Research Article

Predicting Mathematical Performance: The Effect of Cognitive Processes and Self-Regulation Factors

Mariel Musso,[1, 2, 3, 4] Eva Kyndt,[1] Eduardo Cascallar,[1, 3, 4] and Filip Dochy[1]

[1] *Centre for Research on Teaching and Training, Katholieke Universiteit Leuven, Belgium*
[2] *Universidad Argentina de la Empresa, Buenos Aires, Argentina*
[3] *Assessment Group International, Washington, DC, USA*
[4] *Assessment Group International, Brussels, Belgium*

Correspondence should be addressed to Eduardo Cascallar, agi_group@msn.com

A substantial number of research studies have investigated the separate influence of working memory, attention, motivation, and learning strategies on mathematical performance and self-regulation in general. There is still little understanding of their impact on performance when taken together, understanding their interactions, and how much each of them contributes to the prediction of mathematical performance. With the emergence of new methodologies and technologies, such as the modelling with predictive systems, it is now possible to study these effects with approaches which use a wide range of data, including student characteristics, to estimate future performance without the need of traditional testing (Boekaerts and Cascallar, 2006). This research examines the different cognitive patterns and complex relations between cognitive variables, motivation, and background variables associated with different levels of mathematical performance using artificial neural networks (ANNs). A sample of 800 entering university students was used to develop three ANN models to identify the expected future level of performance in a mathematics test. These ANN models achieved high degree of precision in the correct classification of future levels of performance, showing differences in the pattern of relative predictive weight amongst those variables. The impact on educational quality, improvement, and accountability is highlighted.

1. Introduction

Although there is substantial research which has investigated the influences of (a) working memory [1–11]; (b) attentional systems [12, 13], and (c) motivation [14–20], on mathematical performance and self-regulation in general, these studies have looked at the separate effects of these components. Therefore, we have little understanding and data about how they impact performance when taken together, understanding their interactions, and how much each can predict the mathematical performance in an integrated model [21, 22]. New methodologies and technologies, and the emergence of predictive systems, have focused on the possibility of assessments which use a wide range of data or student productions to estimate future student performance without the need for traditional testing [23].

Artificial neural networks (ANNs) have been used in several different fields of research and in applied environments, such as biology [24], business [25], finance [26], medicine [27], defense [28], meteorology and environmental studies [29], and in the prediction of terrorist attacks [30]. During the last few decades, ANNs have been increasingly utilized as a statistical methodology in applied areas such as classification and recognition of patterns in business, finance, and the social sciences [25, 31–35].

However, the literature shows very few studies applying neural networks in education and in educational assessment in particular [36, 37]. Some authors have suggested that traditional statistical methods do not always yield accurate predictions [38]. ANNs may improve the validity and the accuracy of the predictions and/or classifications and also improve the predictive validity of test scores [36, 39, 40].

The purpose of this research was to develop predictive classification models that could identify with sufficient precision three groups of students corresponding to the highest 30%, lowest 30%, and middle 30% of estimated future

performance in a mathematics test, utilizing only cognitive, motivational, and background variables, with no consideration of the mathematics content present in the test or of any measure of previous mathematics performance. Finally, in order to compare the predictive power of this ANN-based approach with more classical statistical methods, discriminant analyses were used.

It was expected that results would enable the development of an "early warning" system which could allow early and prompt intervention with those students most in need of support and remediation in mathematics (at the level of exit from secondary education and/or at the beginning of university studies). Similarly, this approach could serve to identify top or advanced students and improve their placement and/or career choice.

2. Theoretical Background

2.1. Working Memory Capacity and Mathematical Performance. A large body of literature shows working memory as a very important construct in several areas, and several studies have shown its important role in a wide range of complex cognitive behaviors, such as comprehension, reasoning, and problem solving [41]. Working memory (WM) is an important predictive variable of intellectual ability and academic performance, consistent over time [1–11, 21, 22, 42, 43]. Nevertheless, it is still not understood precisely how this basic cognitive mechanism influences specific performance and how it is related to performance levels in particular areas, as is the case with mathematical performance.

Working memory capacity refers to the temporary representation of information that was just experienced or just retrieved from long-term memory but no longer exists in the external environment, and it will be operationalized by the overall measure of the automated operation span [44]. Internal representations are short-lived, but can be maintained for longer periods of time through active rehearsal strategies, and can be subjected to operations that manipulate the information in such a way that it becomes useful for goal-directed behavior. The term working memory is applied to a system of limited capacity, which is capable of storing and handling information necessary for the performance of complex tasks such as learning, comprehension, and reasoning [45]. There are several paradigms to examine the role of working memory in complex tasks. One important approach examines it from the perspective of individual differences, using various working memory span tasks as a research tool [46–49]. These span tasks (reading, operation, and spatial spans) are designed to resemble the working memory demands during the performance of complex cognitive tasks by placing simultaneous demands on both processing and storage. Individual differences in WM influence the performance in complex tasks [50, 51]. It is possible to assume that task complexity has an influence on performance because the increase of complexity demands a greater level of activation for retrieval of information from declarative memory [50].

Mathematical cognition involves complex mechanisms or processes such as identification of relevant quantities, encoding into an internal representation, mental comparisons, and calculations [52]. These cognitive activities are encompassed by working memory. Despite the level of agreement regarding the close relationship between working memory and mathematics processing and learning, further studies on the role of working memory in mathematical cognition are necessary to better understand the participation of task and subject characteristics in the modulation of performance in mathematical processing and learning [1, 53].

There is some supportive but not extensive literature on the critical role of working memory in mathematical performance [54]. Working memory is related to a variety of numerical and mathematical abilities used for counting, which underlie the solution of simple addition and subtraction problems [2–4, 6–8, 11] in [1, 5, 9, 10, 55], as well as the solution of complex arithmetic problems [50, 51]. From a perspective of individual differences, and specifically from the perspective of "math disabilities," results of various studies indicate that children with poor mathematics ability also show low working memory [56–62]. These children have difficulty in holding information in memory while performing another activity (e.g., counting) and in inhibiting irrelevant information [7, 63]. Furthermore, they perform poorly in measures of the visuospatial working memory and the central executive [61, 64–67]. However, they do not show particular problems in phonological working memory tasks [60, 67].

One of the most recent working memory's approaches develops computational models that simulate the effects of individual differences and/or working memory load on participants' performance on various cognitive tasks. Interesting areas of this approach include the model of mental algebra [50] and reasoning and problem solving [68].

2.2. Attention and Mathematical Performance. In cognitive models, attention has been traditionally involved in the control of intended actions. In this sense, attentional control has been identified as an important domain in self-regulation [12, 13].

Specifically, attention problems have been related to mathematical performance. Inattention is considered as a risk factor for poor math achievement [69]. Some research has found that a deficit in sustained attention is correlated with problems in mathematics (from 26% to 60% [70]). Other studies have found a low predictive power of attention considering it together with depressive symptoms and anxiety [71]. However, most of those studies have used classical multiple regression analysis to predict mathematical performance.

Current research findings suggest that attention involves different mechanisms which involve separate brain areas. In particular, attention encompasses three subsystems: (a) orienting, (b) alerting, and (c) executive control. The orienting network allows the selection of information from sensory input, the alerting network refers to a system that achieves and maintains an alert state, and executive control is responsible for resolving conflict among responses [72]. Executive

control has been closely related to working memory capacity [73].

2.3. Self-Regulation of Learning and Mathematical Performance.

Previous research on self-regulated learning focuses primarily on the learning strategies that students need to use in order to guide their learning [23]. Motivation researchers (e.g., [14–16]) argued that students also need their motivation for learning and effort investment, motivation strategies [17, 18] and volitional strategies [74, 75] to comply with social rules and expectations and deal with obstacles en route to the learning goal. Therefore, this study analyzes both learning strategies as motivational components of self-regulation. Learning strategies (LSs) involve any thoughts or behaviors that help the students to acquire new information and integrate these in their existing knowledge [76–79]. LS also help students retrieve stored information. Examples of LS include summarizing, paraphrasing, imaging, creating analogies, note taking, and outlining [77].

