Motivation and Feelings of Competence Among University Students in Introductory Physics

Ana Romero-Abrio[1] · Santiago Hurtado-Bermúdez[2]

Abstract

This work analyzes whether there is a cognitive bias between the ideal perception of the skills and the real performance in an introductory physics class, and additionally, whether predictions of students' performance are related to various motivational variables. We examined through a validated survey and network analysis the relationship between several motivational aspects and volitional variables with the accuracy of their predictions. The results show that the students present a motivational bias when students' desires were considered, mainly in the students with low academic performance. Finally, it is necessary to explore the development of specific interventions that target the motivations of students, in order to be effective, and to reduce the gap between expected and actual grades, increasing students' metacognitive skills and thus their academic performance.

Introduction

Students who are starting their studies in scientific-technical degrees usually have a low opinion of their skills and competences in subjects such as Physics or Calculus. They also underestimate the study effort needed to pass the subject. (Blackmore et al., 2021; Fakcharoenphol et al., 2015; Galloway et al., 2013; Morphew, 2021). It is usual for students to fail in the first exams and choose to drop the course that creates a feeling of frustration (King, 2015; Lippert, 2020; Morphew, 2021). Additionally, the syllabus for high school physics in Spain does not completely overlap with physics syllabus for engineering found in first year of university, which is more oriented toward applied physics. So for most students, the discipline-specific physics course is their first encounter with those topics (Seidel et al., 2008). On the other hand, many students view introductory physics as a required hurdle and focus strictly on how to do exam-like problems. (Buckman et al., 1975).

✉ Santiago Hurtado-Bermúdez
shurtado@us.es

1   Departamento de Educación Y Psicología Social, Facultad de Ciencias Sociales, Universidad Pablo de Olavide de Sevilla, 41013 Sevilla, Spain

2   Departamento de Física Aplicada II, ETSA, Universidad de Sevilla, Av. Reina Mercedes 2, 41012 Sevilla, Spain

One of the possible causes of this academic failure may be because they have an illusion of competence. This cognitive bias is known as the Dunning-Kruger effect (Kruger & Dunning, 2002). This effect shows that low-performing students often over-estimate their performance on exams and tend to be overconfident. Moreover, this lack of conscience regarding their low performance implies that these students do not usually try to improve that performance. That is, these students do not have metacognitive self-awareness of their low performance, nor how to get it (Dunning et al., 2004). However, the students are able to improve their self-perception with metacognitive training (Kruger & Dunning, 2002).

Several studies have focused on the clinical and non-academic field in order to find out the psychological and social mechanisms of this effect, ranging from humor, logical reasoning, and grammar (Kruger & Dunning, 1999), geography (Ehrlinger & Dunning, 2003), to criticism of science (Caputo & Dunning, 2005). Even so, several studies have criticized that this cognitive bias actually consists of a statistical trap based on unreliable measurements (Krueger & Mueller, 2002), but they have subsequently been refuted conclusively (Freund & Kasten, 2012).

In the field of educational psychology, various empirical studies have been carried out (Bol et al., 2005; Nietfeld et al., 2006), although only a few have focused on the field of STEM (Science, Technology, Engineering and Mathematics) education. The presence of the effect has been corroborated, for example, in the reading comprehension (Huff & Nietfeld, 2009); biology (Rachmatullah & Ha, 2019), chemistry (Pazicni & Bauer, 2014), or mathematics course (Labuhn et al., 2010), but, as far as we know, there have been no studies applied to computer engineering. In all these studies, the variables on which the presence of cognitive bias depends have not been fully identified.

As a result, several studies have focused on finding the causes of this cognitive bias, indicating some of them the origin in the motivation and self-image protection (Blanton et al., 2001), while others are inclined toward a limited information processing (Chambers & Windschitl, 2004), or previously conceived beliefs about their skill and knowledge (Critcher & Dunning, 2009). A recent work suggests that metacognitive judgments may be affected by motivation, for example, when students make predictions about their performance on a forthcoming exam, they may, explicitly or implicitly, take into account the desired grades (Bol et al., 2005). That is, student motivations can influence performance predictions (Saenz et al., 2017). Since academic performance is an important aspect of students' self-concept (Elliot & McGregor, 2001) and self-esteem (Crocker et al., 2003), it is interesting to study how motivational variables play a role in self-assessments.

The Present Study

In view of all this, the following objectives are proposed in this paper:

1. To analyze whether there is a cognitive bias between the ideal perception of the skills and the real performance in the exams (Dunning-Kruger effect) in the first year of the Computer Engineering Degree course.
2. To determine whether predictions of students' performance are related to various motivational variables.

Method

Participants

In this study, the students were enrolled in the first course at the Pablo de Olavide University (Spain) during the 2017–2018 academic year. The students attended the face-to-face Fundamentals of Physics subject in the Degree in Computer Engineering. The students were invited to participate in the study during the final semester exam. Only 72 of the 82 enrolled students (76.8%) signed the informed consent; a total of 57 males (90.4%) and 6 females (9.6%) participated in the study, a usual ratio in Computer Science degree. The consent form did not include any specific reference that it was a Dunning-Kruger effect study. The study was approved by the Ethics Committee of the University Pablo Olavide.

Materials

Procedure

The students were informed of the details of the survey and the objective of the study. The instrument was applied before beginning the exam at the end of the semester. The exam grade used a scale from 0 up to 10 points. Participation was voluntary and anonymous, although students who did participate were rewarded with a 0.1 point incentive for completing the questionnaire, which in other studies has been found not to affect the results (Ehrlinger et al., 2008).

For all calculations, the statistical package R version 3.4.3 was used for quantitative analysis of the results including questionnaire reliability and validity statistics, t test for paired samples, and multiple regression. Additionally, to assess the connection between predicted performance and selected variables, a network analysis (NA) approach was applied. NA does not involve any a priori assumption and accepts simultaneous testing of reciprocal interplay between variables. Consequently, this methodological approach overcomes the need to hypothesize about causal latent variables (Borsboom, 2017).

NA was performed using the qgraph package of R (Epskamp et al., 2012). Variables are represented in a network by nodes related through edges between the nodes. Networks are based on partial correlations associating two nodes while considering the influence of all other variables. The edge is colored blue/green or red corresponding to positive or negative correlation/covariance between variables, respectively. And the magnitude of the association is characterized by the thickness of the edge. In a network, the relative importance of a node in the context of the other nodes is provided by centrality indices. There are three different and highly related measures of centrality implemented in qgraph package: strength (referring to the sum of the strength of all connections between a specific node and all other nodes), closeness (referring to the average distance involving indirect connections from that node to all the other nodes), and betweenness (referring to the relevance of that node in the average pathway between two other nodes) (McNally, 2016).

In order to facilitate the interpretation of the results, a "least absolute shrinkage and selection operator" (LASSO) regularization was applied making small correlations automatically reduce to zero and retaining only significant associations (Friedman et al., 2008). In this analysis, extended Bayesian information criterion (EBIC) was set to 0.5 enhancing

both the accuracy and interpretability of network through the degree of regularization/penalty applied to sparse correlations (Chen & Chen, 2008).

