OPEN ACCESS

Manuscript ID:
EDU-2025-13048999

Volume: 13

Issue: 4

Month: September

Year: 2025

P-ISSN: 2320-2653

E-ISSN: 2582-1334

Received: 29.04.2025

Accepted: 12.08.2025

Published Online: 01.09.2025

Citation:
Çetin, M., Asiltürk, E., Polat, M., & Önal Karakoyun, G. (2025). Examining the Relationship between 8th Grade Students' Academic Success Predictions on Acid-Base Subject, Dunning–Kruger Effect, Moral Maturity Level and Life Positions. *Shanlax International Journal of Education, 13*(4), 106-115.

DOI:
https://doi.org/10.34293/education.v13i4.8999

Examining the Relationship between 8th Grade Students' Academic Success Predictions on Acid-Base Subject, Dunning-Kruger Effect, Moral Maturity Level and Life Positions

Meryem Çetin
Fırat University, Turkey
https://orcid.org/0009-0007-2835-6825

Erol Asiltürk
Fırat University, Turkey
https://orcid.org/0000-0001-8126-7812

Mehmet Polat
Fırat University, Turkey
https://orcid.org/0000-0002-5955-5620

Gülen Önal Karakoyun
Van Yüzüncü Yıl University, Turkey
https://orcid.org/0000-0002-7675-0006

Abstract
This study examines the relationship between 8th grade students' academic success estimates on the acid-base subject, the Dunning–Kruger effect, moral maturity, and life positions. The quantitative research method was used, and the Life-Based Skills Test (LFBST), Moral Maturity Scale (MMS), and Life Positions Scale (LPS) served as data collection tools. The study was carried out with 610 students, but after data cleaning 459 valid responses were analyzed. Analyses showed no significant relationship between Dunning–Kruger Effect scores and gender. However, significant relationships were found with school type and grade level. While moral maturity and life positions were positively correlated, the Dunning–Kruger Effect showed no relation to either of these variables. Cluster analysis identified four self-assessment profiles: students who underestimated their performance, students whose estimates matched their performance, students who moderately overestimated, and students who largely overestimated. These profiles show how different patterns of miscalibration may affect self-perception and confidence. Overall, the findings reveal that students at lower achievement levels tend to overestimate their performance, while those at higher levels show underestimation. These results underline the importance of considering cognitive biases when supporting students' self-assessment skills.

Keywords: Academic Achievement, Dunning–Kruger Effect, Life Positions, Moral Maturity, Science Education

Introduction

An individual's ability to evaluate him- or herself realistically is a critical determinant of both quality of life and the nature of social interactions. This capacity requires an objective appraisal of one's strengths and limitations (Dunning et al., 2003). Perceptual distortions and cognitive biases, however, can complicate this process. Chief among these is the Dunning–Kruger Effect (DKE), which reveals a systematic misjudgment of competence: people with low skill levels tend to overestimate their performance, whereas highly skilled

individuals often underestimate theirs (Kruger & Dunning, 1999). The effect is supported by mechanisms such as metacognitive deficits and self-serving attribution biases, yielding important practical consequences in education and the workplace.

Perception is a subjective experience that emerges as the individual interprets sensory data (Plotnik, 2009). Past experiences, schemas, and attitudes play decisive roles in this process, potentially distorting reality (Yeniçeri & Dönmez, 2008). Negative experiences, for instance, may constrain future behaviour and adversely affect self-concept. Within this framework, moral maturity and life positions shape behavioural outcomes by influencing how individuals perceive themselves and their environment (İşgör et al., 2012). High moral maturity is associated with fair judgment and the capacity for empathy (Şengün & Kaya, 2007), whereas life positions determine one's role within social dynamics.

