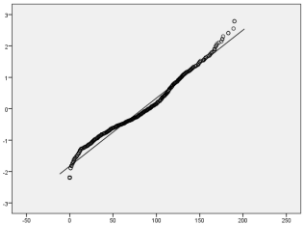
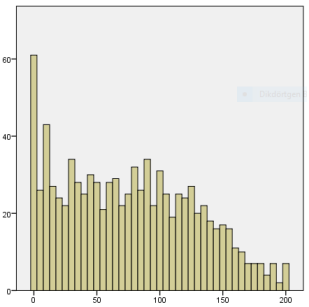
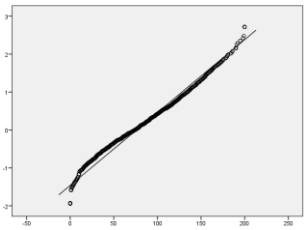
In Graphic 2, all paths are statistically significant (t_i>1.96 and p_i<0.05). Standardized solutions of the errors are under 0.90. Similar to the exploratory studies' results, there are moderate and negative correlations between the factors. Also, goodness of fit statistics show at least acceptable model-data fit (χ^2/sd=2.79, RMSEA=0.056, SRMR=0.036, GFI=0.92, NFI=0.98, CFI=0.99, ECVI Free model =41.45 and Saturated model=0.75, AIC Free model=2344.37 and Saturated Model=420.0).

Total Score Distributions and Descriptive Statistics

The ACTS has 20 items, which are each scored between 0 and 10. The available total scores are between 0 and 200. In Table 3, the total score distributions and descriptive statistics are given for both trial and main applications.

Table 3

ACTS Total Score Distributions and Descriptive Statistics

	Statistics	Value	Histogram	Normal Q-Q Plot
Trial Application	Mean	84.29		
	5% Trim.Mean	84.12		
	SE (Mean)	1.942		
	Median	91		
	Std.Dev	46.127		
	Minimum	0		
	Maximum	190		
	Skewness	-0.146		
	Kurtosis	-0.806		
Main Application	Mean	76.13		
	5% Trim.Mean	74.48		
	SE (Mean)	1.727		
	Median	74		
	Std.Dev.	52.054		
	Minimum	0		
	Maximum	200		
	Skewness	0.288		
	Kurtosis	-0.877		

As seen in Table 3, total scores distribute normally. The mean, trimmed mean, and median are close to each other. Skewness and kurtosis are between (-1, +1). Also, the graphics supporting the normality.

Item Statistics and Item-Total Score Correlations

The item statistics and item-total score correlations of the ACTS items were calculated for both trial and main application data. The results are given in Table 4.

Table 4

Item Statistics and Item-Total Score Correlations of the ACTS Items

Item	Trial Application						Main Application					
	Mean	Med.	Std.D.	Skew.	Kurt.	r	Mean	Med.	Std.D.	Skew.	Kurt.	r
12	4.49	5	3.380	0.144	-1.246	.736**	3.83	3	3.412	0.549	-1.021	.803**
18	4.08	4	3.447	0.320	-1.206	.613**	3.49	2	3.586	0.590	-1.110	.617**
20	4.69	5	3.422	-0.029	-1.329	.752**	4.13	4	3.491	0.315	-1.276	.779**
38	4.26	5	3.215	0.154	-1.129	.766**	3.82	3	3.501	0.440	-1.155	.816**
40	4.66	5	3.649	0.038	-1.435	.771**	3.78	3	3.666	0.488	-1.240	.760**
58	5.06	5	3.464	-0.076	-1.280	.760**	4.65	5	3.865	0.091	-1.559	.796**
61	4.76	5	3.563	0.013	-1.360	.721**	3.94	3	3.774	0.390	-1.378	.714**
64	4.90	5	3.405	-0.094	-1.253	.812**	3.98	4	3.543	0.368	-1.234	.844**
70	4.80	5	3.618	0.047	-1.374	.769**	4.50	4	3.816	0.216	-1.471	.812**
78	4.56	5	3.532	0.119	-1.319	.747**	3.67	3	3.607	0.517	-1.173	.779**
81	4.83	5	3.557	0.008	-1.347	.764**	4.52	4	3.759	0.170	-1.493	.806**
91	4.27	5	3.498	0.203	-1.287	.696**	3.87	3	3.604	0.466	-1.202	.819**
8*	6.69	7	3.210	-0.551	-0.951	.596**	7.27	8	3.093	-0.986	-0.113	.621**
16*	6.79	8	3.104	-0.621	-0.751	.608**	7.41	9	3.105	-1.028	-0.110	.682**
19*	5.32	5	3.545	-0.061	-1.328	.658**	5.77	6	3.538	-0.231	-1.306	.749**
33*	6.64	7	2.880	-0.551	-0.545	.529**	6.43	7	3.461	-0.559	-1.022	.569**
53*	6.58	7	3.120	-0.570	-0.749	.571**	6.80	8	3.395	-0.769	-0.714	.676**
71*	5.60	5	3.310	-0.156	-1.124	.488**	6.05	6	3.516	-0.355	-1.238	.646**
72*	5.77	6	3.303	-0.264	-1.085	.586**	5.83	6	3.626	-0.275	-1.332	.697**
87*	5.88	6	3.398	-0.245	-1.217	.576**	6.50	7	3.459	-0.578	-1.015	.678**