Finally, in order to estimate the accuracy of the network, confidence intervals (95% CI) on the edge-weights were calculated for their estimates (Epskamp et al., 2018). CI were obtained through non-parametric bootstrapping (nboots = 2500) creating new plausible datasets with resampled observations in the data. Network stability analysis was performed using the bootnet package.

Questionnaire

The instrument was an in-house questionnaire and measures the expected exam grade in the subject for each student with respect to the rest of his or her classmates, following a common procedure in other studies (Pazicni & Bauer, 2014), as well as various motivational parameters about the student's desires with respect to his or her assessment (Saenz et al., 2017). The questionnaire consisted of 5 items.

The first item (Expected Grade) was related to the Dunning-Kruger effect and was concerned with measuring the test scores self-predicted by the students, "What grade (out of 10.00) do you think you will score on this exam?". The second item (Expected Range) was associated to the test scores self-predicted by the students relative to the grades of their peers, "Please indicate what range in % you expect your grade to be in with respect to the rest of the class (mark with an X)," where it can take values from 1 ("I'm at the very bottom"), up to 99 ("I'm at the very top") in increments of 10, and going through 50 ("I'm exactly at the average"). It is important to note that students were not aware of the performance of their peers.

Items 3 and 4 explored how students' motivations influence their expected grades. Item 3 studied positive motivation through the ideal grades that students would like to receive (Ideal Grade), "What is your ideal grade (out of 10.00) considering your individual efforts for this exam?". And item 4 (Minimum Grade) considered negative motivation through the minimum grades that students would be happy with, "What is the lowest grade (out of 10.00) you would be happy with on this test."

Finally, item 5 (Effort) explored students' volition through an overall score of their level of preparation, "Compared to your prior efforts, how much did you prepare for this exam?", using a Likert 5-point scale (1: Not at all; 5: More than ever).

A test of the validity of the questionnaire was performed using exploratory factor analysis (EFA). Previously, Bartlett's test of sphericity ($\chi^2 = 196$; $p = <0.001$) revealed the existence of high dependence among the five items (Hair et al., 2010). Additionally, the value of the Kaiser–Meyer–Olkin index was > 0.6; and therefore, it was determined that data were adequate to carry out a factor analysis (Hair et al., 2010).

EFA using maximum likelihood (ML) with oblimin rotation and parallel analysis was performed. The exploratory factor analysis revealed a unique factor that explained 61.5% of the total variance with an eigenvalue of 3.08. As observed in Table 1, items showed factor loadings varying between 0.335 (Effort) and 0.938 (Expected grade), being the lowest factor loading minimally accepted (Hair et al., 2010). Finally, Cattell's scree plot (see Fig. 1) showed that the one-factor model was satisfactory to represent the data. The items seem to be measuring the same construct.

A test of the reliability of the questionnaire was performed using a Cronbach alpha (see Table 2).

Table 1 Factor Loadings for each item of the questionnaire

	Factor	Uniqueness
Ideal grade	0.791	0.375
Minimum grade	0.854	0.270
Expected grade	0.938	0.119
Effort	0.335	0.888
Expected range	0.852	0.273

"Maximum likelihood" extraction method was used in combination with an "oblimin" rotation
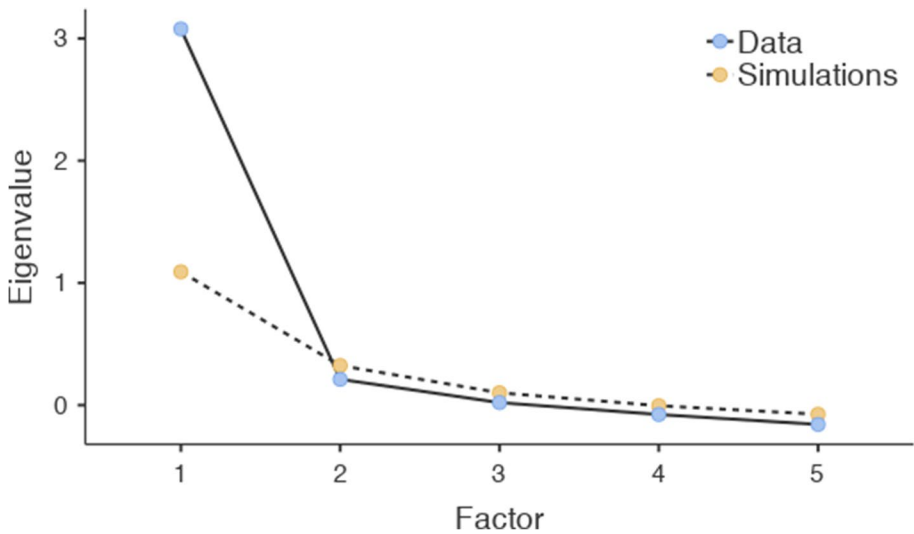

Fig. 1 Cattell's Scree plot

Table 2 Questionnaire reliability statistics

	McDonald's ω (standarized)	Cronbach's α	Gutmann's λ6	Average interitem correlation
Questionnaire	0.880	0.862	0.882	0.555

Of the observations, 63 were used, 0 were excluded listwise, and 63 were provided

The measurement of reliability using Cronbach's alpha assumes that the items (measured in Likert-type scale) measure the same construct, and are highly correlated (Zinbarg et al., 2005). The closer the alpha value is to 1, the greater the internal consistency of the scale items. In the first instance, the value obtained (0.392) highlighted the limitations of the unstandardized Cronbach's alpha coefficient, as it is affected by items measured with very different scales (see item 2). A satisfactory standardized Cronbach's alpha coefficient was obtained (0.862). Additionally, McDonald's omega coefficient was used, which uses factorial loads making the calculations more stable (Gerbing & Anderson, 1988), and does not depend on the number of items. A reliability value of the McDonald's omega

coefficient between 0.70 and 0.90 (Katz, 2003) is normally accepted, so our value of 0.88 strengthens the reliability of the questionnaire used. Guttman's λ_6 was also used as an additional measure of reliability. This coefficient uses the amount of an item's variance which is predictable by all of the other variable, so our value of 0.882 indicates acceptable reliability coefficients. Finally, average inter-item correlation compares correlations between all pairs of items that test the same construct by calculating the mean of all paired correlations. The obtained value exceeds 0.30, thus the construct validity was satisfied (Robinson et al., 1991).

Analyses and Results

First, we have focused on analyzing the trend of these data in a graph, and then perform a statistical analysis of them.

Figure 2 shows the expected grades of the students against the actual grades obtained in the exam. It is important to note that students' grades have not been moderated inflating the average grade. If all students had perfect knowledge of their exam performance, the relationship between predicted and actual grades would be a straight line of 45° slope shown in Fig. 2. However, it can be seen how the relationship follows the usual tendency in studies of the Dunning-Kruger effect. Thus, lower performing students tend to overestimate their grades, and higher performing students tend to underestimate them, as seen by comparing the solid line related to the ideal line trend, and the dashed line that represents the trend of the relationship between the two variables.