Review of Literature

The first studies on the Dunning Kruger effect were in the fields of logical reasoning, humor tests, and grammar (Kruger & Dunning, 1999). Later, studies aiming to examine the Dunning Kruger effect in different fields have appeared in the literature. Many other studies, including sports coaching (Sullivan et al., 2019), geography (Ehrlinger & Dunning, 2003), driving skill (Marottoli & Richardson, 1998), card games (Simons, 2013), and emotional intelligence ability-based tests (Sheldon et al., 2014), have emphasized the importance of the Dunning Kruger effect in various fields. As a result of these studies, it was concluded that low-performing participants overestimated their own performance, while high-performing participants underestimated their own performance. According to Kruger and Dunning (1999), deficiencies in online metacognitive monitoring due to limited knowledge or skills are the main cause of exaggeration. The number of articles investigating the Dunning Kruger effect in science courses is limited (Pazicni & Bauer, 2014). This situation confirms once again the value of the present study.

Objectives

Research on the Dunning Kruger effect in Turkey remains scarce (Kavak, 2013; Turğut & Tanhan, 2019). This gap hampers our understanding of the interplay between academic achievement and psychological variables—particularly among adolescents. The present study seeks to address this void by examining how eighth-grade students' academic performance in science relates to the DKE, moral maturity, and life positions. Specifically, it analyses inconsistencies between students' predicted and actual performance on the Acids–Bases topic and explores how these discrepancies connect to psychological variables.

- Is the difference between students' pre- and post-test accuracy estimates and their actual performance statistically significant across percentile ranks?
- Do moral maturity levels and life positions differ significantly among percentile groups?
- Is there a significant relationship among the Dunning–Kruger Effect, moral maturity, and life positions?

Methodology

Research Design

This study adopts a mixed design that integrates relational and descriptive survey models within the quantitative research paradigm. The relational survey model offers an analytical approach aimed at uncovering the nature and direction of interactions between independent and dependent variables, whereas the descriptive survey model focuses on the systematic observation and interpretation of extant phenomena within their natural contexts (Karasar, 2005). The relational component tests, with quantitative data, how psychological variables interact with academic achievement; the descriptive component enables an in-depth analysis of students' self-assessment tendencies and perceptual dynamics.

Study Group

The universe of this research consists of 8th grade students studying in secondary schools within the borders of Solhan district of Bingöl province. The sample of the research consists of 459 students who were selected by simple random sampling

technique and gave their consent to participate in the study. Although data were initially collected from 610 students, incomplete or inconsistent responses were excluded, and the final analyses were carried out with 459 valid cases. Simple random sampling is a sampling procedure without replacement in which every unit in the population has an equal probability of selection. Because this method is free from bias and systematic error, it yields a representative sample that most accurately reflects the structure of the population. It is widely regarded as one of the most reliable approaches for enhancing the generalisability and validity of research findings (Büyüköztürk et al., 2014), particularly for heterogeneous populations where balanced representation of participants' demographic and sociocultural diversity is essential. Ethics committee approval and informed consent were obtained meticulously, and all data-collection procedures complied with ethical principles.

Table 1 Demographic Characteristics of the Participants

		Frequency (n)	Percentage (%)
Gender	Male	241	52,4
	Female	218	47,6
Age	14	434	94,5
	15	25	5,5

According to the data presented in Table 1, 52.4% of the students in the study group are male, whereas 47.6% are female. With respect to age, the vast majority of participants are 14 years old (94.5%), and only a small proportion of the sample is 15 years old (5.5%).

Data Collection Instruments

In this study, three scales were used in the data collection process: Life-Based Skills Test related to the Acid-Base topic (LFBST), Moral Maturity Scale (MMS) and Life Positions Scale(LPS).

Within the scope of the research, a 17-item multiple-choice test called "Life-Based Skills Test" was developed to measure academic achievement on the subject of "acid-base" in the 8th grade science curriculum. The content validity of the test was ensured by the specification table created in line with the curriculum of the Ministry of National Education and the opinions of field experts. An initial 23-item pilot was analysed; five items were discarded based on item difficulty (.30–.60), discrimination (> .30), and item–total correlations ($p < .01$). The final test showed a mean difficulty of .44, a discrimination index of .42, and a KR-20 reliability coefficient of .85, indicating moderate difficulty, high discriminative power, and strong reliability. Scores were used as objective indicators of conceptual understanding.