Motivational self-regulation includes motivational beliefs, motivation strategies, and motivational regulatory strategies. Motivational beliefs involve (a) values that students attach to a particular domain, (b) the students' opinion of the efficiency and effectiveness of learning and teaching methods, (c) beliefs about internal control transformed into self-efficacy beliefs (opinions that students hold about their own ability in relation to a specific domain), (d) outcome expectations: beliefs about the success or failure of specific actions, (e) goal orientation: orientation to learning tasks versus egoorientation: the intention is to demonstrate success (approach ego orientation) or to hide failure (avoidance ego-orientation), and (f) effort beliefs. Domain-specific self-efficacy beliefs influence effort investment, and not the other way round [80].

Research shows that epistemic and motivational beliefs that students hold play an important role in self-regulation [16, 81, 82]. Shoenfeld [83] recognized the existence of a system of beliefs that influences the mathematics problem-solving behavior. Several studies have identified beliefs about mathematics as a discipline, about the learning of mathematics, and about the social context of mathematical activities [19, 20].

Other categories of beliefs have been identified about the self in relation to mathematical learning: achievement goal orientation [84], responsibility beliefs for failure and success [85], self-efficacy beliefs [86–88], and value beliefs [88].

2.4. Neural Networks and Performance.

Conceptually, a neural network is a computational structure consisting of several highly interconnected computational elements, known as neurons, perceptrons, or nodes. Each neuron carries out a very simple operation on its inputs and transfers the output to a subsequent node or nodes in the network topology [89]. Neural networks exhibit polymorphism in structure and parallelism in computation [90], and it can be construed as a highly connected structure of processing elements that attempts to mimic the parallel computation ability of the biological brain [91–94].

Predictive streams analyses [21], based in this case on neural network (ANN) models, have several strengths: (a) because these are machine learning algorithms, the assumptions required for traditional statistical predictive models (e.g., ordinary least squares regression) are not necessary. As such, this technique is able to model nonlinear and complex relationships among variables. ANNs aim to maximize classification accuracy and work through the data in an interactive process until maximum accuracy is achieved, automatically modeling all interactions among variables; (b) ANNs are robust, general function estimators. They usually perform prediction tasks at least as well as other techniques and sometimes perform significantly better [95]; (c) ANNs can handle data of all levels of measurement, continuous or categorical, as inputs and outputs. Because of the speed of microprocessors in even basic computers, ANNs are more accessible today than they were when originally developed.

The ANN learns by examining individual training case, then generating a prediction for each testing case, and making adjustments to the weights whenever it makes an incorrect prediction. Information is passed back through the network in iterations, gradually changing the weights. As training progresses, the network becomes increasingly accurate in replicating the known outcomes. This process is repeated many times, and the network continues to improve its predictions until one or more of the stopping criteria have been met. A minimum level of accuracy can be set as the stopping criterion, although additional stopping criteria may be used as well (e.g., number of iteration and amount of time). Once trained, the network can be applied to future cases (validation or holdout sample) for validation and implementation [96].

2.5. Measures to Evaluate the Neural Network System Performance.

In order to evaluate the performance of the neural network system, there are a number of measures used which provide a means of determining the quality of the solutions offered by the various network models tried. The traditional measures include the determination of actual numbers and rates for true positive (TP), true negative (TN), false positive (FP), and false negative (FN) outcomes, as products of the ANN analysis. In addition, certain summative evaluative algorithms have been developed in this field of work, to assess overall quality of the predictive system.

These overall measures are Recall, which represents the proportion of correctly identified targets, out of all targets presented in the set, and is represented as Recall = P/(TP + FN); Precision which represents the proportion of correctly identified targets, out of all identified targets by the system, and is represented as Precision = TP/(TP + FP). Two other measures have been used to report the characteristics of the detection sensitivity of the system. One of them is Sensitivity (similar to Recall: the proportion of correctly identified targets, out of all targets presented in the set), which is expressed as Sensitivity = TP/(TP + FN). The other is Specificity, defined as the proportion of correctly rejected targets from all the targets that should have been rejected by the system and which is expressed as Specificity = TN/(TN + FP). These measures are typically represented in what is called

Table 1: Testing phase of the neural network predicting lowest 30% math scores.

| | | Prediction of performance | |
		∼30% lowest	30% lowest
Observed performance	∼30% lowest	71.40%	28.60%
	30% lowest	0%	100%

a "confusion matrix" representing all four outcomes (see Table 1).

In addition, the evaluation of ANN performance is carried out with a summative measure, which is used to account for the somewhat complementary relationship between Precision and Recall. This measure is defined as F1 = (2 ∗ Precision ∗ Recall)/(Precision + Recall). Such a definitional expression of F1 assumes equal weights for Precision and Recall. This assumption can be modified to favor either Precision or Recall, according to the utility and cost/benefit ratio of outcomes favoring either Precision or Recall for any given predictive circumstance.

3. Method

The sample included 800 university students, of both genders, ages between 18 and 25, enrolled in the first year in several different disciplines (psychology, engineering, medicine, law, social communication, business, and marketing), in three universities, during the 2009-2010 academic year.

3.1. Instruments

3.1.1. Attention Network Test (ANT) [72]. This task provides a measure for each of the three anatomically defined attentional networks: alerting, orienting, and executive. Participants are asked to determine when a central arrow points left or right. The ANT's responses were collected via two mouse buttons (left-right). They were instructed to focus on a centrally located fixation cross throughout the task and to respond as quickly and accurately as possible. During the practice trials, but not during the experimental trials, subjects received feedback from the computer on their speed and accuracy. The practice trials took approximately 2 minutes, and each of the three experimental blocks was approximately 5 minutes long. The whole experiment took about twenty minutes. The measure for (general) attention is the average response time regardless of the cues or flankers. To analyse the effect of the three attentional networks, a set of cognitive subtractions described by Fan et al. [72] were used. The efficiency of the three attentional networks is assessed by measuring how response times are influenced by alerting cues, spatial cues, and flankers [72]. The alerting effect was calculated by subtracting the mean response time of the double-cue conditions from the mean response time of the no-cue conditions. For the orienting effect, the mean response time of the spatial cue conditions (up and down) was subtracted from the mean response time of the centre cue condition. Finally, the effect of the executive control (conflict effect) was calculated by subtracting the

mean response time of all congruent flanking conditions, summed across cue types, from the mean response time of incongruent flanking conditions [72].

3.1.2. Automated Operation Span [44]. This is a computer-administered version of the Ospan instrument [44] that measures working memory capacity. The responses were collected via click of a mouse button. This study reported Absolute Ospan score that is interpreted as the measure of overall working memory capacity and one Reaction Time's score (operations). The task took approximately 20–25 minutes to complete [44].

3.1.3. Learning Strategies Questionnaire [77–79]. A validated Spanish version was administered. It is a 77-item questionnaire with 10 scales that assesses the students' awareness about, and use of, learning and study strategies related to skill, will, and self-regulation components of strategic learning. The Attitude Scale assesses students' attitudes and interest in college and academic success. It examines how facilitative or debilitative their approach to college and academics is for helping them get their work done and for succeeding in college (sample item: I feel confused and undecided as to what my educational goals should be). The Motivation Scale assesses students' diligence, self-discipline, and willingness to exert the effort necessary to successfully complete academic requirements (sample item: When work is difficult I either give up or study only the easy parts). The Time Management Scale assesses students' application of time management principles to academic situations (sample item: I only study when there is the pressure of a test). The Anxiety Scale assesses the degree to which students worry about school and their academic performance. Students who score low on this scale are experiencing high levels of anxiety associated with school (note that this scale is reverse scored). The Concentration Scale assesses students' ability to direct and maintain attention on academic tasks (sample item: I find that during lectures I think of other things and do not really listen to what is being said). The Information Processing Scale assesses how well students can use imagery, verbal elaboration, organization strategies, and reasoning skills as learning strategies to help build bridges between what they already know and what they are trying to learn and remember, that is, knowledge acquisition, retention, and future application (sample item: I translate what I am studying into my own words). The Selecting Main Ideas Scale assesses students' skill at identifying important information for further study from among less important information and supporting details (sample item: Often when studying I seem to get lost in details and cannot see the forest for the trees). The Study Aids Scale assesses students' use of supports or resources to help them learn or retain information (sample item: I use special helps, such as italics and headings that are in my textbooks). The Self-Testing Scale assesses students' use of reviewing and comprehension monitoring techniques to determine their level of understanding of the information to be learned (sample item: I stop periodically while reading and mentally go over or review what was said). The Test Strategies Scale assesses students' use of test

preparation and test taking strategies (sample item: In taking tests, writing themes, etc., I find I have misunderstood what is wanted and lose points because of it). Items were scored on a 5-point Likert scale ranging from "Always" to "Never."