On the other hand, students have been divided into quartiles according to their performance, from the lowest 25% (Q1) to the highest 25% (Q4) (see Fig. 3). Students in the lower quartile tend to overestimate their performance on the exam, their actual performance being around the 10th percentile, but they perceive their performance to be around the 26th percentile. However, students in the higher performance quartiles underestimate their exam grades, so students in the higher quartile underestimate their performance on
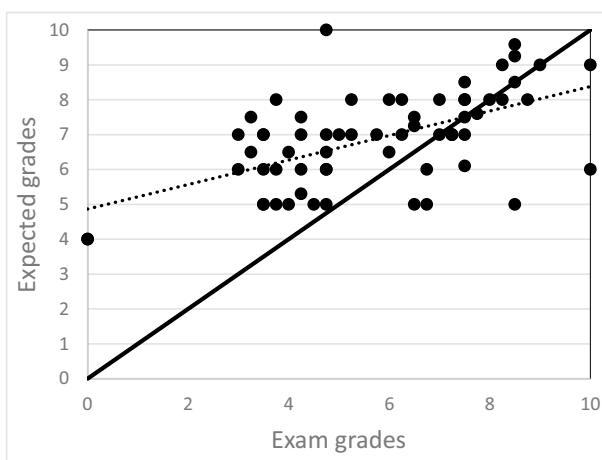

Fig. 2 Expected student exam grades versus actual grades. The solid line indicates the ideal relationship between the two grades, while the dashed line represents the adjustment to a linear function of both variables
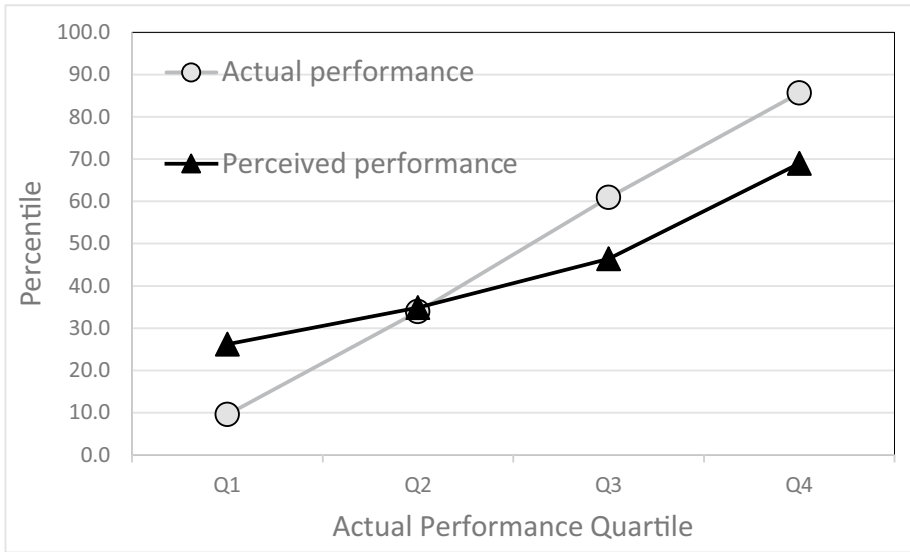
Fig. 3 Perceived and actual percentiles for student performance on the exam based on actual performance rank. Students' exam grades were grouped into four quartiles (Q1=lowest, Q4=highest). The averages of actual and perceived performance (both expressed as percentiles) for each rank were represented as a function of the actual performance rank

the exam, so that students in the 90th percentile perceive their performance around the 69th percentile.

As explained above, students were ranked according to their performance based on actual grades. Differences between actual and perceived scores were analyzed using a t test for paired samples. The difference between the two grades was significant for both the worst quartile Q1, $t=11.170$ and $p<0.001$, and for the next quartile Q2, $t=6.180$ and $p<0.0001$. On the other hand, these differences disappear in students with higher performance, $t=0.848$ and $p=0.410$ for Q3, and $t=1.275$ and $p=0.222$ for Q4 (see Table 3).

In contrast, in the case of female students, as they are still a small percentage of computer engineering students (Buse et al., 2017), the sample is very small, which makes it impossible to obtain reliable statistical parameters.

This study has gone a step further by examining the contribution of students' motivations to their predictions about their exam grades. The students' desired grades were

Table 3 Mean, standard deviation, and t test for expected and actual grades by performance quartile

	Q1		Q2		Q3		Q4	
	EG	AG	EG	AG	EG	AG	EG	AG
Mean	6.033	3.017	6.581	4.703	6.984	6.766	7.908	8.375
SD	1.217	1.255	1.306	0.458	0.959	0.504	1.296	0.796
t	11.170		6.180		0.848		1.275	
p	<0.001		<0.001		0.410		0.222	

SD standard deviation, *EG* expected grade, *AG* actual grade

obtained through the questions described above where they had to indicate the grade they would like to receive (Ideal Grade), along with the lowest grade they would not mind obtaining (Minimum Grade). In this way, both the positive motivation was evaluated, measuring the desired grade they expected to obtain given their own circumstances and preparation (Ideal Grade), and the negative motivation through the lowest grade with which the students would be happy.

To compare whether other types of variables might be related to the desired or expected grades, students were also asked to indicate the effort they made prior to the examination. In this way, it is possible to analyze whether this information is not associated with the perceived grades, which would reinforce the presence of desires and motivations in the metacognitive processes of the students.

Thus, all variables, both motivational (Ideal Grade and Minimum Grade) and volitional (Effort), were used as data in a multiple regression analysis to determine how they could predict the grades (absolute prediction) or range (relative prediction) perceived by students. The multiple regression analysis was preceded by an analysis of multicollinearity computing a variance inflation factor (VIF) for each independent variable. In the presence of multicollinearity, regression estimates are unstable and have high standard errors. VIF measures how much the variance of a regression coefficient is inflated due to multicollinearity in the model. The obtained VIF values for the predictor variables (see Table 4) was low (VIF < 4) indicating lack of multi-collinearity (Hair et al., 2010).

Regarding multiple regression analysis, it was obtained that all the variables together constituted a good predictive model ($F = 48.38$, $R^2 = 0.711$, $p < 0.001$) of the grades perceived by the students. Among these variables, volitional one was not a good predictor of the model (Effort: $\beta = -0.027$, $t = -0.202$, $p = 0.840$) (see Table 4). On the other hand, the variables that reflect the student's motivation were very good predictors of the behavior of the expected grades (Ideal Grade: $\beta = 0.360$, $t = 4.266$, $p < 0.001$; Minimum Grade: $\beta = 0.545$, $t = 5.176$, $p < 0.001$). In addition, the co-linearity between both explanatory variables of the model was checked, obtaining a condition index lower than 20, so that both variables are associated with a weak co-linearity (Goldstein, 1993).

Regarding the range perceived by the students, multiple regression analysis produced also a good predictive model ($F = 30.53$, $R^2 = 0.608$, $p < 0.001$). There was only one variable that was a very good predictor of the model (Minimum Grade: = 0.410, $t = 6.049$, $p = <0.001$) (see Table 4), whereas the remaining variables did not perform as well (Ideal Grade: $\beta = 0.078$, $t = 1.436$, $p = 0.156$; Effort: $\beta = -0.059$, $t = -0.680$, $p = 0.499$).