MMS is a 66-item tool developed by Şengün and Kaya (2007) that measures the moral maturity levels of individuals with five-point Likert-type response options (1:Never–5:Always). Psychometric evaluation with a sample of 830 secondary-school students indicated a unidimensional structure, with factor loadings ranging from .40 to .82. Item–total correlations fell between .45 and .78, all statistically significant at $p < .001$. Reliability was supported by test–retest ($r = .84$), split-half ($r = .89$), and Cronbach's alpha ($\alpha = .95$). Fifty-two items are scored positively and fourteen are reverse-scored, yielding total scores from 66 to 330 (normative mean=273.11, SD=28.35). These indices demonstrate that the MMS validly and reliably measures moral maturity.

LPS was developed by Boholst based on the theoretical foundations of Eric Berne and adapted to Turkish by İşgör et al, (2012). It evaluates attitudes toward self and others along four dimensions:

- I'm OK – You're OK: mutually positive self and other-perceptions
- I'm OK – You're Not OK: self-enhancement at others' expense
- I'm Not OK – You're OK: self-devaluation
- I'm Not OK – You're Not OK: mutually negative perceptions

Each dimension comprises five items rated on a five-point Likert scale (1 = Never, 5 = Always). The highest score among the four dimensions reflects the respondent's dominant life position. The Turkish version has a Cronbach's alpha of .84, indicating satisfactory internal consistency.

In this study, in addition to descriptive and comparative analyses, cluster analysis was also conducted. This analysis created four distinct groups based on the correspondence between students' predicted and actual achievement. The groups are: those who underestimated their performance, those

who accurately predicted their performance, those who moderately overestimated their performance, and those who overestimated their performance.

Results

Data were analysed with SPSS. The dataset was first screened for missing and outlying values; anomalous cases were removed following data-cleaning protocols. Prior to scoring, the fourteen negatively worded MMS items were reverse-coded. Total MMS scores were then subjected to cluster analysis, yielding three moral-maturity categories:

- Low: 92–237 (n = 112, 24.4%)
- Moderate: 238–276 (n = 231, 50.3%)
- High: 277–326 (n = 116, 25.3%)

LPS data were analysed separately for each of the four subscales; descriptive statistics (means, standard deviations) were computed to identify participants' dominant life positions.

For the LFBST, each correct answer was scored as 1 and each incorrect answer as 0 (max = 17). Total scores were divided into quartiles for performance grouping:

- 1st Quartile: lowest 25% (0–4 points)
- 2nd Quartile: 26–50% (5–8 points)
- 3rd Quartile: 51–75% (9–12 points)
- 4th Quartile: highest 25% (13–17 points)

Concordance between students' self-estimates and actual scores on pre- and post-tests was evaluated using Pearson correlation. An alpha level of $p < .05$ was applied to all statistical tests. Visualisations (tables and graphs) were produced for clarity. Throughout the analysis, ethical data-handling principles were observed and cross-validation techniques employed to minimise bias.

The cluster analysis revealed four distinct student profiles. The first group underestimated their own achievement. The second group consisted of students whose estimates and actual achievements were similar. The third group moderately overestimated their achievements. The fourth group consisted of students who greatly overestimated their achievements.

LFBST scores were analyzed by dividing them into four quadrants to evaluate students' performance. The results are presented in Table 2.

Table 2 Distribution of Students Across Performance Quartiles

Student Group	Correct Answers	Frequency (n)	%
First quartile	14-17	68	14,2
Second quartile	10-13	91	19,8
Third quartile	6-9	136	29,6
Fourth quartile	1-5	164	36,4
Total		459	100

Table 2 shows that test scores are spread across four quartiles: A small upper quartile of 14.2% with 14-17 correct answers; a second quartile of 19.8% with 10-13 correct answers; a third quartile of 29.6% with 6-9 correct answers; and the largest lower quartile of 36.4% with 1-5 correct answers. The findings reveal that the achievement level is mostly clustered in the middle-low range, while students with high performance are relatively few.