3.1.4. Online Motivation Questionnaire (OMQ) [97]. The last version of the On-Line Motivation Questionnaire, namely, the OMQ91 [97], was used to study motivational variables. This is a self-report questionnaire that is applied to a specific task (in this study, a mathematical task), and it consists of two parts (Part 1: before the task; Part 2: after the task). This study reports the Part 1 (appraisal part) that included 23 items that measure three aspects of task motivation: appraisals, emotions, and learning intention. Appraisals are registered by 13 items that measure three aspects of task judgment: personal relevance of the curricular task (e.g., How useful do you consider this task), subjective competence (e.g., How good are you at doing this type of task?), and task attraction (e.g., How much do you like this type of task?). Six items refer to emotional state (e.g., How do you feel now? Nervous-not nervous; happy-not happy). Four items measure learning intention (e.g., How much effort are you going to invest in the task?). Students were asked to answer the questions focusing on a multiple choice mathematics task. Items were scored on a 4-point Likert scale ranging (e.g., not at all well, not so well, well, very well).

3.1.5. Mathematics Test. This test consisted of 65 multiple choice items with four or five options and only one correct answer (50 items were taken from a national test [98], and 15 items were extracted from disclosed items of the Trends in International Mathematics and Science Study [99]. The local calibration for the test was done applying a 3-parameter Item Response Theory model (IRT model). The equating to the TIMSS results was done using a 1-parameter IRT model, centred on ability, because the TIMSS statistical data uses only one parameter for the analysis and reporting. The items measure simple algorithms for arithmetic problems: some items required the use of percentages or proportions, decimal numbers, and a few others are algebraic and geometric questions. There was no time limit to take the test, but its duration for all students was under two hours. In order to guarantee that it was not a speeded test (which would violate the IRT assumptions), the allotted time was such that more than 95% of the students could complete the test by at least attempting a response to all the questions; this number was actually much higher, closer to 99% of the students.

In addition, a questionnaire was administered in order to collect background variables: gender, highest level of education of mother and father (i.e., did not complete mandatory primary school, completed primary school, completed secondary school, completed undergraduate university studies, completed postgraduate studies), occupation of parents, and secondary school from which the student graduated (i.e., public, private religious school, private nonreligious school, bilingual school, foreign community school).

3.2. Analyses Procedure. The ANN model used was a backpropagation multilayer perceptron neural network, that is, a multilayer network composed of nonlinear units, which computes its activation level by summing all the weighted activations it receives and which then transforms its activation into a response via a nonlinear transfer function. During their training phase, these systems evaluate the effect of the weight patterns on the precision of their classification of outputs, and then, through backpropagation, they adjust those weights in a recursive fashion until they maximize the precision of the resulting classifications. A predictive classification architecture based on neural networks (ANNs) model development was developed for each targeted future mathematical performance group: lowest 30%, middle 30%, and highest 30% of student performance groups. ANN parameters and variable groupings, as well as all other network architecture parameters, were manipulated to maximize predictive precision and total accuracy. Confusion matrices have been determined for each ANN, as well as receiver operating characteristic (ROC) curves to determine the discrimination level of the model. ROC analyses provide a very useful measure to establish the performance of the classifier at various levels of true positive and true negative rates, using sensitivity and specificity values. Parameters such as learning rate, momentum, number of hidden layers, stopping rules, transfer functions, and number of nodes were specified and manipulated in the model construction phase in order to maximize the overall performance of the models.

3.3. Architecture of the Neural Networks. Three different neural networks (ANNs) were developed as predictive systems for the mathematics task of this study. ANN1 was developed to maximize the predictive classification of the lowest 30% of students, which would be scoring the lowest in the mathematics test. ANN2 was developed to maximize the predictive classification of the highest 30% of students, which would be scoring the highest in the mathematics test. ANN3 was developed to predict the middle 30% of students, which would be scoring in the middle level of performance in the mathematics test. The specific architecture of each of the three neural networks developed is as follows.

ANN1 (low 30%): all cognitive, motivational, and background variables were introduced in the analysis. They were used for the development of the vector-matrix containing all predictor variables for each student. The resulting network contained all the input predictors; some of them collapsed into subscales to maximize predictive classification, with a total of 36 input units. The model built contained one hidden layer, with 8 units. The output layer contained two units (categories corresponding to "belongs to lowest 30%" or "belongs to highest 70%"). A standardized method for the rescaling of covariates was used. The hidden layers had hyperbolic tangent activation functions, which is the most common activation function used for neural networks because of its greater numeric range (from −1 to 1) and the shape of its graph. For the output layer, the activation function chosen was identity, and the error function the sum of squares.

ANN2 (high 30%): all cognitive, motivational, and background variables were introduced in the analysis. They were used for the development of the vector-matrix containing all

predictor variables for each student. The resulting network contained all the input predictors; some of them collapsed into subscales to maximize predictive classification, with a total of 36 input units. The model built contained two hidden layers, with 8 and 6 units, respectively, and an output layer with two units (categories corresponding to "belongs to highest 30%" or "belongs to lowest 70%"). A standardized method for the rescaling of covariates was used. The hidden layer and output layer had a hyperbolic tangent activation function, and the error function the sum of squares.

ANN3 (middle 30%): all cognitive, motivational, and background variables were introduced in the analysis. They were used for the development of the vector-matrix containing all predictor variables for each student. The resulting network contained all the input predictors; some of them collapsed into subscales to maximize predictive classification, with a total of 36 input units. The model built contained one hidden layer with 1 unit and one output layer with two units (categories corresponding to "belongs to middle 30%" or "belongs to extreme 30%'s"). A standardized method for the rescaling of covariates was used. The hidden layer had a hyperbolic tangent activation function, and the output layer applied a softmax activation function.

The software used was SPSS v.19, Neural Network Module, for the development and analysis of all predictive models in this study. The usual three development phases of the predictive system were carried out: training of the network, testing of the network developed, and validation of the network. During the training phase several models were attempted, and several modifications of the neural network parameters were tried, manipulating learning persistence, learning rate, momentum, and other criteria. These tests continued until achieving desired levels of classification, maximizing the benefits of the model chosen. In this analysis both precision and recall, as outcome measures of the network, were given equal weight. There was no need to trim the number of predictor inputs in the three models.

3.4. Discriminant Analyses. Discriminant analyses (DA) were carried out using the same data and the same categories of mathematical performance used in the neural networks analyses. The DA1 was performed to discriminate between the students belonging to the lowest 30% of mathematical performance and those not in that category. The DA2 has been focused on identifying students in the highest 30% versus those not in that group, and the DA3 was calculated to discriminate the students belonging to the middle 30% and those not in that category. In order to give every variable the opportunity to contribute significantly to the prediction, a stepwise discriminant analysis was calculated for each category including all independent variables. In addition, we calculated three discriminant analyses, one for each category including the independent variables of the maximised neural networks of each category.

4. Results

The ANN1 was able to reach 100% correct identification of all students that belong to the target group (lowest 30%) in

TABLE 2: Relative importance of the top variables participating in the model for the predictive classification of the lowest 30% of scores in the mathematics test.