However, due to the significance of Minimum Grade, and the non-significance of Ideal Grade, further information is needed to produce a complete understanding of these findings. We examined the predictions corresponding to the low performance quartile (Q1)

Table 4 Estimates of expected grade and expected range through multiple regression using the volitional variable (Effort) and motivational variables (Ideal Grade and Minimum Grade)

	Expected grade			Expected range			
	β	t	p	β	t	p	VIF
Intercept	1.291	2.306	0.025	0.311	0.864	0.391	–
Ideal grade	0.360	4.266	<0.001	0.078	1.436	0.156	2.084
Minimum grade	0.545	5.176	<0.001	0.410	6.049	<0.001	1.864
Effort	−0.027	−0.202	0.840	−0.059	−0.680	0.499	1.216

and the high-performance quartile (Q4), and the items related to them. All items were entered into a multiple regression including the volitional item (Effort) and the motivational items (Minimum and Ideal Grade). At the end, we obtained a significant model for Q1 ($F = 12.92$, $R^2 = 0.779$, $p < 0.001$) and Q4 ($F = 18.09$, $R^2 = 0.819$, $p < 0.001$). Results showed that different motivational items were associated with grade predictions for Q1 and Q4 (see Table 5).

Thus, it appears that high-performing students were strongly associated with Ideal Grade, and low-performing students with Minimum Grade.

Regarding network analysis, the network structure of Ideal Grade, Minimum Grade, Effort, Predicted Grade, and Predicted Range is reported in Fig. 4 (Network-1). Network analysis allows us to examine how the items relate to each other and can reveal important structural relationships that regression cannot reveal. Green lines indicate positive correlations while red lines represented negative correlation.

The network centrality indices are plotted in Fig. 5. Expected grade and Expected range had the highest strength value (or degree) and higher closeness values because they have strong associations to the nodes nearby. Expected grade and Expected range play an essential role in the network and their activation influences strongly to the other nodes. However, Ideal Grade had the highest betweenness values acting as the bridge between the communities of nodes.

Central stability is shown in Fig. 6. The percentage of the students included in the calculation of the centrality indices was decreased obtaining the correlation between the indices from the subsample and the indices from the original entire sample of students. The centrality indices become unstable when the correlation goes below 0.7 for the subset sample involving 50% of the original sample. Strength and closeness had the highest stability; therefore, both indices are interpretable with some care, while betweenness is not (Epskamp et al., 2018).

The bootstrapped confidence intervals of estimated edge-weights are reported in Fig. 7. The resulting plots reveal sizable bootstrapped CIs around the estimated edge-weights, indicating that many edge-weights likely do not significantly differ from one-another. The generally large bootstrapped CIs imply that interpreting the order of most edges in the network should be done with care. The edges Expected grade-Expected range, Expected range-Minimum grade, and Effort-Expected range are reliably the three strongest edges since their bootstrapped CIs do not overlap with most other of the bootstrapped CIs. The network should be interpreted with caution due to the shape of Cis obtained for the edge weights.

Additionally, since there is a positive correlation between motivation as a key factor in academic performance (Pintrich & de Groot, 1990), the relationship between the

Table 5 Coefficients of a multiple regression parameters for Q1 and Q4 performance quartiles showing the strongest correlated item

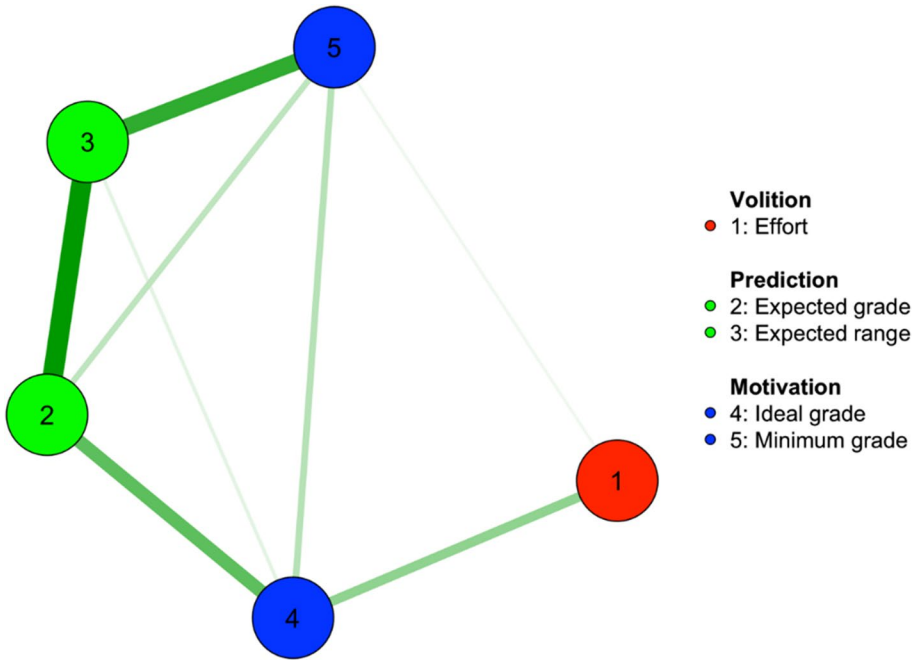|	Q1			Q4		
	β	t	p	β	t	p
Intercept	1.075	1.208	0.252	0.479	0.449	0.662
Ideal grade	0.198	1.708	0.116	0.614	3.287	0.006
Minimum grade	0.863	4.783	<0.001	0.392	2.893	0.013
Effort	−0.129	−0.642	0.534	−0.144	−0.584	0.570

Fig. 4 Partial correlation network constructed using EBIC-glasso depicting the association between volitional and motivational variables, and predictions (Network-1). Red lines denote negative associations between variables; green lines imply positive associations between variables; thicker lines reveal stronger associations
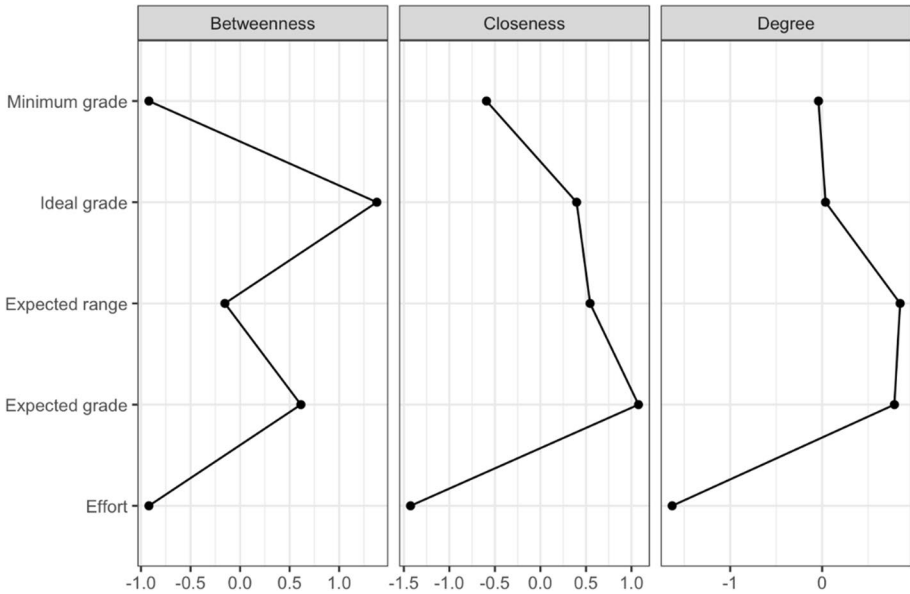