Tables 3-5 present the findings regarding the students' pre-test and post-test percentile estimates and the percentiles they actually reached.

Table 3 Changes in Percentile Estimates Before and After the Test

Actual quartile		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Pre-test percentile estimate	68	29,9	25,3	-1,41	67	0,062	0,12
	Post-test percentile estimate	68	33,4	26,3				
Second quartile	Pre-test percentile estimate	91	35,5	23,4	-2,44	90	0,008	0,19
	Post-test percentile estimate	91	40,8	24,4				
Third quartile	Pre-test percentile estimate	136	42,3	25,9	-2,04	135	0,018	0,24
	Post-test percentile estimate	136	47,3	26,4				
Fourth quartile	Pre-test percentile estimate	164	43,3	25,5	-2,86	163	0,001	0,35
	Post-test percentile estimate	164	52,1	28,4				

The t-test results presented in Table 3 reveal that students' pre- and post-exam predictions of percentile position changed differently as they progressed up the achievement ladder. Those in the top performance bracket (14-17 correct responses) predicted their rank accurately before the exam; no significant change in predictions was found in this group, [t(67)=-1.41, p>.05]. On the other hand, the predictions of the participants with 10-13 correct answers increased from 35.5% to 40.8% [t(90)=-2.44, p<.05, d=.19], and the predictions of the group with 6-9 correct answers increased from 42.3% to 47.3% [t(135)=2.04, p<.05, d=.24], and the predictions of students in the lowest quartile with 1-5 correct answers jumped from 43.3% to 52.1% [t(163)=-2.86, p<.05, d=.35]. Although the effect sizes remained small in all three groups, the findings suggest that as the achievement level decreased, students tended to systematically overestimate their post-test position.

Table 4 Changes in Pre-Test Percentile Estimates

Actual Percentage Quarter		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Actual percentile	68	12,3	6,7	4,95	67	0,000	0,78
	Predicted percentile before the test	68	26,2	21,5				
Second quartile	Actual percentile	91	34,1	4,7	-0,95	90	0,289	0,14
	Predicted percentile before the test	91	31,8	19,7				
Third quartile	Actual percentile	136	54,2	6,2	-6,31	135	0,000	0,85
	Predicted percentile before the test	136	38,6	22,2				
Fourth quartile	Actual percentile	164	78,5	6,5	-16,25	163	0,000	2,13
	Predicted percentile before the test	164	39,6	21,9				

Table 4 summarizes the difference between participants' percentile predictions before the exam and their actual achievement by quartile. Participants in the first quartile significantly underestimated their achievement ($\bar{X}_{prediction}$=26.2; $\bar{X}_{actual}$=12.3), [t(67)=4.95, p<.05, d=0.78]. There was no significant difference in the second quarter ($\bar{X}_{prediction}$=35.5; $\bar{X}_{actual}$=40.8), [t(90)= -0.95, p > .05]. Participants in the third quartile overestimated their achievements ($\bar{X}_{prediction}$=38.6; $\bar{X}_{actual}$=54.2), [t(135)= -6.31, p<.05, d=0.85]. In the fourth quarter, the predicted tranches were well below the actual tranches ($\bar{X}_{prediction}$=39.6; $\bar{X}_{actual}$=78.5), [t(163)= -16.25, p<.05, d=2.13]. Thus, the forecast-performance mismatch is high in the first, third and fourth quarters and statistically insignificant in the second quarter.