Independent variable importance, low 30% group		
	Importance	Normalized importance
Gender	.035	34.2%
Mother's educational level	.028	28.2%
Father's educational level	.024	23.9%
Mother's occupation	.065	64.5%
Father's occupation	.059	58.8%
Age	.062	61.5%
Competence-related attribution for success	.041	40.3%
Personal relevance of task	.029	28.3%
Subjective competence	.043	42.7%
Task attraction	.042	41.8%
Learning intention	.052	51.8%
Reported effort	.062	61.1%
Expected result of assessment	.099	97.7%
Emotional state	.062	61.3%
Alerting attention	.029	29.1%
Orienting attention	.018	17.4%
Executive attention	.067	66.4%
Working memory	.081	80.6%
Reaction time (operations)	.101	100.0%

TABLE 3: Testing phase of the neural network predicting highest 30% math scores.

		Prediction of performance	
		~30% highest	30% highest
Observed performance	~30% highest	66.70%	33.30%
	30% highest	0%	100%

both the training and testing phase. The precision of NN1 equalled .75 on a maximum of 1 (see Table 1). The sensitivity of the network equalled 1, and the specificity was equal to .714 from a maximum of 1. The area under the curve equalled .953 from a maximum of 1.

Table 2 shows the actual predictive weight of each input, and Figure 1 shows the normalised importance of the different variables for the ANN1 predictive classification. These results indicate that reaction time (operations), expected result of assessment, and working memory were the most important predictors.

The ANN2 reached an accuracy of 90% and 100% in the training and testing phase, respectively. The precision of ANN2 equalled .80 from a maximum of 1 (see Table 3). The sensitivity of the network equalled 1, and the specificity amounted to .667 from a maximum of 1. The area under the curve equalled .958 from a maximum of 1.

Figure 2 shows the normalised importance of the different variables for the ANN2 predictive classification. The most

Normalized importance

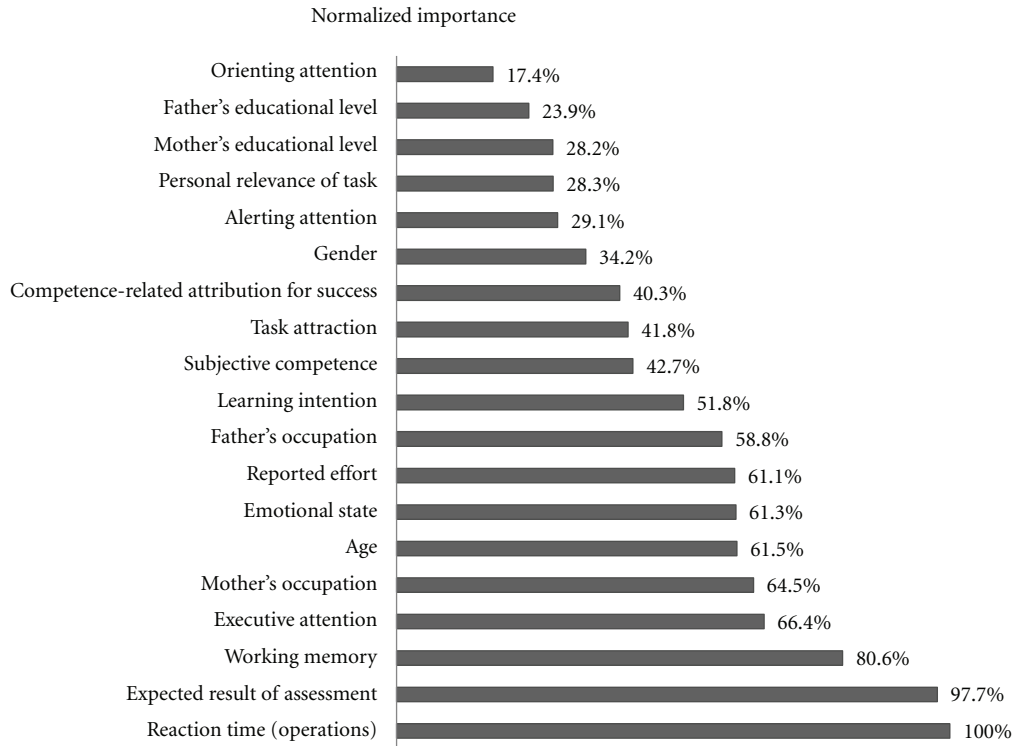| Variable	Value
Orienting attention	17.4%
Father's educational level	23.9%
Mother's educational level	28.2%
Personal relevance of task	28.3%
Alerting attention	29.1%
Gender	34.2%
Competence-related attribution for success	40.3%
Task attraction	41.8%
Subjective competence	42.7%
Learning intention	51.8%
Father's occupation	58.8%
Reported effort	61.1%
Emotional state	61.3%
Age	61.5%
Mother's occupation	64.5%
Executive attention	66.4%
Working memory	80.6%
Expected result of assessment	97.7%
Reaction time (operations)	100%

FIGURE 1: Normalized importance of the top variables participating in the model for the predictive classification of the lowest 30% of scores in the mathematics test.

important variables for the prediction of ANN2 (high 30%) were the task attraction and father's occupation.

Both networks showed interesting differences in the pattern of relative normalized importance of those variables with the highest participation in the predictive model. For the low performers (those predicted to be in the lowest 30% of scores), several basic cognitive variables were most important in attaining a correct classification, such as "reaction-time," "working memory capacity," and the closely related "executive attention," all having to do with the control and the speed of processing. In fact, three out of the top four variables in terms of relative predictive importance correspond to basic cognitive processing variables, with high relative values. Among the self-regulation variables, only "expected results of the assessment" appeared among the most predictive.

On the other hand, results from the predictive model for those expected to be in the highest 30% of the scores, the top three predictors with the most significant participation were "task attraction," "father's occupation," and "reported effort," all among the self-regulation and background variables. Only "working memory capacity" (as measured by "absolute AOSPAN") among the basic cognitive processing variables appeared among the top five predictors and then with a much lower relative importance than for the low 30% group. It is quite evident the relative lower importance of all cognitive control and speed of processing variables, which are not discriminated well for the predictive classification in the highest 30% group. It is also worth noting the relative high

TABLE 4: Testing phase of the neural network predicting for middle 30% math scores.

		Prediction of performance	
		~30% middle	30% middle
Observed performance	~30% middle	67.6%	32.4%
	30% middle	29.4%	70.6%

importance of parents' occupation in both low and high groups, particularly in the first neural network.

4.1. ANN3 (Middle 30%). The ANN3 showed an accuracy of 74.5% and 70.6% in the training and testing phase, respectively. The precision of ANN3 equalled .70 from a maximum of 1 (see Table 4). The sensitivity of the network equalled .70, and the specificity amounted to .67 from a maximum of 1. The area under the curve equalled .725 from a maximum of 1.

The most important variables for the prediction of ANN3 (middle 30%) were positive learning strategies and study techniques, reaction time (natural logarithm) of attentional networks, time management, and subjective competence (see Figure 3).