* Second factor items were recoded before the estimations.
**p<0.01

As seen in Table 4, trial and main application results are similar. Item means and medians are close to each other. Skewness values are between (-1, +1). Kurtosis values are mostly under -1, but between (-1.5, +1.5). Tabachnick and Fidel (2013, p.80) stated that non-strong skewness and kurtosis violations cannot lead the difference for the statistics in large samples. As a result, item scores show reasonable normality. Also, all item-total score correlations are significant at the level of 0.01. This is strong evidence of the discriminative validity of the items and the test.

Reliability Studies

For the reliability of the ACTS, first α inconsistency coefficients were calculated for both trial and main applications. If there is no violation of normality, α is an available estimation. The results are given in Table 5.

Table 5

a Coefficients for the ACTS

Factors	k	Trial Application		Main Application	
		n	α	n	α
Negative perception of exams and grades (NEGALGI)	12	711	0.950	909	0.955
Ethical value (ETIK)	8	711	0.884	909	0.907
Total	20	711	0.942	909	0.955

As seen in Table 5, the inconsistency of the ACTS is very high for both the factor level and the total test. Besides the inconsistency estimates, the test–retest reliability was considered. Test–retest application was conducted with 95 students in December of 2017 after two weeks from the main application. There was no normality violation, and the Pearson product-moment correlation coefficient was used to estimate. Results show that the ACTS has a high level of test–retest reliability ($r=0.804$ ve $p<0.001$).

Classification Accuracy

In the trial and main applications of the ACTS, participants were asked whether they copied answers, gave answers for others to copy, and/or witnessed answer-copying. Similar results were obtained from both the trial and main applications. So, just the main application data was considered for further analysis.

Table 6

Taking, Giving, and Witnessing the Answer-Copying among Undergraduates

	I did		I gave		I witnessed	
	n	%	n	%	n	%
Yes	444	48.8	360	39.6	126	13.9
No	465	51.2	549	60.4	783	86.1
Total	909	100.0	909	100.0	907	99.8

As seen in Table 6, most of the students stated that they did not copy answers (48.8%) or did not gave answers for others to copy (39.6%). On the other hand, the vast majority stated that they witnessed answer-copying (86.1%). The proportion of copying answers (60.4%) was more than giving answers for others to copy (51.2%). As validity evidence, it is expected that the ACTS can classify students according to their answer-copying positions. For this purpose, "discriminant function analysis (DFA)" was used. A total of four discriminant models were identified depending on the

answer-copying positions and ACTS factors. Before the analyses, the main assumptions were checked. As mentioned before, there are no violations about missignes, outliers, and normalities. Furthermore, the homogeneity of the variance-covariance matrices was checked by using Box's M statistics. These values are not significant at the level of 0.001, and there is no violation (Hair, Black, Babin, & Anderson, 2014, p.250). Descriptive statistics for each model are given in Table 7. And the test result for group differences and canonical discriminant functions are given in Table 8.

Table 7

Descriptive Statistics for Classification Models

Model	Predictor(s)	Grouping Variable		n	Mean	Std.Dev.
1[a]	ACTS_Total	Answer-Copying	Yes	444	52.80	48.132
			No	465	98.40	45.529
			Total	909	76.13	52.054
2[b]	ACTS_NEGALGI	Answer-Copying	Yes	444	31.28	32.178
			No	465	64.31	31.156
			Total	909	48.18	35.694
	ACTS _ETIK	Answer-Copying	Yes	444	58.48	20.291
			No	465	45.91	20.173
			Total	909	52.05	21.175
3[c]	ACTS _Total	Giving answer-copy	Yes	360	54.43	48.529
			No	549	90.35	49.345
			Total	909	76.13	52.054
4[d]	ACTS _NEGALGI	Giving answer-copy	Yes	360	33.09	33.052
			No	549	58.07	33.880
			Total	909	48.18	35.694
	ACTS _ETIK	Giving answer-copy	Yes	360	58.66	19.800
			No	549	47.72	20.941
			Total	909	52.05	21.175

[a]Box' M=1.402, F=1.400, df_1=1, df_2=2464420.891 and p=0.237
[b]Box' M=8.786, F=2.922, df_1=3, df_2=158752599.4 and p=0.033
[c]Box' M=0.120, F=0.120, df_1=1, df_2=2194294.214 and p=0.729
[d]Box' M=7.901, F=2.627, df_1=1, df_2=30459579.07 and p=0.049