Table 5 Changes in Percentile Estimates After the Test

Actual Percentage Quarter		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Actual percentile	68	12,4	6,8	6,01	67	0,000	0,92
	Predicted percentile before the test	68	29,6	22,6				
Second quartile	Actual percentile	91	34,1	4,7	1,19	90	0,164	0,18
	Predicted percentile before the test	91	37,1	20,7				
Third quartile	Actual percentile	136	54,2	6,1	-4,01	135	0,000	0,57
	Predicted percentile before the test	136	43,5	22,7				
Fourth quartile	Actual percentile	164	78,4	6,5	-11,67	163	0,000	1,47
	Predicted percentile before the test	164	48,4	24,7				

Table 5 summarizes the agreement between the participants' post-exam percentile predictions and their actual achievement percentiles by quartile. Individuals in the first quartile significantly overestimated their achievement ($\bar{X}_{prediction}$ = 29.6, $\bar{X}_{actual}$ = 12.4), [t(67) = 6.01, p < .05, d = 0.92]. In the second quarter, perception and reality overlapped, but the difference did not reach statistical significance, [t(90) = 1.19, p > .05]. In the third quarter, participants slightly underestimated

their achievement ($\bar{X}_{prediction}$=43.5, $\bar{X}_{actual}$=54.2), [t(90)=-4.01, p<.05, d=0.57]. In the fourth quarter, a significant underestimation was observed ($\bar{X}_{prediction}$=48.4, $\bar{X}_{actual}$=78.4), [t(163) = 11.67, p < .05, d = 1.47]. This finding suggests that the accuracy of the post-exam predictions was maintained only in the second quarter, whereas there was a significant mismatch in the other three quarters.

Table 6 Changes in Number of Correct Estimates Before and After the Test

Actual Percentage Quarter		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Estimation of the number of correct before the test	68	13,21	2,49	7,06	67	0,000	0,76
	Estimation of the number of correct after the test	68	10,37	4,03				
Second quartile	Estimation of the number of correct before the test	91	13,10	2,74	6,32	90	0,000	0,80
	Estimation of the number of correct after the test	91	9,95	4,08				
Third quartile	Estimation of the number of correct before the test	136	12,89	2,87	7,09	135	0,000	0,88
	Estimation of the number of correct after the test	136	8,86	5,00				
Fourth quartile	Estimation of the number of correct before the test	164	12,54	3,74	11,48	163	0,000	1,32
	Estimation of the number of correct after the test	164	5,72	5,30				

Table 6 compares participants' predictions of the number of correct answers reported before and after the exam by quartile. Individuals in the first quartile expected an average of 13.21 correct answers before the exam and reduced this value to 10.37 after the exam; the difference is significant, [t(67)=7.06, p<;.05, d=0.76]. A similar reduction was observed in the second quarter ($\bar{X}_{prediction}$ = 13.10; $\bar{X}_{actual}$ = 9.95), [t(90)=6.32, p<;.05, d=0.80]. Participants in the third quarter reduced their prediction from 12.89 to 8.86, and the difference was again significant, [t(135)=7.09, p<;.05, d=0.88]. The most significant deviation occurred in the fourth quarter; while the pre-exam prediction was 12.54, the post-exam prediction decreased to 5.72, [t(163)=11.48, p<;.05, d=1.32]. Thus, in all quarters, participants systematically overestimated their actual performance before the exam and significantly corrected their estimates after the exam, but the miscalibration was not fully corrected.

Table 7 Table of Changes in Number of Correct Predictions Before Test

Actual Percentage Quarter		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Estimated number of correct before the test	68	13,21	2,49	5,33	67	0,000	0,83
	Actual number of correct	68	11,21	1,67				
Second quartile	Estimated number of correct before the test	91	13,10	2,74	13,41	90	0,000	2,10
	Actual number of correct	91	8,44	0,44				
Third quartile	Estimated number of correct before the test	136	12,89	2,87	19,62	135	0,000	2,81
	Actual number of correct	136	6,24	0,74				
Fourth quartile	Estimated number of correct before the test	164	12,55	3,74	23,12	163	0,000	3,00
	Actual number of correct	164	3,01	1,38				

Table 7 summarizes the differences between pre-test predictions and actual test performance. Participants in the first quartile predicted that they would get 13.21 correct on average, while in reality they got 11.21 correct; the difference was significant, [t(67)=5.34, p<.05, d = 0.83]. Similarly, predictions were systematically overestimated in the second ($\bar{X}_{prediction}$=13.10; $\bar{X}_{actual}$=8.44), third ($\bar{X}_{prediction}$=12.89; $\bar{X}_{actual}$=6.24) and fourth ($\bar{X}_{prediction}$=12.55; $\bar{X}_{actual}$=3.01) quarters; the corresponding t-values are 13.41, 19.62 and 23.12, respectively, all significant at p<.05. Cohen's d coefficients range between 2.10-3.00, indicating large effect sizes. These findings suggest that the participants significantly overestimated their performance and that the overestimation was especially evident at low performance levels.