4.2. Discriminant Analyses. DA1 focused on the lowest 30% of the students and the rest. One of the restrictions of this analyses refers to the assumption of equality of

Normalized importance

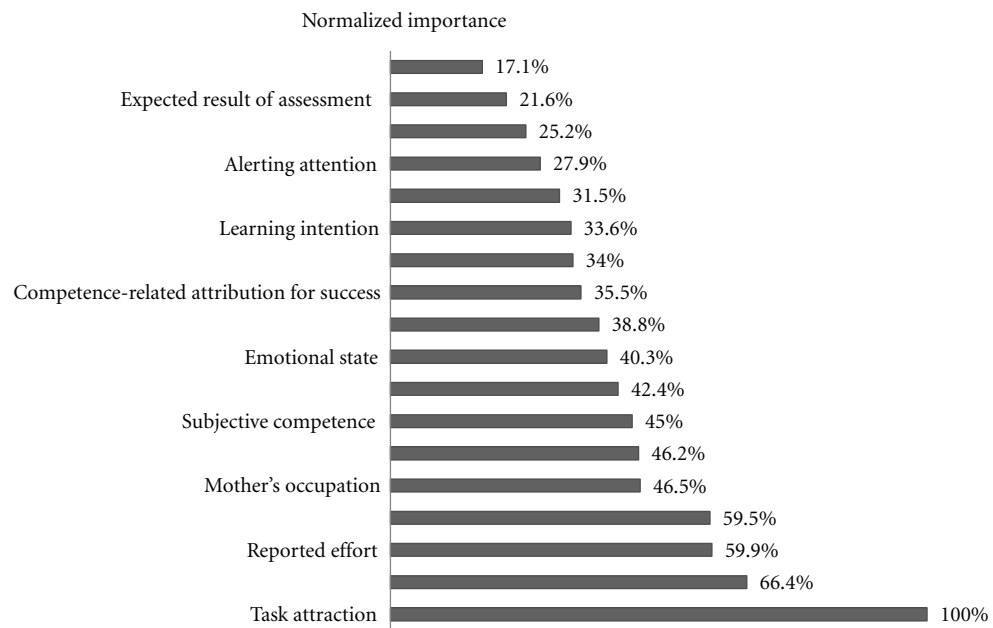|	
	17.1%
Expected result of assessment	21.6%
	25.2%
Alerting attention	27.9%
	31.5%
Learning intention	33.6%
	34%
Competence-related attribution for success	35.5%
	38.8%
Emotional state	40.3%
	42.4%
Subjective competence	45%
	46.2%
Mother's occupation	46.5%
	59.5%
Reported effort	59.9%
	66.4%
Task attraction	100%

FIGURE 2: Normalized importance of the top variables participating in the model for the predictive classification of the highest 30% of scores in the mathematics test.
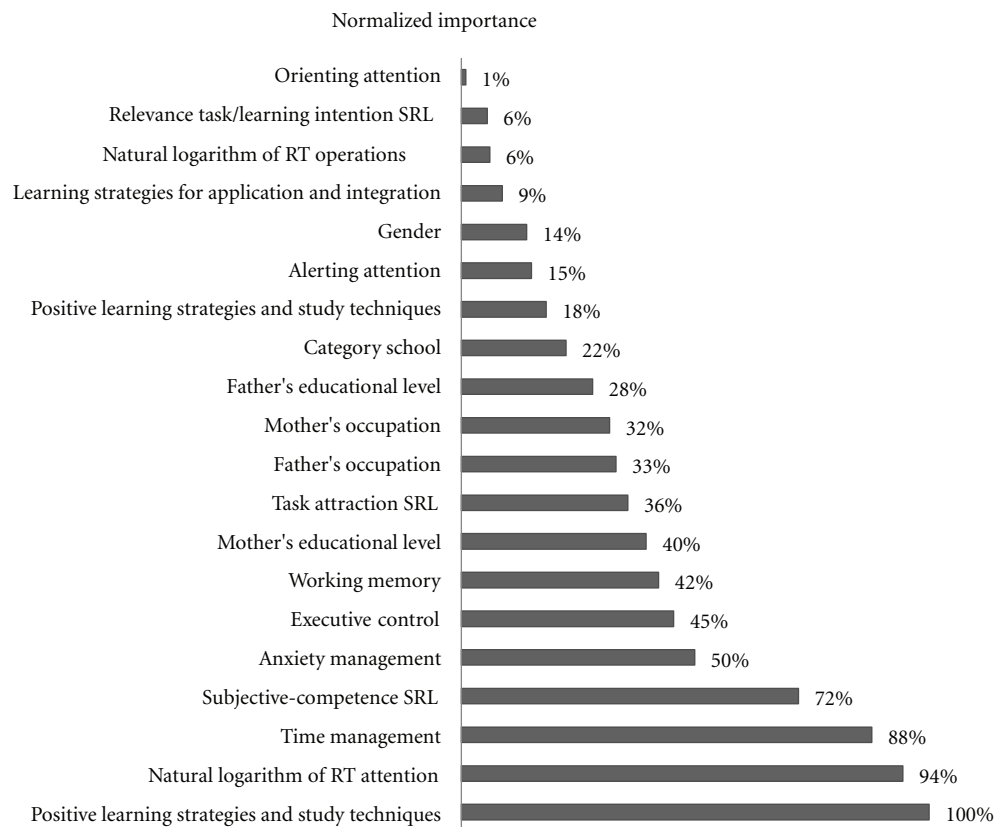
Normalized importance

Orienting attention	1%
Relevance task/learning intention SRL	6%
Natural logarithm of RT operations	6%
Learning strategies for application and integration	9%
Gender	14%
Alerting attention	15%
Positive learning strategies and study techniques	18%
Category school	22%
Father's educational level	28%
Mother's occupation	32%
Father's occupation	33%
Task attraction SRL	36%
Mother's educational level	40%
Working memory	42%
Executive control	45%
Anxiety management	50%
Subjective-competence SRL	72%
Time management	88%
Natural logarithm of RT attention	94%
Positive learning strategies and study techniques	100%

FIGURE 3: Normalized importance of the top variables participating in the model for the predictive classification of the middle 30% of scores in the mathematics test.

covariance matrices that, in this case, is violated (Box's $M = 302.730$, $F = 1.382$, $P = .001$). None of the variables, that were also included in the ANN1, were able to discriminate between the two groups of the students. The squared canonical correlation (CR^2) gives the amount of variation between the groups that is explained by the discriminating variables; in this case it was low (Wilk's $\lambda = .786$, $\chi^2 = 135.050$, df = 20, $P = .001$, $CR^2 = .21$).

DA2 was calculated to discriminate between the highest 30% of mathematical performance and the rest 70% of students, entering the same independent variables that were used into ANN2. Results show that the independent variables were not able to discriminate between both groups of students. The Box's M statistic is significant (Box's $M = 321.229$, $F = 1.471$, $P = .001$), meaning that the assumption of equality of covariance matrices is violated. In this analysis the squared canonical correlation indicated that the strength of the function is low (Wilk's $\lambda = .756$, $\chi^2 = 157.390$, df = 20, $P = .001$, $CR^2 = .24$).

DA3 involved the same variable as the ANN3 to predict the middle 30% of mathematical performance. The assumption of equality of covariance matrices was violated (Box's $M = 303.897$, $F = 1.395$, $P = .001$). In this case, also none of the variables were able to discriminate between the students in the mid 30% group and the extreme 30% groups; the explained variance was also low (Wilk's $\lambda = .951, \chi^2 = 28.336$,df = 20, $P = .102$, $CR^2 = .049$).

5. Discussion

It is clear from these results that besides the high predictive power of the three neural networks to model the expected performance of low, middle, and high performance groups of students, this methodology has also detected important differences in the factors that seem to underlie the students' performance. Among the student groups with the lowest 30% of math performance, the main determinants of performance appear to be basic cognitive processing variables, indicating the degree to which they represent the areas of relative weakness in the group and more discriminating from the rest of the students. This seems to indicate that it is the area of basic cognitive abilities, in other words, the basic processing capacity of the cognitive system in these students that best provides the information necessary to correctly identify this group. There is an extensive literature indicating a strong correlation between poor math performance and low working memory [56–62]. Children with mathematical disorders have difficulty in holding information in their memory while performing another activity (e.g., counting) or when they must inhibit irrelevant information [7, 63]. They perform poorly in tasks that measure visual-spatial working memory and central executive functions [61, 64–67].

On the other hand, among the student groups with the highest 30% of math performance, the main determinants of performance appear to be self-regulation and background variables (particularly, how interested students were in the task and social indicators such as parents' occupation). In

this group cognitive processing variables had much lower levels of importance in terms of their predictive weights, probably due to the fact that this group was much stronger in its levels of cognitive processing, and therefore these variables are less discriminating when the model attempts to classify the students according to their performance level. Working memory, reaction time, and attentional networks seem to be much less discriminating among students who reach certain threshold levels needed for basic mathematical problem solving. The items about self-regulated learning were important for students with high performance, reflecting their appreciation of the content and context of the math test. The model of adaptive learning [100–102] explains how the mental representation that the student forms of a specific (math) task depends on the context sensitivity, and this context depends on the links between specific appraisals and one or more domain-specific motivational beliefs. Consequently, the performance will be moderated by these interactions, and in this study in particular, SRL-variable results become even more predictive when the basic cognitive variables become less discriminating (high performance group).