Table 8

Test Results for Group Difference and Canonical Discriminant Functions

Model	Predictors	Equality of group means					Canonical Discriminant Functions					
		Wilks ′λ	F	df_1	df_2	p	Wilks ′λ	χ^2	df	p	Eigen value	Cann. Corr.
1	ACTS_Total	0.808	215.5	1	907	0.000	0.808	193.2	1	0.000	0.238	0.438
2	ACTS_NEGALGI	0.786	247.2	1	907	0.000	0.786	218.4	2	0.000	0.273	0.463
	ACTS_ETIK	0.912	87.7	1	907	0.000						
3	ACTS _Total	0.886	116.7	1	907	0.000	0.886	109.8	1	0.000	0.129	0.338
4	ACTS _NEGALGI	0.883	120.5	1	907	0.000	0.881	114.5	2	0.000	0.135	0.345
	ACTS _ETIK	0.936	62.0	1	907	0.000						

As seen in Table 8, each model shows significant group differences. Also, each model is statistically significant. After that, the maximum chance criteria were calculated according to the recommendation of Tabachnick and Fidel (2012, p.406) for unequal group sizes. The proportions of the correct classifications and related criteria are given in Table 9.

Table 9

Proportions of the Correct Classifications and Maximum Chance Criteria

Model	Correct Classification (%)	Maximum Chance (%)
1	69.9	50.0
2	71.9	50.0
3	68.2	52.2
4	68.4	52.2

As seen in Table 9, each model provides correct classifications beyond chance. For 1st and 2nd models, this ratio is higher. It is expected that the correct classification values should be higher, at least over 10% of maximum chance values (Hair et al.., 2014, p.261). According to this criteria, all models have high power for classifications. As a result, the ACTS can predict significantly both answer-copying and giving the answers to copy.

Differential Item Functioning Studies

DIF studies for the ACTS were conducted on the main application data. Gender sub-groups and faculties are considered the group variable. For gender, females were defined as the reference group (n=674, 74.1%), and males were defined as the focal group (n=235, 25.9%). For faculties, education faculties were defined as the reference group (n=440, 48.4%), and other faculties were defined as the focal group (n=469, 51.6%). "Logistic Ordinal Regression Differential Item Functioning using IRT (lordif)" (Choi, Gibbons, & Crane, 2011) was used for the estimations. This technique was developed for polythomous items and based on IRT with GPCM or GRM models. Both uniform and nonuniform DIF can be detected. DIF results for the ACTS are given in Table 10 and Table 11.

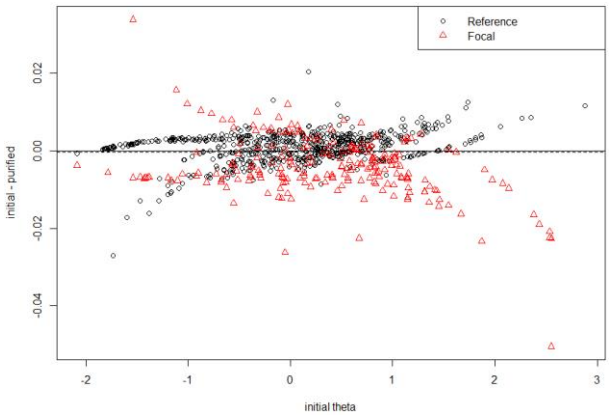
Table 10

*DIF Results for Gender Sub-Groups Obtained with lordif**

	Probability			Cox & Snell			Nagelkerke			McFadden			β_{12}
	χ^2_{12}	χ^2_{13}	χ^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	
1	0.8367	0.9478	0.7992	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0e+00	0.0013
2	0.0162	0.0239	0.1940	0.0017	0.0022	0.0005	0.0017	0.0022	0.0005	0.0014	0.0018	4e-04	0.0036
3	0.9611	0.9694	0.8069	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0003
4	0.9632	0.1598	0.0556	0.0000	0.0023	0.0023	0.0000	0.0024	0.0024	0.0000	0.0010	1e-03	0.0003
5	0.0106	0.0351	0.6859	0.0032	0.0033	0.0001	0.0033	0.0033	0.0001	0.0016	0.0017	0e+00	0.0201
6	0.0248	0.0800	0.9156	0.0017	0.0017	0.0000	0.0017	0.0017	0.0000	0.0012	0.0012	0e+00	0.0054
7	0.3088	0.4712	0.4933	0.0008	0.0011	0.0004	0.0008	0.0012	0.0004	0.0003	0.0004	1e-04	0.0071
8	0.6653	0.5713	0.3342	0.0001	0.0003	0.0003	0.0001	0.0003	0.0003	0.0000	0.0003	2e-04	0.0013
9	0.5654	0.7148	0.5593	0.0001	0.0003	0.0001	0.0001	0.0003	0.0001	0.0001	0.0002	1e-04	0.0024
10	0.0263	0.0225	0.1037	0.0030	0.0046	0.0016	0.0031	0.0047	0.0016	0.0013	0.0020	7e-04	0.0201
11	0.2684	0.3812	0.4016	0.0004	0.0007	0.0002	0.0004	0.0007	0.0002	0.0003	0.0005	2e-04	0.0033
12	0.6771	0.8728	0.7534	0.0001	0.0001	0.0000	0.0001	0.0001	0.0000	0.0000	0.0001	0e+00	0.0023
13	0.1974	0.3201	0.4324	0.0004	0.0005	0.0001	0.0004	0.0005	0.0001	0.0004	0.0006	2e-04	0.0065
14	0.1736	0.1418	0.1517	0.0006	0.0012	0.0006	0.0006	0.0012	0.0006	0.0005	0.0010	5e-04	0.0076
15	**0.0044**	**0.0067**	**0.1691**	**0.0052**	**0.0064**	**0.0012**	**0.0053**	**0.0065**	**0.0012**	**0.0020**	**0.0025**	**5e-04**	**0.0246**
16	0.6195	0.8237	0.7070	0.0002	0.0002	0.0001	0.0002	0.0002	0.0001	0.0001	0.0001	0e+00	0.0030
17	0.2253	0.4451	0.7001	0.0005	0.0006	0.0001	0.0006	0.0006	0.0001	0.0004	0.0004	0e+00	0.0048
18	0.5429	0.4023	0.2284	0.0001	0.0005	0.0004	0.0001	0.0005	0.0004	0.0001	0.0004	4e-04	0.0021
19	0.3798	0.5318	0.4832	0.0005	0.0008	0.0003	0.0005	0.0008	0.0003	0.0002	0.0003	1e-04	0.0065
20	0.0664	0.1539	0.5412	0.0009	0.0010	0.0001	0.0009	0.0010	0.0001	0.0008	0.0009	1e-04	0.0096