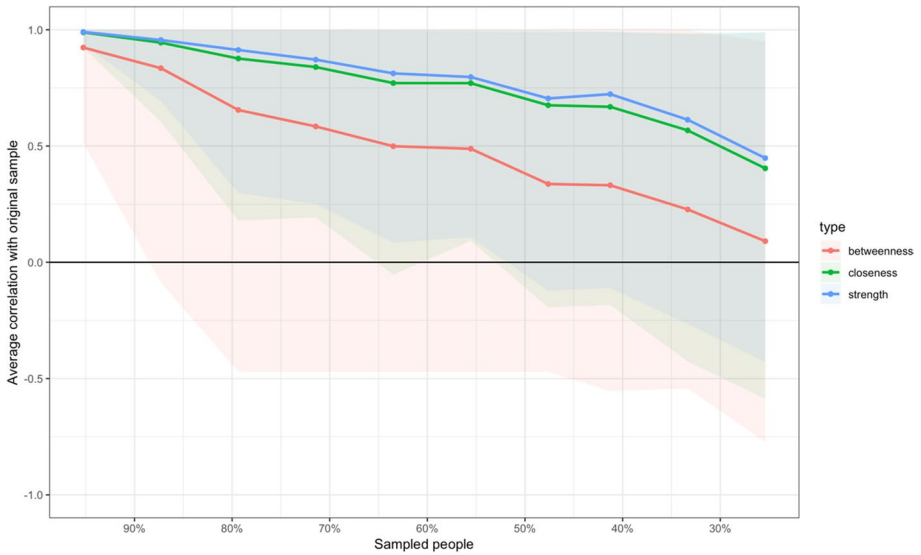**Fig. 5** Centrality plot of the Network-1. The plot is standardized where larger scores indicate greater centrality

Fig. 6 Average correlations between centrality indices sampled with a subset containing from 95 to 25% of the full sample and the original sample for Network-1. Lines show the means and areas show the range from the 2.5th quantile to the 97.5th quantile obtained through the bootstrap method
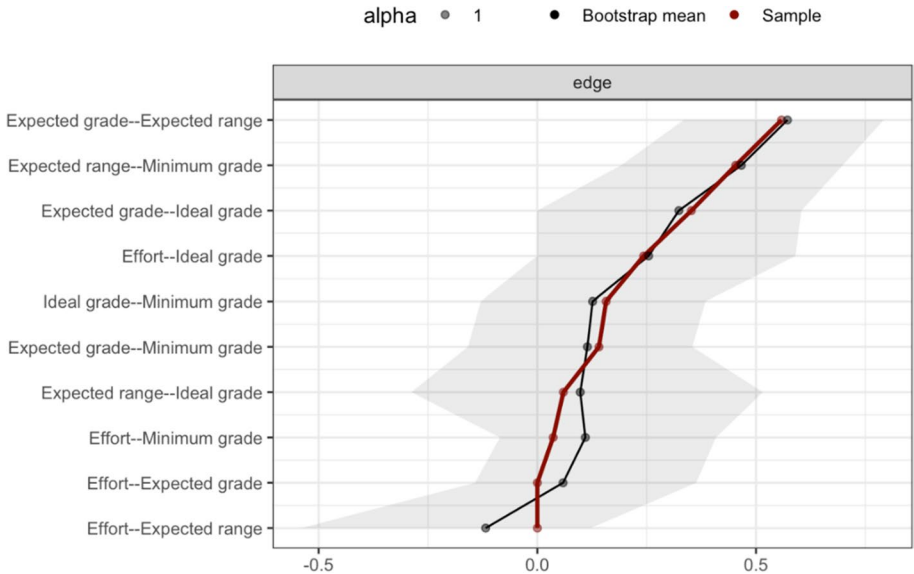
Fig. 7 Bootstrapped confidence intervals (CIs) of estimated edge-weights for the estimated network of 5 variables. The red line indicates the sample values and the gray area the bootstrapped Cis

variables and the student performance was studied. The students carried out a midterm exam near the middle of the semester (Previous Grade) and a final exam at the end of the semester (Actual Grade). A multiple regression analysis showed that the set of

explanatory variables (Ideal Grade, Minimum Grade, Effort, Expected Grade, and Previous Grade) formed an acceptable model ($F = 14.191$, $R^2 = 0.555$, $p < 0.001$), but only the Ideal Grade variable ($\beta = 0.549$, $t_{(58)} = 4.547$, $p < 0.001$), and the Previous Grade variable ($\beta = 0.452$, $t_{(58)} = 5.808$, $p < 0.001$) perform as good model predictors.

The network structure of Ideal Grade, Minimum Grade, Effort, Predicted Grade, Predicted Range, Actual grade and Previous grade is reported in Fig. 8 (Network-2).

Expected grade and Expected range had also the highest strength value (or degree) as in Network-1 (Fig. 9). Ideal grade had the highest closeness value showing a strong association to the nodes nearby, and it had also the highest betweenness value acting as the bridge between the communities of nodes.

Central stability is shown in Fig. 10 for Network-2. Strength had the highest stability; therefore, it is interpretable with some care, while betweenness and closeness are not.

The bootstrapped confidence intervals of estimated edge-weights are reported in Fig. 11 for Network-2. The edges Expected grade-Expected range, Expected range-Minimum grade, and Effort-Previous grade, and Ideal grade-Previous grade are reliably the five strongest edges since their bootstrapped CIs do not overlap with most other of the bootstrapped CIs. The network should be interpreted with caution due to the shape of CIs obtained for the edge weights.
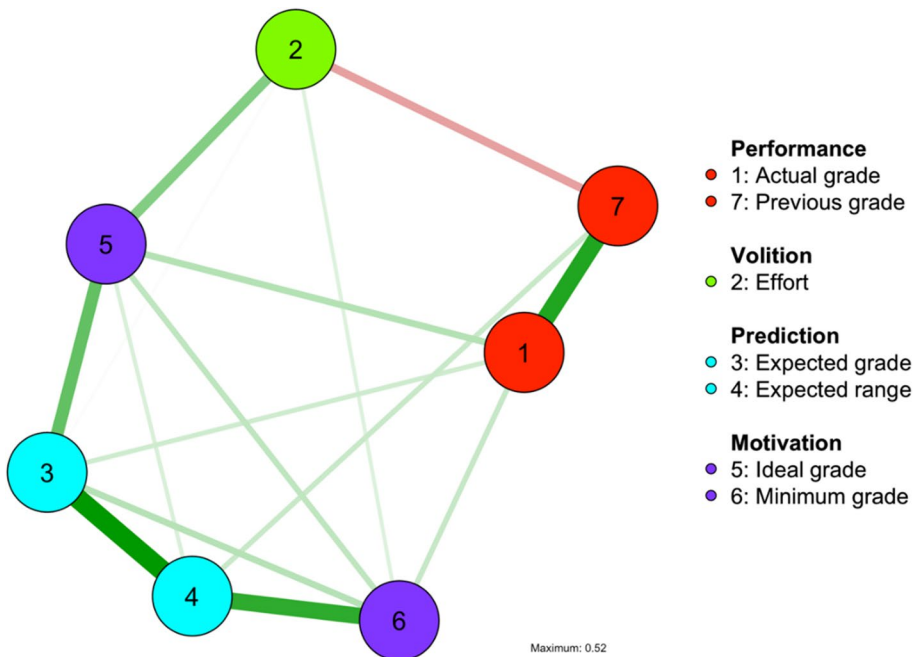