The prediction for the middle 30% level of math performance group of students shows a particular pattern involving learning strategies and self-efficacy (as important motivational beliefs), together with attentional resources, as important predictors. Moreover, working memory does not seem to improve the prediction of performance for this middle group of students, indicating that their mathematical performance is more determined by processes related to self-regulated learning (i.e., learning strategies, motivational beliefs, and attention). These are variables more related to environmental, instructional, and training constructs, rather than to basic cognitive processes such as working memory.

The results of the discriminant analyses (DA) confirm the lack of significant linear relations between the independent variables analysed here and mathematical performance. Neural network models have an important advantage in this area, because ANN models are able to model nonlinear and complex relationships among variables. Another assumption required for traditional statistical predictive models (e.g., equality of covariance matrices) was violated for the three stepwise discriminant analyses that were performed to predict a specific category (lowest 30% or not, highest 30% or not, and middle 30% or not). Even with this restriction, the amount of variance explained was low in the three DA analyses. None of the variables were able to discriminate between the different categories of mathematical performance. When we compare these results with the ANNs analysed in this study, it can be concluded that ANNs are more robust and perform significantly better than other classical techniques, as prior studies have indicated [36, 95].

The predictive systems approach allows for the conceptualization and development of new modes of assessment which could facilitate breaking away from traditional forms of testing while at the same time improving the quality of the assessment process [23, 42, 103]. In addition, this new methodology allows for the understanding of the students' individual characteristics and for the prediction of

expected performance levels. This opens major possibilities for improvement of evaluation procedures and the planning of interventions. In addition, it has implications for the application of these methods in educational research and in the implementation of diagnostic "early-warning" programs in educational settings, as well as informing cognitive theory and the development of automated tutoring and learning systems. The capacity to very accurately classify students, which is also what tests attempt to do, without the performance sampling issues of traditional testing, and using a much broader spectrum of all factors influencing a student's overall performance, is a major advantage of the ANNs methodology. In fact, such an approach also represents a much more valid approach to educational evaluation due to its overall accuracy and the breadth of the constructs considered to classify the expected performance.
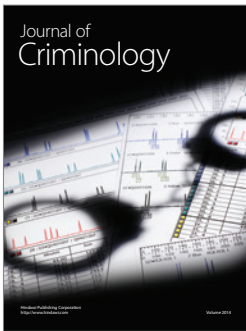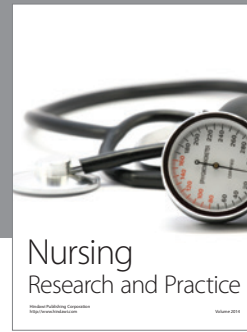
References

[1] M. H. Ashcraft, "Cognitive psychology and simple arithmetic: a review and summary of new directions," *Mathematical Cognition*, vol. 1, pp. 3–34, 1995.

[2] D. C. Geary, "A componential analysis of an early learning deficit in mathematics," *Journal of Experimental Child Psychology*, vol. 49, no. 3, pp. 363–383, 1990.

[3] D. C. Geary and K. F. Widaman, "Numerical cognition: on the convergence of componential and psichometric models," *Inteligence*, vol. 16, pp. 47–80, 1992.

[4] G. J. Hitch, "The role of short-term working memory in mental arithmetic," *Cognitive Psychology*, vol. 10, no. 3, pp. 302–323, 1978.

[5] P. Lemaire, H. Abdi, and M. Fayol, "The role of working memory resources in simple cognitive arithmetic," *European Journal of Cognitive Psychology*, vol. 8, no. 1, pp. 73–103, 1996.

[6] R. H. Logie, K. J. Gilhooly, and V. Wynn, "Counting on working memory in arithmetic problem solving," *Memory and Cognition*, vol. 22, no. 4, pp. 395–410, 1994.

[7] M. C. Passolunghi, C. Cornoldi, and S. De Liberto, "Working memory and intrusions of irrelevant information in a group of specific poor problem solvers," *Memory and Cognition*, vol. 27, no. 5, pp. 779–790, 1999.

[8] M. C. Passolunghi and F. Pazzaglia, "Individual differences in memory updating in relation to arithmetic problem solving," *Learning and Individual Differences*, vol. 14, no. 4, pp. 219–230, 2004.

[9] S. J. Pickering, *Working Memory and Education*, Academic Press, 2006.

[10] K. F. Widaman, D. C. Geary, P. Cormier, and T. D. Little, "A componential model for mental addition," *Journal of Experimental Psychology*, vol. 15, no. 5, pp. 898–919, 1989.

[11] J. W. Adams and G. J. Hitch, "Working memory and children's mental addition," *Journal of Experimental Child Psychology*, vol. 67, no. 1, pp. 21–38, 1997.

[12] M. I. Posner and M. K. Rothbart, "Attention, self-regulation and consciousness," *Philosophical Transactions of the Royal Society B*, vol. 353, no. 1377, pp. 1915–1927, 1998.

[13] M. R. Rueda, M. I. Posner, and M. K. Rothbart, "Attentional control and self regulation," in *Handbook of Self Regulation: Research, Theory, and Applications*, R. F. Baumeister and K. D. Vohs, Eds., vol. 14, pp. 283–300, Guilford Press, New York, NY, USA, 2004.

[14] M. Boekaerts, "Being concerned with well-being and with learning," *Educational Psychologist*, vol. 28, no. 2, pp. 149–167, 1993.

[15] L. Corno, "Student volition and education: outcomes influences, and practices," in *Self-Regulation of Learning and Performance*, B. J. Zimmerman and D. H. Schunk, Eds., pp. 229–254, Lawrence Erlbaum, Hillsdale, NJ, USA, 1994.

[16] P. R. Pintrich, "The role of goal orientation in self-regulated learning," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 452–502, Academic Press, San Diego, Calif, USA, 2000.

[17] R. M. Ryan and E. L. Deci, "An overview of self-determination theory: an organistic-dialectical perspective," in *Handbook of Self-Determination Research*, E. L. Deci and R. M. Ryan, Eds., pp. 3–34, University of Rochester, Rochester, NY, USA, 2000.

[18] C. Sansone and J. M. Harackiewicz, "'I don't feel like it': the function of interest in self-regulation," in *Striving and Feeling: Interactions Among Goals, Affect, and Selfregulation*, L. L. Martin and A. Tesser, Eds., pp. 203–228, Lawrence Erlbaum, Mahwah, NJ, USA, 1996.

[19] E. De Corte, P. Op't Eynde, and L. Verschaffel, "Knowing what to believe: the relevance of mathematical beliefs for mathematics education," in *Personal Epistemology: The Psychology of Beliefs About Knowledge and Knowing*, B. K. Hofer and P. R. Pintrich, Eds., pp. 297–320, Erlbaum, Mahwah, NJ, USA, 2002.

[20] K. R. Muis, "Personal epistemology and mathematics: a critical review and synthesis of research," *Review of Educational Research*, vol. 74, no. 3, pp. 317–377, 2004.

[21] E. C. Cascallar and M. F. Musso, "Classificatory stream analysis in the prediction of expected reading readiness: understanding student performance," *International Journal of Psychology*, vol. 43, no. 43-44, p. 231, 2008, 29th International Congress of Psychology ICP.

[22] M. F. Musso and E. C. Cascallar, "New approaches for improved quality in educational assessments: using automated predictive systems in reading and mathematics," *Journal of Problems of Education in the 21st Century*, vol. 17, pp. 134–151, 2009.

[23] M. Boekaerts and E. Cascallar, "How far have we moved toward the integration of theory and practice in self-regulation?" *Educational Psychology Review*, vol. 18, no. 3, pp. 199–210, 2006.