*Replication=100, α=0.01, ΔR^2 =0.02, Δβ=0.1

As seen in Table 10, except the 15th item, there is no significant DIF. The 15th item shows DIF with all χ^2 , R^2 , and β values. These differences are observed between the 1st-2nd and 1st-3rd models. So, it is possible that DIF should be uniform. The 15th item of the ACTS is "defending copying answers is completely nonsense". Distributions of females' and males' responses with their tendency levels (θ) are given in Graphic 3.



Graphic 3. Females' (Reference) and Males' (Focal) Responses with their Tendency Levels (θ) in the 15th Item of the ACTS

As seen in Graphic 3, at each θ level, females are located in the middle, whereas males have different locations. At the lower θ levels, males show more admittance to the 15th item. At the higher θ levels, the opposite is the case. This can show real differences between the gender sub-groups. Indeed, some research reported that males are more prone to cheating and copying (Crown & Spiller, 1998; Genereux & McLeod, 1995; Selçuk, 1995; Whitley, 1998). If there is a real difference in the 15th item, it can be stated that the ACTS has no DIF between gender sub-groups.

Table 11

*DIF Results for Faculty Sub-Groups Obtained with lordif**

	Probability			Cox & Snell			Nagelkerke			McFadden			β_{12}
	χ^2_{12}	χ^2_{13}	χ^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	R^2_{12}	R^2_{13}	R^2_{23}	
1	0.0998	0.2560	0.8968	0.0018	0.0019	0.0000	0.0019	0.0019	0.0000	7e-04	0.0008	0.0000	0.0029
2	0.6035	0.3700	0.1898	0.0001	0.0006	0.0005	0.0001	0.0006	0.0005	1e-04	0.0005	0.0004	0.0003
3	0.9914	0.4724	0.2207	0.0000	0.0009	0.0009	0.0000	0.0009	0.0009	0e+00	0.0004	0.0004	0.0000
4	0.9297	0.9890	0.9044	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0000	0.0000	0.0001
5	**0.5920**	**0.0083**	**0.0023**	**0.0001**	**0.0047**	**0.0046**	**0.0001**	**0.0048**	**0.0046**	**1e-04**	**0.0024**	**0.0023**	**0.0003**
6	0.7174	0.8093	0.5888	0.0000	0.0001	0.0001	0.0000	0.0001	0.0001	0e+00	0.0001	0.0001	0.0003
7	0.6390	0.2473	0.1086	0.0002	0.0021	0.0020	0.0002	0.0021	0.0020	1e-04	0.0007	0.0007	0.0003
8	0.8153	0.6698	0.3874	0.0000	0.0002	0.0002	0.0000	0.0002	0.0002	0e+00	0.0002	0.0002	0.0001
9	0.0867	0.1553	0.3741	0.0012	0.0015	0.0003	0.0012	0.0015	0.0003	8e-04	0.0010	0.0002	0.0027
10	0.1785	0.3856	0.7576	0.0011	0.0012	0.0001	0.0011	0.0012	0.0001	5e-04	0.0005	0.0000	0.0028
11	0.2993	0.4566	0.4837	0.0004	0.0005	0.0002	0.0004	0.0005	0.0002	3e-04	0.0004	0.0001	0.0009
12	0.2539	0.1787	0.1433	0.0006	0.0016	0.0010	0.0006	0.0017	0.0010	3e-04	0.0009	0.0006	0.0006
13	0.3633	0.5833	0.6160	0.0002	0.0002	0.0001	0.0002	0.0002	0.0001	2e-04	0.0003	0.0001	0.0006
14	0.7429	0.9080	0.7700	0.0000	0.0001	0.0000	0.0000	0.0001	0.0000	0e+00	0.0000	0.0000	0.0001
15	0.6603	0.9051	0.9373	0.0001	0.0001	0.0000	0.0001	0.0001	0.0000	0e+00	0.0000	0.0000	0.0000
16	0.6691	0.8854	0.8053	0.0001	0.0001	0.0000	0.0001	0.0002	0.0000	0e+00	0.0001	0.0000	0.0001
17	0.8401	0.3451	0.1485	0.0000	0.0008	0.0008	0.0000	0.0008	0.0008	0e+00	0.0005	0.0005	0.0000
18	0.7289	0.9405	0.9600	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0e+00	0.0000	0.0000	0.0002
19	0.1360	0.0623	0.0680	0.0014	0.0034	0.0021	0.0014	0.0035	0.0021	6e-04	0.0014	0.0009	0.0022
20	0.5212	0.3237	0.1744	0.0001	0.0006	0.0005	0.0001	0.0006	0.0005	1e-04	0.0006	0.0005	0.0004