Fig. 8 Partial correlation network constructed using EBIC-glasso depicting the association between volitional and motivational variables, predictions, and academic performance (Network-2)
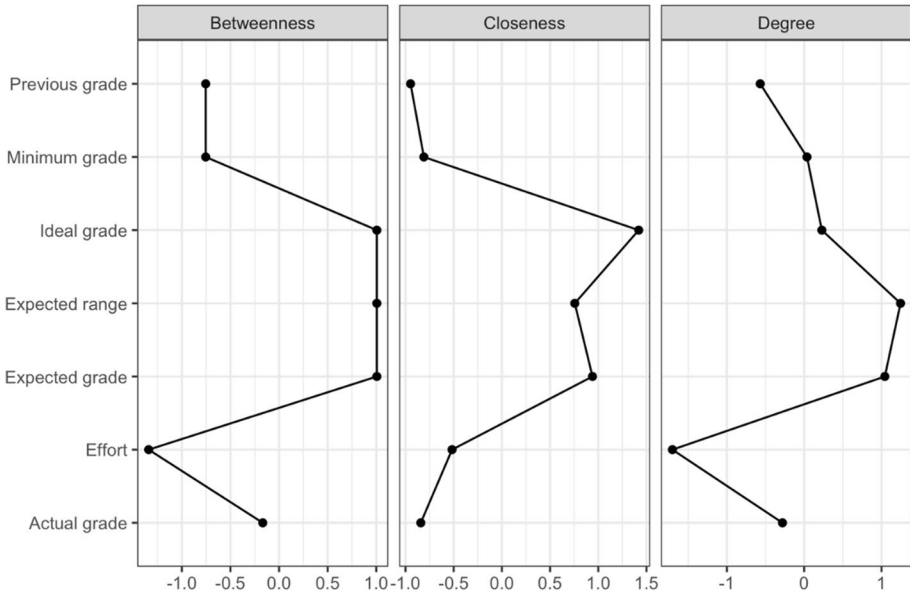
Fig. 9 Centrality plot of Network-2. The plot is standardized where larger scores indicate greater centrality
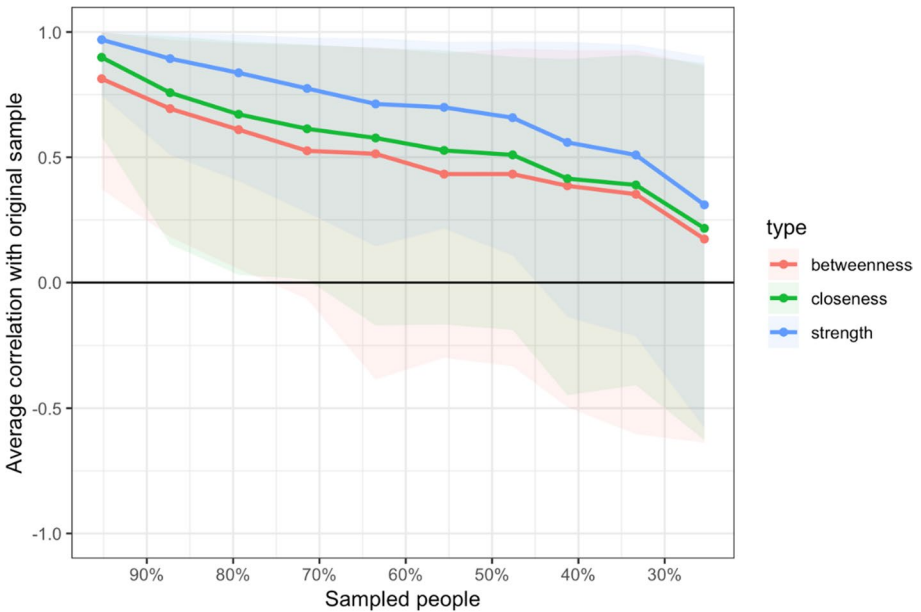

Fig. 10 Average correlations between centrality indices sampled with a subset containing from 95 to 25% of the full sample and the original sample for Network-2
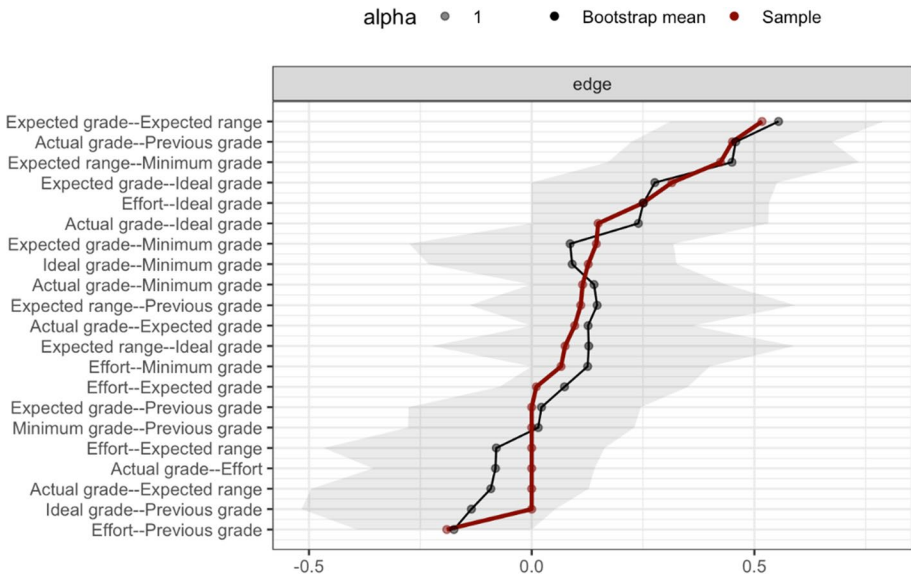
Fig. 11 Bootstrapped confidence intervals (CIs) of estimated edge-weights for the estimated network of 7 variables

Discussion and Conclusions

The objectives of the present study were, firstly, to analyze whether there was a cognitive bias between the ideal perception of knowledge and performance, and the real ones presented in the exam scores, and secondly, to determine whether the predictions of the students' performance were related to different motivational variables.