[24] S. L. Gulliford, S. Webb, C. G. Rowbottom, D. W. Corne, and D. P. Dearnaley, "Use of artificial neural networks to predict biological outcomes for patients receiving radical radiotherapy of the prostate," *Radiotherapy and Oncology*, vol. 71, no. 1, pp. 3–12, 2004.

[25] W. D. Neal and J. Wurst, "Advances in market segmentation," *Marketing Research*, vol. 13, no. 1, pp. 14–18, 2001.

[26] C. Jiang and F. Song, "Forecasting chaotic time series of exchange rate based on nonlinear autoregressive model," in *Proceedings of the IEEE International Conference on Advanced Computer Control (ICACC '10)*, Shangai, China, 2010.

[27] P. J. G. Lisboa, "A review of evidence of health benefit from artificial neural networks in medical intervention," *Neural Networks*, vol. 15, no. 1, pp. 11–39, 2002.

[28] A. S. Andreou and G. A. Zombanakis, "A neural network measurement of relative military security: the case of Greece and Cyprus," MPRA Paper 14539, 2009, http://mpra.ub.uni-muenchen.de/14539/.

[29] S. E. Haupt, A. Pasini, and C. Marzban, Eds., *Artificial Intelligence Methods in the Environmental Sciences*, vol. 8, 2009.

[30] E. C. Cascallar, "Neural networks applications for the predictive classification of complex group behavior in terrorist activities," Tech. Rep., USA Government, Washington, DC, USA, 2003.

[31] H. M. Al-Deek, "Which method is better for developing freight planning models at seaports—neural networks or multiple regression?" *Transportation Research Record*, no. 1763, pp. 90–97, 2001.

[32] N. Nguyen and A. Cripps, "Predicting housing value: a comparison of multiple regression and artificial neural networks," *Journal of Real Estate Research*, vol. 22, no. 3, pp. 313–336, 2001.

[33] H. White and J. Racine, "Statistical inference, the bootstrap, and neural-network modeling with application to foreign exchange rates," *IEEE Transactions on Neural Networks*, vol. 12, no. 4, pp. 657–673, 2001.

[34] M. Laguna and R. Martí, "Neural networks prediction in a system for optimizing simulations," *IIE Transactions*, vol. 34, no. 3, pp. 273–282, 2002.

[35] K. B. DeTienne, D. H. DeTienne, and S. A. Joshi, "Neural networks as statistical tools for business researchers," *Organizational Research Methods*, vol. 6, no. 2, pp. 236–265, 2003.

[36] H. T. Everson, D. Chance, and S. Lykins, "Exploring the use of artificial neural networks in educational research," in *Proceedings of the Annual Meeting of the American Educational Research Association*, New York, NY, USA, 1994.

[37] R. L. Wilson and B. C. Hardgrave, "Predicting graduate student success in a MBA program: regression vs. classification," *Educational and Psychological Measurement*, vol. 55, no. 2, pp. 186–195, 1995.

[38] H. T. Everson, "Modeling the student in intelligent tutoring sytems: the promise of a new psychometrics," *Instructional Science*, vol. 23, no. 5-6, pp. 433–452, 1995.

[39] S. M. Weiss and C. A. Kulikowski, *Computer Systems that Learn*, Morgan Kaufmann, San Mateo, Calif, USA, 1991.

[40] K. Perkins, L. Gupta, and R. Tammana, "Predict item difficulty in a reading comprehension test with an artificial neural network," *Language Testing*, vol. 12, no. 1, pp. 34–53, 1995.

[41] R. W. Engle, "Working memory capacity as executive attention," *Current Directions in Psychological Science*, vol. 11, no. 1, pp. 19–23, 2002.

[42] E. C. Cascallar, M. Boekaerts, and T. E. Costigan, "Assessment in the evaluation of self-regulation as a process," *Educational Psychology Review*, vol. 18, no. 3, pp. 297–306, 2006.

[43] M. F. Musso and E. C. Cascallar, "Predictive systems using artificial neural networks: an introduction to concepts and applications in education and social sciences," in *Research in Behavioural Sciences*, M. C. Richaud and J. E. Moreno, Eds., vol. 1, pp. 433–459, CIIPME/CONICET, 2009.

[44] N. Unsworth, R. P. Heitz, J. C. Schrock, and R. W. Engle, "An automated version of the operation span task," *Behavior Research Methods*, vol. 37, no. 3, pp. 498–505, 2005.

[45] G. Arteaga Díaz and H. Pimienta Jimenez, "Working memory and cortical pathways," *Revista De La Facultad De Medicina De La Universidad Nacional De Colombia*, vol. 54, pp. 248–268, 2006.

[46] M. Daneman and P. A. Carpenter, "Individual-differences in working memory and reading," *Journal of Verbal Learning and Verbal Behaviour*, vol. 19, pp. 450–466, 1980.

[47] R. W. Engle, J. Cantor, and J. J. Carullo, "Individual differences in working memory and comprehension: a test of four hypotheses," *Journal of Experimental Psychology*, vol. 18, no. 5, pp. 972–992, 1992.

[48] T. A. Salthouse and R. L. Babcock, "Decomposing adult age differences in working memory," *Developmental Psychology*, vol. 27, no. 5, pp. 763–776, 1991.

[49] P. Shah and A. Miyake, "The separability of working memory resources for spatial thinking and language processing: an individual differences approach," *Journal of Experimental Psychology*, vol. 125, no. 1, pp. 4–27, 1996.

[50] J. R. Anderson, L. M. Reder, and C. Lebiere, "Working memory: activation limitations on retrieval," *Cognitive Psychology*, vol. 30, no. 3, pp. 221–256, 1996.

[51] T. A. Salthouse, "Why do adult age differences increase with task complexity?" *Developmental Psychology*, vol. 28, no. 5, pp. 905–918, 1992.

[52] J. LeFevre, D. DeStefano, B. Coleman, and T. Sanan, "Mathematical cognition and working memory," in *Handbook of Mathematical Cognition*, J. I. D. Campbell, Ed., pp. 361–411, Psychology Press, New York, NY, USA, 2005.

[53] D. DeStefano and J. A. LeFevre, "The role of working memory in mental arithmetic," *European Journal of Cognitive Psychology*, vol. 16, no. 3, pp. 353–386, 2004.

[54] M. H. Ashcraft and E. P. Kirk, "The relationships among working memory, math anxiety, and performance," *Journal of Experimental Psychology*, vol. 130, no. 2, pp. 224–237, 2001.

[55] M. H. Ashcraft, R. D. Donley, M. A. Halas, and M. Vakali, "Working memory, automaticity, and problem difficulty," in *The Nature and Origins of Mathematical Skills*, J. I. D. Campbell, Ed., pp. 301–329, North- Holland, Oxford, UK, 1992.

[56] D. C. Geary, "Mathematical disabilities: cognitive, neuropsychological, and genetic components," *Psychological Bulletin*, vol. 114, no. 2, pp. 345–362, 1993.

[57] J. G. Hitch and E. McAuley, "Working memory in children with specific arithmetical learning difficulties," *The British Journal of Psychology*, vol. 82, part 3, pp. 375–386, 1991.

[58] M. C. Passolunghi and F. Pazzaglia, "A comparison of updating processes in children good or poor in arithmetic word problem-solving," *Learning and Individual Differences*, vol. 15, no. 4, pp. 257–269, 2005.

[59] M. C. Passolunghi and L. S. Siegel, "Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving," *Journal of Experimental Child Psychology*, vol. 80, no. 1, pp. 44–57, 2001.

[60] M. C. Passolunghi and L. S. Siegel, "Working memory and access to numerical information in children with disability in mathematics," *Journal of Experimental Child Psychology*, vol. 88, no. 4, pp. 348–367, 2004.

[61] L. S. Siegel and E. B. Ryan, "The development of working memory in normally achieving and subtypes of learning disabled children," *Child Development*, vol. 60, no. 4, pp. 973–980, 1989.