*Replication=100, α=0.01, ΔR^2 =0.02, Δβ=0.1

As seen in Table 11, there is no DIF on any items except the 5th item. The 5th item shows DIF according to the χ^2 values. It is known that χ^2 is affected by sample size (Tabachnick & Fidel, 2013; Hair et al., 2014). On the other hand, the R^2 and β values are very small and close to 0. This shows that there are no significant DIF on the items of the ACTS.

Discussion and Conclusions

In this study, it was thought that the answer-copying tendency would be one of the indicators of the answer-copying behaviors. It is highly possible that the students with higher levels of tendency will be answer-copying. It is obvious that the statistical detection techniques couldn't provide exact solutions. These known techniques need initial response patterns for defining the suspicious focal group of answer-copying.

Therefore, we need such indirect solutions at least in order to understand the nature of answer-copying.

The "Answer-Copying Tendency Scale in University Students (ACTS)" was developed within this context. Unlike similar examples, it was studied on large groups and used more comprehensive techniques to obtain psychometric evidence. Results show that the validity and reliability levels of the ACTS are very high. The ACTS can be used to understand the nature of answer copying. Also, and more importantly, it is thought that the ACTS can be used to define suspicious answer patterns for prior distributions.

References

Aaron, R.M. (1992). Student academic dishonesty: Are collegiate institutions addressing the issue? *NASPA Journal, 29(2)*, 107-113.

Aaron, R.M., & Georgia, R.T. (1994). Administrator perceptions of student academic dishonesty in collegiate institutions. *NASPA Journal, 31(2),* 83-91.

Angoff, W.H. (1974). The development of statistical indices for detecting cheaters. *Journal of the American Statistical Association, 69,* 44–49.

Bolin, A.U. (2004). Self-control, perceived opportunity, and attitudes as predictors of academic dishonesty. *The Journal of Psychology, 138(2),* 101-114.

Burke, J. L. (1997). *Faculty perceptions of and attitudes toward academic dishonesty at a two year college.* (Unpublished doctoral dissertation). University of Georgia.

Choi, S.W., Gibbons, L.E., & Crane, P.K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software, 39*(8), 1-30.

Crown, D.F., & Spiller, M.S. (1998). Learning from the literature on collegiate cheating: A review of empirical research. *Journal of Business Ethics, 17*, 683-700.

DeVellis, R.F. (2003). *Scale development, theory and applications (Second edition).* Sage Publication, Inc.

Dwyer, D.J., & Hecht, J.B. (1994). *Cheating detection: Statistical, legal and policy implications*. Illinois State University.

Eminoglu, E., & Nartgun, Z. (2009). Universite ogrencilerinin akademik sahtekarlık egilimlerinin olculmesine yonelik bir olcek gelistirme calısması [A scale development study to measure academic dishonesty tendency of university students]. *Uluslararası Insan Bilimleri Dergisi, 6(1),* 215-240.

Frary, R.B., Tideman, T.N., &Watts, T.M. (1977). Indices of cheating on multiple-choice tests. *Journal of Educational Statistics, 6,* 152–165.

Gardner, W.M. & Kenneth, B.M. (1988). A scale for measuring attitude toward cheating. *Bulletin of the Psychonomic Society, 26(5),* 429-432.

Genereux, R.L., & McLeod, B.A. (1995). Circumstances surrounding cheating: A questionnaire study of college students. *Research in Higher Education, 36(6),* 687-704.

Gerdeman, R.D. (2000). *Academic dishonesty and the community college.* ERIC Digest, ED447840.

Hair Jr., J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2014). *Multivariate data analysis (seventh edition)*. Edinburg Gate, England: Pearson Education Limited.