With respect to the first objective, the results showed that there is indeed a cognitive bias between the ideal perception of their performance and the actual performance. It is noted that lower performing students tend to overestimate their grades, while better performing students tend to underestimate their performance. This result coincides with previous work in which similar results have been obtained (de Bruin et al., 2017).

The results also show that the students' perception of their academic performance does not seem to be influenced by the feedback provided by professors (Previous Grade variable). Students who held illusions of competence on the midterm exam tended to do so throughout the end of semester, failing feedback to obtain a more accurate self-awareness of their low performance. That is, as Critcher and Dunning, (2009) point out, cognitive bias on performance is not affected by students' concrete experiences of similar exams, but by preconceived notions of their ability. There are different forms of intervention that can reduce this effect and that can serve as metacognitive training, such as home activities or group work, with the aim of achieving a better perception of one's own performance (de Bruin et al., 2017).

Network analysis found that performance variables had relatively few connections to motivation and prediction variables. However, performance variables are positively interconnected, not being affected by feedback. Students who overestimate their ability are not likely to change their study methods. The overestimation effect might obscure the student's perception about the effort required to pass the exam (Boekaerts & Rozendaal, 2010). This

connection can be observed in the negative association between volition and performance variables. The more effort students reported expending, lower the grade earned on the mid-term exam.

Regarding the second objective, volition is, however, positively related to motivational variables. Students report a high Ideal grade according to the effort expended. However, there is no connection between Effort and Minimum grade variable. Students who have studied more than they usually do have based their grade prediction on optimistic form of motivation.

Our findings show that there is a strong relationship between students' motivation (Ideal grade and Minimum grade), and the grades that students expect to obtain in the exam (Expected grade and Expected range). Thus, motivational variables play a more important role than academic variables in predicting future student performance. The network structures illustrate that the nodes that are specifically strongly paired are Ideal grade-Expected grade, and Minimum grade-Expected range. The optimistic form of motivation materialized in students' optimal level of performance (Ideal grade) is directly connected to their predicted grade. Their desires are directly connected with their metacognitive judgments. On the other hand, the pessimistic form of motivation expressed by students' lowest level of performance (Minimum grade) is connected to their perceived performance ranking. The minimum grade that students desire to earn is strongly related to student's ranking indicating their position among peers within the same class.

Additionally, multiple regression provides us with additional information about the Dunning-Kruger effect. The relationship between desired grades and expected grades has been corroborated. If we focused on students' ranking, Minimum Grade variable was a significant unique predictor for low-performing students (Q1), capturing the students' fear of failure and establishing the floor of students' desire. Those low-performing students may be content with just a passing grade. On the other hand, high-performing students (Q4) tend to focus on Ideal Grades, as a possible desire of self-improvement.

Recent works also suggests that individuals with stronger cognitive knowledge may simply be more cautious in their self-evaluations, providing lower judgements of their performance (de Carvalho Filho, 2009), or they may avoid appearing too competent for social reasons (Schunk & Pajares, 2004). On the other hand, students who underestimate their performance ability report tend to be more conformist (Bol et al., 2005).

However, this study presents limitations as the questionnaire had to have few questions because students had to fill it before the semester exam. Therefore, more information is needed to achieve a complete understanding of the results of this work, which precisely identifies the variables that influence both students' perception of their competences and their academic performance. Limitations of the study include a small sample size in which the gender distribution is not equal. There are recent studies that have looked at how gender influences the Dunning-Kruger effect. However, the results are inconclusive, as the results range from no gender difference (Pirttinen et al., 2020; Rachmatullah & Ha, 2019) to women having a greater effect than men (Harrington et al., 2018; Pazicni & Bauer, 2014). Future studies on the Dunning-Kruger effect may focus on gender differences and the influence of the mentioned variables, as well as on cognitive bias. On the other hand, students who participated in this study were not split into several streams. The influence of stream on the Dunning-Kruger effect has been scarcely studied in the literature (Bewes & Sharma, 2006; Muller & Sharma, 2012). The main results indicate that the students in advanced streams showed evidence of smaller bias or were better calibrated. Finally, it is necessary to explore the development of specific interventions that target the motivations of students, in order to be effective, and to reduce the gap between expected and actual grades.

Declarations

Conflict of Interest The authors declare no conflict of interest.

References

Bewes, J., & Sharma, M. D. (2006). A preliminary study on how accuracy relates to student self reports of confidence on a conceptual physics test. In *Proceedings of The Australian Conference on Science and Mathematics Education* (12th ed.). https://openjournals.library.sydney.edu.au/IISME/article/view/6377

Blackmore, C., Vitali, J., Ainscough, L., Langfield, T., & Colthorpe, K. (2021). A review of self-regulated learning and self-efficacy: The key to tertiary transition in science, technology, engineering and mathematics (STEM). *International Journal of Higher Education, 10*(3), 169–177.

Blanton, H., Pelham, B. W., DeHart, T., & Carvallo, M. (2001). Overconfidence as dissonance reduction. *Journal of Experimental Social Psychology, 37*(5), 373–385. https://doi.org/10.1006/jesp.2000.1458

Boekaerts, M., & Rozendaal, J. S. (2010). Using multiple calibration indices in order to capture the complex picture of what affects students' accuracy of feeling of confidence. *Learning and Instruction, 20*(5), 372–382. https://doi.org/10.1016/j.learninstruc.2009.03.002

Bol, L., Hacker, D. J., O'Shea, P., & Allen, D. (2005). The Influence of overt practice, achievement level, and explanatory style on calibration accuracy and performance. *The Journal of Experimental Education, 73*(4), 269–290. https://doi.org/10.3200/JEXE.73.4.269-290

Borsboom, D. (2017). A network theory of mental disorders. *World Psychiatry, 16*(1), 5–13. https://doi.org/10.1002/wps.20375

Buckman, W. G., Parks, J. E., & Coohill, T. P. (1975). An introductory physics and biophysics course for life science students. *American Journal of Physics, 43*(1), 77–80. https://doi.org/10.1119/1.10001

Buse, K., Hill, C., & Benson, K. (2017). Establishing the research agenda for increasing the representation of women in engineering and computing. *Frontiers in Psychology, 8*, 598. https://doi.org/10.3389/fpsyg.2017.00598

Caputo, D., & Dunning, D. (2005). What you don't know: The role played by errors of omission in imperfect self-assessments. *Journal of Experimental Social Psychology, 41*(5), 488–505. https://doi.org/10.1016/j.jesp.2004.09.006

Chambers, J. R., & Windschitl, P. D. (2004). Biases in social comparative judgments: The role of non-motivated factors in above-average and comparative-optimism effects. *Psychological Bulletin, 130*(5), 813–838. https://doi.org/10.1037/0033-2909.130.5.813

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika, 95*(3), 759–771. https://doi.org/10.1093/biomet/asn034

Critcher, C. R., & Dunning, D. (2009). How chronic self-views influence (and mislead) self-assessments of task performance: Self-views shape bottom-up experiences with the task. *Journal of Personality and Social Psychology, 97*(6), 931–945. https://doi.org/10.1037/a0017452