Table 8 Table of Changes in Number of Correct Predictions After the Test

Actual Percentage Quarter		n	$\bar{X}$	ss	t	sd	p	Cohen d
First quartile	Estimated number correct after the test	68	10,37	4,03	-1,65	67	0,059	0,24
	Actual number correct	68	11,21	1,67				
Second quartile	Estimated number correct after the test	91	9,95	4,08	3,00	90	0,000	0,46
	Actual number correct	91	8,44	0,44				
Third quartile	Estimated number correct after the test	136	8,86	5,00	4,63	135	0,000	0,64
	Actual number correct	136	6,24	0,74				
Fourth quartile	Estimated number correct after the test	164	5,72	6,30	4,93	163	0,000	0,62
	Actual number correct	164	3,01	1,37				

The comparisons presented in Table 8 show the extent to which post-test performance predictions corresponded to actual performance. For participants in the first quartile, there was no significant difference between predicted and actual results [t(67)= -1.65, p>;.05, d = .24]. On the other hand, individuals in the second quartile overestimated their achievements in a statistically significant way [t(90)=3.00, <;.05, d = .46]. The same trend continues in the third quartile; participants rated themselves as more successful than they actually were [t(135) = 4.63, <;.05, d = .64]. The deviation in the fourth quarter continues to be significant [t(163) = 4.93, <;.05, d = .62]. The means of $\bar{X}_{prediction}$=10.37; $\bar{X}_{actual}$=11.21; $\bar{X}_{prediction}$=9.95; $\bar{X}_{actual}$=8.44; $\bar{X}_{prediction}$=8.86; $\bar{X}_{actual}$=6.24 and $\bar{X}_{prediction}$=5.72; $\bar{X}_{actual}$=3.01, respectively, indicate that prediction bias increases as performance decreases to lower levels.

Table 9 Chi-square Test Analysis of the Difference Between Life Positions and Moral Maturity on the Groups in the First and Last Quarters

		Actual Percentile Quarter				χ2	sd	P
		First quartile		Last Quartile				
		n	%	n	%			
Moral Maturity Categorical	Low	12	17,6	36	22,0	1,357	2	0,467
	Medium	31	45,6	57	34,8			
	High	25	36,7	71	43,3			
Life Position	I'm OK You're OK	29	42,6	83	50,6	0,378	3	0,835
	I'm OK You're not OK	18	26,5	48	29,3			
	I'm not OK You're OK	13	19,1	19	11,6			
	I'm not OK You're not OK	8	11,8	14	8,5			

The chi-square analysis presented in Table 9 examined the relationship between levels of moral maturity and being in the bottom (4th quartile) and top (1st quartile) of the achievement distribution. The findings showed that there was no significant relationship between the variables, $\chi^2(2, N=232)=1.36, p > .05$; Cramér V≈.08. In other words, participants in the highest and lowest achievement groups were statistically indistinguishable in terms of moral maturity. Similarly, the analysis presented in Table 9 did not indicate a significant relationship between life position levels and the same extreme performance groups, $\chi^2(3, N=232)=0.38$, p>.05; Cramér V≈.04). As seen in Table 9, no significant relationship was found between moral maturity and life positions and achievement quartiles. Although this finding was one of the aims of the study, it was not as expected. This result suggests that these variables do not directly affect achievement levels. Furthermore, other variables may be intervening. Therefore, future research should re-examine these relationships with larger samples and different methods.