Holland, P.W. (1996). *Assessing unusual agreement between the incorrect answers of two examinees using the K-index: Statistical theory and empirical support (Research Report RR 96-7).* Princeton, NJ: Educational Testing Service.

Hughes, J.M.C., & McCabe, D.L. (2006). Understanding Academic Misconduct. *Canadian Journal of Higher Education, 36(1)*, 49 – 63.

Maeda, H., & Zhang, B. (2016). An algorithm to improve test answer copying detection using the omega statistic. *International Journal of Testing, 17(1)*, 55-73.

May, K.M., & Loyd, B.H. (1993). Academic dishonesty: The honor system and students' attitudes. *Journal of College Student Development, 34*, 125-129.

McCabe, D.L. (1993). Academic integrity: What the latest research shows. *Synthesis: Law and Policy in Higher Education, 5(1),* 340-343.

McCabe, D.L., & Trevino, L.K. (1996). What we know about cheating in college: Longitudinal trends and recent developments. *Change, 28(1),* 28-33.

McCabe, D.L., & Trevino, L.K. (1997). Individual and contextual influences on academic dishonesty: A multi-campus investigation. *Research in Higher Education, 38*(3), 379-396.

Mullens, A. (2000). Cheating to win. *University Affairs, 41*(10), 22-28.

Ozturk, B. & Yesilyaprak, B. (1997). Universite ogrencilerinin kopya cekme olayına iliskin tutumları [Attitudes of the university students toward copying]. *Egitim Yonetimi, 3*(2), 227-240.

Selcuk, Z. (1995). Bir egitim ve rehberlik sorunu: Okullarda kopya cekme [An education and consultansy problem: Copying in schools]. *Kuram ve Uygulamada Egitim Yonetimi, 1*(3), 397-418.

Tabachnick, B.G., & Fidell, L.S. (2013). *Using multivariate statistics (sixth edition)*. New Jersey, USA: Pearson Education.

van der Linden, W.J., & Sotaridona, L.S. (2006). Detecting answer copying when the regular response process follows a known response model. *Journal of Educational and Behavioral Statistics, 31(3)*, 283–304.

Wesolowsky, G. (2000). Detecting excessive similarity in answers on multiple choice exams. *Journal of Applied Statistics, 27*(7), 909-921.

Whitley Jr, B.E. (1998). Factors associated with cheating among college students: A review. *Research in Higher Education, 39*(3), 235-274.

Wollack, J.A. (1997). A nominal response model approach to detect answer copying. *Applied Psychological Measurement, 21,* 307–320.

Appendix A.

Kopya Çekme Eğilimleri Ölçeği (KÇE)

[Answer-Copying Tendency Scale in University Students (ACTS)]

Aşağıdaki ifadelere katılma düzeylerinize 0 (hiç katılmıyorum) ile 10 (tamamen katılıyorum) arasında puan veriniz. [Point your participation levels below between 0 (I do not agree) and 10 (I fully agree).]

		Katılma Düzeyiniz [Participation Level]
1*	Kopya çekmek, insanın kendisini kandırmasıdır. [Copying answers makes a fool of a person.]	
2	Fırsatını yakaladığım durumlarda kopya çekerim. [I'll copy answers when I get the chance.]	
3*	Kopya çekmeyi haksız bir kazanç olarak görüyorum. [I see copying answers as an unfair advantage.]	
4	Kopya çekerek yüksek puanlar aldığını bildiğim arkadaşların olması bende de kopya çekme hırsı uyandırıyor. [Having friends I know that get higher scores by copying answers, I also have the ambition to copy answers.]	
5*	Kopya çekilmesine kesinlikle karşıyım. [I'm absolutely against copying answers.]	
6	Yeterince hazırlanamadığım sınavlarda kopya çekmeyi düşünürüm. [I think about copying answers for the examinations that I have not prepared enough for.]	
7*	Kopya çekmek, öğretmenin emeğine saygısızlıktır. [Copying answers is disrespectful to the teacher's endeavors.]	
8	Dersi geçmeye yetecek kadar kopya çekilebilir. [Copying answers can be enough to pass the exam.]	
9	Bazı derslerde kendimi kopya çekmeye mecbur hissediyorum. [I feel compelled to copy answers myself in some lessons.]	