Crocker, J., Karpinski, A., Quinn, D. M., & Chase, S. K. (2003). When grades determine self-worth: Consequences of contingent self-worth for male and female engineering and psychology majors. *Journal of Personality and Social Psychology, 85*(3), 507–516. https://doi.org/10.1037/0022-3514.85.3.507

de Bruin, A. B. H., Kok, E. M., Lobbestael, J., & de Grip, A. (2017). The impact of an online tool for monitoring and regulating learning at university: Overconfidence, learning strategy, and personality. *Metacognition and Learning, 12*(1), 21–43. https://doi.org/10.1007/s11409-016-9159-5

de Carvalho Filho, M. K. (2009). Confidence judgments in real classroom settings: Monitoring performance in different types of tests. *International Journal of Psychology, 44*(2), 93–108. https://doi.org/10.1080/00207590701436744

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*(3), 69–106. https://doi.org/10.1111/j.1529-1006.2004.00018.x

Ehrlinger, J., & Dunning, D. (2003). How chronic self-views influence (and potentially mislead) estimates of performance. *Journal of Personality and Social Psychology, 84*(1), 5–17. https://doi.org/10.1037//0022-3514.84.1.5

Ehrlinger, J., Johnson, K., Banner, M., Dunning, D., & Kruger, J. (2008). Why the unskilled are unaware: Further explorations of (absent) self-insight among the incompetent. *Organizational Behavior and Human Decision Processes, 105*(1), 98–121. https://doi.org/10.1016/j.obhdp.2007.05.002

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods, 50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48*(4), 18.

Fakcharoenphol, W., Morphew, J. W., & Mestre, J. P. (2015). Judgments of physics problem difficulty among experts and novices. *Physical Review Special Topics - Physics Education Research, 11*(2), 020128. https://doi.org/10.1103/PhysRevSTPER.11.020128

Freund, P. A., & Kasten, N. (2012). How smart do you think you are? A meta-analysis on the validity of self-estimates of cognitive ability. *Psychological Bulletin, 138*(2), 296–321. https://doi.org/10.1037/a0026556

Friedman, J., Hastie, T., & Tibshirani, R. (2008). Sparse inverse covariance estimation with the graphical lasso. *Biostatistics, 9*(3), 432–441. https://doi.org/10.1093/biostatistics/kxm045

Galloway, R. K., Bates, S. P., Parker, J., & Usoskina, E. (2013). The effect of research-based instruction in introductory physics on a common cognitive bias. *AIP Conference Proceedings, 1513*(1), 138–141.

Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research, 25*, 186–192. https://doi.org/10.2307/3172650

Goldstein, R. (1993). Conditioning diagnostics: Collinearity and weak data in regression. *Technometrics, 35*(1), 85–86. https://doi.org/10.1080/00401706.1993.10484997

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson.

Harrington, B., Peng, S., Jin, X., & Khan, M. (2018). Gender, confidence, and mark prediction in CS examinations. In *Proceedings of the 23rd Annual ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE 2018)* (pp. 230–235). Association for Computing Machinery. https://doi.org/10.1145/3197091.3197116

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition and Learning, 4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8

Katz, M. H. (2003). Multivariable analysis: A primer for readers of medical research. *Annals of Internal Medicine, 138*(8), 644–650. https://doi.org/10.7326/0003-4819-138-8-200304150-00012

King, B. (2015). Changing college majors: Does it happen more in STEM and do grades matter? *Journal of College Science Teaching, 44*(3), 44–51. JSTOR.

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. *Journal of Personality and Social Psychology, 82*(2), 180–188. https://doi.org/10.1037/0022-3514.82.2.180

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77*(6), 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Kruger, J., & Dunning, D. (2002). Unskilled and unaware–but why? A reply to Krueger and Mueller (2002). *Journal of Personality & Social Psychology, 82*(2), 189–192. https://doi.org/10.1037//0022-3514.82.2.189

Labuhn, A. S., Zimmerman, B. J., & Hasselhorn, M. (2010). Enhancing students' self-regulation and mathematics performance: The influence of feedback and self-evaluative standards. *Metacognition and Learning, 5*(2), 173–194. https://doi.org/10.1007/s11409-010-9056-2

Lippert, H. P. (2020). A *mixed methods approach to understanding math anxiety among engineering students in calculus using their own words and by analysis of correlations among math anxiety, math attitude, and math competence (Order No. 28260392)*. Available from ProQuest Dissertations & Theses Global; Publicly Available Content Database (2469850847). https://www.proquest.com/dissertations-theses/mixed-methods-approach-understanding-math-anxiety/docview/2469850847/se-2

McNally, R. J. (2016). Can network analysis transform psychopathology? *Contributions from Experimental Psychopathology to the Understanding and Treatment of Mental Disorders, 86*, 95–104. https://doi.org/10.1016/j.brat.2016.06.006

Morphew, J. W. (2021). Changes in metacognitive monitoring accuracy in an introductory physics course. *Metacognition and Learning, 16*(1), 89–111. https://doi.org/10.1007/s11409-020-09239-3

Muller, D. A., & Sharma, M. D. (2012). Tackling misconceptions in introductory physics using multimedia presentations. In *Proceedings of The Australian Conference on Science and Mathematics Education* (13th ed.). https://openjournals.library.sydney.edu.au/index.php/IISME/article/view/6345

Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning, 1*(2), 159. https://doi.org/10.1007/s10409-006-9595-6

Pazicni, S., & Bauer, C. F. (2014). Characterizing illusions of competence in introductory chemistry students. *Chemistry Education Research and Practice, 15*(1), 24–34. https://doi.org/10.1039/C3RP00106G

Pintrich, P. R., & de Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*(1), 33–40. https://doi.org/10.1037/0022-0663.82.1.33

Pirttinen, N., Hellas, A., Haaranen, L., & Duran, R. (2020). Study major, gender, and confidence gap: Effects on experience, performance, and self-efficacy in introductory programming. In *2020 IEEE Frontiers in Education Conference (FIE)* (pp. 1–7). IEEE Press. https://doi.org/10.1109/FIE44824.2020.9273884

Rachmatullah, A., & Ha, M. (2019). Examining high-school students' overconfidence bias in biology exam: A focus on the effects of country and gender. *International Journal of Science Education, 41*(5), 652–673. https://doi.org/10.1080/09500693.2019.1578002

Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.). (1991). *Measures of personality and social psychological attitudes*. Academic Press.

Saenz, G. D., Geraci, L., Miller, T. M., & Tirso, R. (2017). Metacognition in the classroom: The association between students' exam predictions and their desired grades. *Consciousness and Cognition, 51*, 125–139. https://doi.org/10.1016/j.concog.2017.03.002

Schunk, D. H. & Pajares, F. (2004). Self-efficacy in education revisited: Empirical and applied evidence. In D. M. McInerney & S. V. Etten (Eds.), *Big theories revisited* (pp. 115–138). Information Age Publishing.

Seidel, L., Benito, R. M., Camara, M. E., Losada, J. C., & Arranz, F. J. (2008). Development of an interactive learning environment for introductory physics in engineering. *WSEAS Transactions on Advances in Engineering Education, 5*(5), 306–315.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and Mcdonald's ωH: Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*(1), 123–133. https://doi.org/10.1007/s11336-003-0974-7