Discussion and Conclusion

This study examines the relationship between individuals' academic performance predictions and their actual achievement and how their self-evaluation skills differ according to their competency levels. The findings revealed that there was a moderate positive correlation between the participants' pre- and post-test predictions and their actual performance. Somyürek & Çelik (2018) evaluated the personal evaluations of 128 participants between the ages of 22-45 towards the academic achievement test of instructional technologies and material development course. As a result of the study in which data were collected with a 40-question academic achievement test, it was determined that 87.5% of the participants' actual achievement and predicted achievement were different from each other; 71.9% of the participants predicted their predictions about their achievement at a higher level than their actual achievement. In the study, it was determined that the Dunning-Kruger effect scores of the participants did not differ statistically significantly according to their gender and departments; however, there was a negative and strong relationship between overconfidence scores and actual achievement. Dhindra (2024), in his qualitative study aiming to explain the effect of the Dunning-Kruger effect on academic achievement and self-assessment accuracy, found that the Dunning-Kruger effect was widely observed in different grade levels and courses, and that individuals with low achievement levels overestimated their abilities at a statistically significant level, while individuals with high achievement levels underestimated their current potential. Modranský (2016), in his mixed study aiming to reveal the Dunning-Kruger effect, which has a lack of knowledge and does not recognize this deficiency, both evaluated the studies published between 1999 and 2016 and compared the exam performances and self-evaluations of 141 university students. As a result of the study, it was determined that individuals with low academic achievement had higher self-evaluation scores, while individuals with higher academic achievement had lower self-evaluation scores. This result is consistent with the "self-evaluation fallacy" theory proposed by Kruger and Dunning (1999). However, this relationship showed significant differences according to percentile groups. The fact that participants in the first quartile significantly underestimated their actual performance can be explained by the "false consensus effect". This situation is recognized in the literature as the relative underestimation of their own skills by individuals with high competence levels. On the other hand, the overestimation of performance by participants in the fourth quartile reflects the "double curse" phenomenon described by Kruger and Dunning (1999). This paradox emphasizes that the lack of self-evaluation skills of individuals with low levels of competence has a negative impact on both performance and self-confidence. It is noteworthy that in post-test predictions, participants in the second quartile made predictions that were closest to their actual performance. This finding indicates that individuals with intermediate competence exhibit a more realistic attitude towards self-evaluation. However, the statistically significant differences between pre-test predictions and actual performance in all groups indicate that self-evaluation skills are generally limited.

The study revealed that individuals' self-assessment skills are directly related to their competence levels and this relationship is consistent with the "Dunning-Kruger effect" in the literature. The prediction errors, which are particularly evident in the low and high performance groups, indicate that individuals are not able to objectively assess their own level of knowledge. Participants in the second quartile made more accurate predictions, suggesting that intermediate competence has a positive effect on self-awareness. The trend of increased self-confidence in post-test predictions suggests that performance feedback may temporarily change individuals' self-efficacy perceptions. However, the permanence of this change and its impact on educational outcomes should be examined in depth in future studies.

In this context, the following suggestions can be made: First of all, modules that develop self-awareness and feedback analysis skills should be designed especially for students in low and high achievement groups. These modules can increase students' competencies in realistic goal setting and performance evaluation. Second, activities that support children's self-assessment skills (e.g. self-

reporting and peer assessment) should be integrated into the curriculum in pre-primary and primary education. Third, educational institutions should develop digital platforms that instantly analyze discrepancies between students' exam predictions and actual performance and provide personalized feedback. Furthermore, long-term longitudinal studies examining the change in self-assessment skills over time and their impact on academic achievement should be supported. Finally, it is critical to train teachers to recognize students' self-assessment errors and develop strategies to correct them.

The findings of the study suggest that developing individuals' self-evaluation competencies will optimize not only academic achievement but also decision-making processes in social and professional life. However, the impact of cultural and contextual factors on these dynamics should be addressed from an interdisciplinary perspective in future research.