10*	Kopya çekmem benim başarımı değil hilekârlığımı gösterir. [Copying answers is not my achievement but my deceit.]	
11	Zaten unutacağım ve ileride kullanmayacağım konuları içeren derslerde kopya çekmek bana mantıklı geliyor. [It makes sense for me to copy answers in lessons that I will forget and not use in the future.]	
12	Çekmesem bile sınavda kopya çekebilecek olduğumu bilmek bana güven verir. [Even if I do not do it, it gives me confidence to know that I can copy answers at the exam.]	
13	Bilmediğim konulardan gelen sorularda kopya çekmek bana makul geliyor. [It makes sense for me to copy answers to the questions I do not know.]	
14	Herhangi bir ceza almayacağımı bilsem kopya çekerim. [I will copy answers if I know I will not be punished.]	
15*	Kopya çekmenin savunulması tamamen saçmalıktır. [Defending copying answers is completely nonsense.]	
16*	Kopya çekilmesini önlemeye yönelik gerçekten caydırıcı önlemler alınmasını istiyorum. [I would like to take real deterrent measures to prevent students from copying answers.]	
17	İşime yaramayacak onca bilgiyi ezberleyeceğime kopya çekmek daha mantıklı geliyor. [It makes more sense to copy answers when I memorize so much information that will not work in my own life.]	
18	Not kaygısı beni kopya çekmeye itiyor. [Anxiety about earning high grades pushes me to copy answers.]	
19*	Kopya çekmenin ahlak dışı bir davranış olduğuna inanıyorum. [I believe that the copying answers is immoral behavior.]	
20	Yakalanma korkum olmasa kopya çekerim. [If I am not afraid of getting caught, I'll copy answers.]	

*İkinci faktöre (ETIK) ait maddelerdir. Ölçek toplam puanı için ters kodlanması gerekir. [Belongs to the second factor (ETIK). Should be recode for total score.]

Potansiyel Bir Hata Kaynağı Olarak Üniversite Öğrencilerinde Kopya Çekme Eğilimlerinin Ölçülmesi: Bir Ölçek Geliştirme Çalışması

Atıf:

Demir, E. (2018). As a potential source of error, measuring the tendency of university students to copy the answers: A scale development study. *Eurasian Journal of Educational Research, 75 (2018),* 37-58, DOI: 10.14689/ejer.2018.75.3

Özet

Problem Durumu: Kopya çekmeyi de içeren akademik sahtekarlık okullarda giderek daha önemli bir sorun haline gelmektedir. Bu sorun, sınıf büyüklüklerinin artırılması ve öğretim kaynaklarının azaltılması ile artmaktadır. Kopya çekmenin tespit edilmesi

için bazı istatistiksel teknikler olmasına rağmen, bunlar sınırlıdır ve ancak bir hata, olası bir hata kaynağı veya bir olasılık gösterebilmektedir. Kopya çekmenin yapısını ve doğasını anlamak için bir başka yol, tutumlar ya da eğilimler gibi ilişkili diğer faktörleri dikkate almaktır. Çalışma ahlakı yüksek olan, sınav kaygısı düşük olan, öz-benlik saygısı yüksek olan vb. öğrencilerin daha az kopya çektiği gözlenmektedir. Diğer taraftan kopya çekmenin yaygınlaşması, normal bir davranış olarak görülmesi ve kabul edilmesi durumlarında, kopya çekme davranışları daha sık görülmektedir. Pek çok çalışmada öğrencilerin sahtekarlık ya da kopya çekmeye yönelik görüş ve önerileri, ancak anketler aracılığıyla gözlenmiştir. Ölçek düzeyinde olan yani toplam puan alınabilen çok az sayıda araç vardır. Bu araçların kullanımı ve erişimi oldukça sınırlıdır. Ayrıca bu araçların genellikle, küçük ve sınırlı gruplarda geliştirildiği görülmektedir. Kapsamları geniştir. Çoğunluğu, geçerlik ve güvenirlik düzeylerine yönelik az sayıda kanıt içermektedir. Çoğunlukla bu araçlar geliştireli çok zaman geçmiştir. Anlaşıldığı üzere, uzun süredir konu üzerinde çalışılıyor olmasına ragmen, geçerli ve güvenilir araçlara olan ihtiyaç devam etmektedir.

Araştırmanın Amacı: Bu çalışmanın amacı, üniversite öğrencilerinin kopya çekme eğilimlerini gözlemlemek için ölçek olarak geçerli ve güvenilir bir ölçme aracı geliştirmektir. Ayrıca, daha kapsamlı geçerlilik ve güvenilirlik çalışmaları ile kanıt sağlamayı amaçlamaktadır.

Araştırmanın Yöntemi: Bu bir ölçek geliştirme çalışmasıdır. "Klasik Test Teorisi" altında "Üniversite Öğrencilerinde Kopya Çekme Eğilimi Ölçeği (KÇE)" geliştirilmiştir. Ayrıca diğer teoriler, özellikle "Madde Tepki Kuramı" da dikkate alınmıştır. Ön çalışmalar ve madde yazıma çalışmalarında, 80 civarı öğrenciden kopya çekmeye yönelik görüşlerini kompozisyon biçiminde yazmaları istenmiş, bu dokumanlar analiz edilerek 123 maddeden oluşan bir taslak form oluşturulmuştur. Bu form eğitim bilimleri alanında çalışan 6 akademisyenin görüşlerine sunulmuştur. Görüşler doğrultusunda gerekli düzenlemeler yapılarak 93 maddelik deneme formu oluşturulmuştur. Bu çalışmalardan sonra 711 öğrenci ile deneme uygulaması ve 909 öğrenci ile ana uygulamalar yapılmıştır. Elde edilen veriler kullanılarak, yapı geçerliği, madde ve test betimsel istatistikleri, madde ayırıcılıkları, içtutarlılık ve test-tekrar test güvenirliği, sınıflandırma doğruluğu ve değişen madde fonksiyonu ile madde yanlılığı incelenmiştir.