This study has some limitations. The research was conducted in only one district. Only the topic of acid-base was examined. Because the study was cross-sectional, the results cannot demonstrate causality. Future research should use larger samples from different regions. Different science topics should be examined. Furthermore, longitudinal studies should be conducted to track changes in students' self-assessment skills over time.

Acknowledgment

This study was derived from the first author's master's thesis. AI-assisted tools were employed to enhance the language, grammar, and spelling accuracy of this manuscript.

References

Büyüköztürk, Ş., Kılıç, Ç. E., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2014). *Scientific research methods*. Pegem Publishing.

Dhindra, G. (2024). The Dunning–Kruger effect: A hidden obstacle to student success: A comprehensive study. *World Academy of Informatics and Management Sciences*, *13*(3), 13301-13306.

Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognise their own incompetence. *Current Directions in Psychological Science*, *12*(3), 83-87.

Ehrlinger, J., Johnson, K., Banner, M, Dunning, D., and Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes*, *105*, 98-121.

Isgor, I. Y., Kaygusuz, C., & Ozpolat, A. R. (2012). Life positions scale: Language equivalence, reliability, and validity analysis. *Procedia - Social and Behavioral Sciences*, *47*, 284-291.

Karasar, N. (2005). Scientific research method. Nobel Academic Publishing.

Kavak, K. (2013). Conservative politics in the grip of the Dunning–Kruger syndrome. *Turkish Academy for Political, Social and Strategic Studies*.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognising one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121-1134.

Marottoli, R. A., & Richardson, E. D. (1998). Confidence in, and self-rating of, driving ability among older drivers. *Accident Analysis & Prevention*, *30*(3), 331-336.

Modranský, R. M. (2016). The Dunning–Kruger effect: A review of significant findings. *Journal of Global Science*, *1*(1), 1-6.

Pazicni, S., & Bauer, C. F. (2014). Characterizing illusions of competence in introductory chemistry students. *Chemistry Education Research and Practice*, *15*(1), 24-34.

Plotnik, R. (2009). *Introduction to psychology*. Kaknüs Publishing.

Sheldon, O. J., Dunning, D., & Ames, D. R. (2014). Emotionally unskilled, unaware, and uninterested in learning more: Reactions to feedback about deficits in emotional intelligence. *Journal of Applied Psychology*, *99*(1), 125-137.

Simons, D. J. (2013). Unskilled and optimistic: Overconfident predictions despite calibrated knowledge of relative skill. *Psychonomic Bulletin & Review*, *20*, 601-607.

Somyürek, S., & Çelik, İ. (2018). Dunning–Kruger syndrome and subjective judgments. *Educational Technology: Theory and Practice, 8*(1), 141-157.

Sullivan, P. J., Ragogna, M., & Dithurbide, L. (2019). An investigation into the Dunning–Kruger effect in sport coaching. *International Journal of Sport and Exercise Psychology*, *17*, 591-599.

Şengün, M., & Kaya, M. (2007). The scale of moral-maturity: The study of its validity and reliability. *Ondokuz Mayıs University Review of the Faculty of Divnity*, 24(24-25), 51-64.

Turğut, E., & Tanhan, F. (2019). An investigation of the Dunning–Kruger effect among high-school students in terms of moral maturity and life position. *Journal of Human Sciences*, *16*(1), 102-117.

Yeniçeri, Z., & Dönmez, A. (2008). Perception of terrorism and terrorist: How important who holds the gun?. *Turkish Journal of Psychology*, *23*(62), 93-107.

Author Details

Meryem Çetin, *Fırat University, Turkey,* ***Email ID****: meryemcetin23@hotmail.com*

Erol Asiltürk, *Fırat University, Turkey,* ***Email ID****: ecil@firat.edu.tr*

Mehmet Polat, *Fırat University, Turkey,* ***Email ID****: hvpltyzbsmplt@gmail.com*

Gülen Önal Karakoyun, *Van Yüzüncü Yıl University, Turkey,* ***Email ID****: gulenonal@yyu.edu.tr*