Araştırmanın Bulguları: Deneme ve esas uygulama verileri üzerinde ayrı ayrı yürütülen "Temel Bileşenler Analizi (TBA)" sonuçları birbirini destekler niteliktedir. Bu sonuçlara göre KÇE, 2 faktör ve 20 maddeden oluşmaktadır. İlk faktör 12 maddeden oluşmaktadır ve "negatif sınav ve not algısı (NEGALGI)" olarak tanımlanmıştır. İkinci faktör 8 maddeden oluşmaktadır ve "etik değerler (ETIK)" olarak tanımlanmıştır. Açıklanan varyans yüzdeleri, birinci faktör için %48.42, ikinci faktör için %13.29 ve toplamda %61.71'dir. Bu faktörler arasında negatif yönlü ve orta düzey bir ilişki

bulunmaktadır (r=-0.49 ve p<0.01). TBA'nın yanı sıra hem destekleyici kanıt sağlamak hem görsel sunum sağlamak için ALSCAL tekniği ile Öklit uzaklıklarına dayalı çok boyutlu ölçekleme çalışması da yapılmıştır. Elde edilen saçılma grafikleri de yapının belirgin bir şekilde 2 boyuttan oluştuğunu göstermiştir. KÇE'de her bir madde 0 ile 10 arasında puanlanmaktadır. Toplam puanlar 0 ile 200 arasında değişmektedir. Toplam puanlar ve madde puanlarına yönelik dağılımlar, normaldir. Ayrıca madde-toplam puan korelasyonlarına yönelik olarak madde ayırıcılıkları, çok yüksek ve 0,40'ın üzerinde gözlenmiştir. Güvenirlik çalışmaları olarak iki ayrı yöntem kullanımıştır. Öncelikle hem deneme uygulaması hem esas uygulama verileri üzerinde, hem test geneli hem faktörler düzeyinde α içtutarlılık katsayıları hesaplanmıştır. Bu değerler; deneme uygulaması sonuçlarına göre birinci faktör için 0.950, ikinci faktör için 0.884 ve ölçek geneli için 0.942; esas uygulama sonuçlarına göre ise sırasıyla 0.955, 0.907 ve 0.955'tir. Bu değerler çok yüksek güvenirlik düzeylerine işaret etmektedir. Ayrıca esas uygulamaya katılan 95 öğrenciye tekrar test uygulaması yapılarak test-tekrar test güvenirliği hesaplanmıştır. Bu değer de 0.804'tür ve yüksek düzeyde güvenirliğe işaret etmektedir. Bir diğer inceleme olarak KÇE'nin öğrencileri kopya çekme ve kopya verme durumlarına göre ne düzeyde doğru sınıflandırabildiği "Diskriminant Fonksiyon Analizi" ile analiz edilmiştir. Bu kapsamda test edilen dört ayrı modelde de yüksek sınıflama doğrlukları elde edilmiştir. Madde yanlılığı çalışmaları her bir madde düzeyinde değişen madde fonksiyonlaşması incelenerek yapılmıştır. Bu amaçla, çok kategorili maddelere yönelik olarak MTK'ya dayalı bir şekilde geliştirilmiş "lordif" tekniği kullanılmıştır. Analiz sonuçlarına göre maddeler üzerinde manidar ve ciddi düzeyde bir DIF gözlenmemiştir.

Araştırmanın Sonuçları ve Önerileri: Bu çalışmada ana hipotez, kopya çekme eğiliminin kopya çekme davranışlarının bir göstergesi olacağı yönündeydi. Daha yüksek düzeyde bir eğilime sahip olan öğrencilerin, kopya çekme davranışı göstermeleri olasılığı yüksektir. İstatistiksel tespit teknikleri kesin çözümler sağlayamadığından, kopya çekmenin doğasını anlamak için en azından bu tür dolaylı çözümlere ihtiyacımız var. Bu kapsamda bu çalışmada "Üniversite Öğrencilerinde Kopya Çekme Eğilim Ölçeği (KÇE)" geliştirilmiştir. Benzer örneklerden farklı olarak, bu çalışmada büyük gruplar üzerinde çalışılmış ve kanıt elde etmek için daha kapsamlı teknikler kullanılmıştır. Sonuçlar, KÇE'nin geçerlilik ve güvenilirlik düzeylerinin çok yüksek olduğunu göstermektedir. KÇE, kopya çekmenin doğasını anlamak için kullanılabilir. Ayrıca ve daha önemlisi, KÇE'nin kopya belirleme tekniklerinin öncül dağılımları için şüpheli cevap örüntülerini tespit etmek amacıyla da kullanılabileceği düşünülmektedir.

Anahtar Sözcükler: kopya çekme eğilimi, sahtekarlık, ölçek geliştirme, kopya-belirleme