[62] H. L. Swanson, "Working memory in learning disability subgroups," *Journal of Experimental Child Psychology*, vol. 56, no. 1, pp. 87–114, 1993.

[63] M. C. Passolunghi and C. Cornoldi, "Working memory and cognitive abilities in children with specific difficulties in arithmetic word problem solving," *Advances in Learning and Behavioral Disabilities*, vol. 14, pp. 155–178, 2000.

[64] S. E. Gathercole and S. J. Pickering, "Working memory deficits in children with low achievements in the national curriculum at 7 years of age," *British Journal of Educational Psychology*, vol. 70, no. 2, pp. 177–194, 2000.

[65] S. E. Gathercole and S. J. Pickering, "Assessment of working memory in six- and seven-year-old children," *Journal of Educational Psychology*, vol. 92, no. 2, pp. 377–390, 2000.

[66] D. C. Geary, M. K. Hoard, and C. O. Hamson, "Numerical and arithmetical cognition: patterns of functions and deficits in children at risk for a mathematical disability," *Journal of Experimental Child Psychology*, vol. 74, no. 3, pp. 213–239, 1999.

[67] J. F. McLean and G. J. Hitch, "Working memory impairments in children with specific arithmetic learning difficulties," *Journal of Experimental Child Psychology*, vol. 74, no. 3, pp. 240–260, 1999.

[68] M. A. Just, P. A. Carpenter, and D. D. Hemphill, "Constraints on processing capacity: architectural or implementational?" in *Mind Matters: A Tribute To Allen Newell*, D. M. Steier and T. M. Mitchell, Eds., pp. 141–178, Erlbaum, Mahwah, NJ, USA, 1996.

[69] R. Tannock, "Inattention and working memory: effects on academic performance," in *Proceedings of the 23rd Annual Learning Differences Conference*, Harvard Graduate School of Education, Cambridge, Mass, USA, March 2008.

[70] R. Barkley, *Attention-Deficit Hyperactivity Disorder: A Handbook For Diagnosis and Treatment*, Guilford Press, New York, NY, USA, 1998.

[71] A. Fernández-Castillo and M. E. Gutiérrez-Rojas, "Selective attention, anxiety, depressive symptomatology and academic performance in adolescents," *Electronic Journal of Research in Educational Psychology*, vol. 7, no. 1, pp. 49–76, 2009.

[72] J. Fan, B. D. McCandliss, T. Sommer, A. Raz, and M. I. Posner, "Testing the efficiency and independence of attentional networks," *Journal of Cognitive Neuroscience*, vol. 14, no. 3, pp. 340–347, 2002.

[73] T. S. Redick and R. W. Engle, "Working memory capacity and attention network test performance," *Applied Cognitive Psychology*, vol. 20, no. 5, pp. 713–721, 2006.

[74] L. Corno, "Volitional aspects of self-regulated learning," in *Self-Regulated Learning and Academic Achievement: Theoretical Perspectives*, B. J. Zimmerman and D. H. Schunk, Eds., pp. 191–226, Lawrence Erlbaum, Mahwah, NJ, USA, 2nd edition, 2001.

[75] C. A. Wolters and H. Rosenthal, "The relation between students' motivational beliefs and their use of motivational regulation strategies," *International Journal of Educational Research*, vol. 33, no. 7-8, pp. 801–820, 2000.

[76] C. E. Weinstein and R. E. Mayer, "The teaching of learning strategies," in *Handbook of Research on Teaching*, M. C. Wittrock, Ed., Macmillan, New York, NY, USA, 3rd edition, 1986.

[77] C. E. Weinstein, D. R. Palmer, and A. C. Schulte, *Learning and Study Strategies Inventory*, H & H, Clearwater, Fla, USA, 1987.

[78] C. E. Weinstein, A. C. Schulte, and E. C. Cascallar, "The learning and studies strategies inventory (LASSI): initial design and development," Tech. Rep., US Army Research Institute for the Social and Behavioural Sciences, Alexandria, Va, USA, 1982.

[79] C. E. Weinstein and D. R. Palmer, *LASSI: User's Manual*, H & H, Clearwater, Fla, USA, 2nd edition, 2002.

[80] M. Boekaerts, *Motivation To Learn. Educational Practices Series-10*, International Academy of Education, 2002, http://www.ibe.unesco.org/publications/EducationalPracticesSeriesPdf/prac10e.pdf.

[81] E. De Corte, L. Verschaffel, and P. Op 't Eynde, "Self-regulation: a characteristic and a goal of mathematics education," in *Handbook of Self-Regulation*, M. Boekaerts, P. R. Pintrich, and M. Zeidner, Eds., pp. 687–726, Academic Press, San Diego, Calif, USA, 2000.

[82] K. R. Muis, "The role of epistemic beliefs in self-regulated learning," *Educational Psychologist*, vol. 42, no. 3, pp. 173–190, 2007.

[83] A. H. Schoenfeld, "Beyond the purely cognitive: belief systems, social cognitions, and metacognitions as driving forces in intellectual performance," *Cognitive Science*, vol. 7, no. 4, pp. 329–363, 1983.

[84] A. J. Elliot, "A conceptual history of the achievement goal construct," in *Handbook of Competence and Motivation*, A. Elliot and C. Dweck, Eds., pp. 52–72, Guilford Press, New York, NY, USA, 2005.

[85] B. Weiner, "A theory of motivation for some classroom experiences," *Journal of Educational Psychology*, vol. 71, no. 1, pp. 3–25, 1979.

[86] A. Bandura, *Social Foundations of Thought and Action: A Social Cognitive Theory*, Prentice Hall, Englewood Cliffs, NJ, USA, 1986.

[87] F. Pajares, "Motivational role of self-efficacy beliefs in self-regulated learning," in *Motivation and Self-Regulated Learning*, D. H. Schunk and B. J. Zimmerman, Eds., pp. 111–139, Erlbaum, Mahwah, NJ, USA, 2008.

[88] A. Wigfield and J. S. Eccles, "The development of competence beliefs and values from childhood through adolescence," in *Development of Achievement Motivation*, A. Wigfield and J. S. Eccles, Eds., pp. 92–120, Academic Press, San Diego, Calif, USA, 2002.

[89] D. F. Specht, "A general regression neural network," *IEEE Transactions on Neural Networks*, vol. 2, no. 6, pp. 568–576, 1991.

[90] M. L. Mavrovouniotis and S. Chang, "Hierarchical neural networks," *Computers and Chemical Engineering*, vol. 16, no. 4, pp. 347–369, 1992.

[91] S. Grossberg, "How does a brain build a cognitive code?" *Psychological Review*, vol. 87, no. 1, pp. 1–51, 1980.

[92] S. Grossberg, *Studies of Mind and Brain: Neural Principles of Learning, Perception, Development, Cognition and Motor Control*, Reidel Press, Boston, Mass, USA, 1982.

[93] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, no. 6088, pp. 533–536, 1986.

[94] D. E. Rumelhart, J. L. McClelland, and P. D. P. research group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, vol. 1, MIT Press, Cambridge, Mass, USA, 1986.

[95] L. Marquez, T. Hill, R. Worthley, and W. Remus, "Neural network models as an alternative to regression," in *Proceedings of the IEEE 24th Annual Hawaii International Conference on Systems Sciences*, vol. 4, pp. 129–135, 1991.

[96] R. P. Lippmann, "An introduction to computing with neural nets," *IEEE ASSP Magazine*, vol. 4, no. 2, pp. 4–22, 1987.

[97] M. Boekaerts, "The on-line motivation questionnaire: a self-report instrument to assess students' context sensitivity. Advances in motivation and achievement," in *New Directions in Measures and Methods*, vol. 12, pp. 77–120, Elsevier, New York, NY, USA, 2002.

[98] N. Cortada de Kohan and G. Macbeth, "Construction of a mathematical test for adolescents and adults," *Interdisciplinary*, vol. 24, no. 1, pp. 43–64, 2007.

[99] Trends in international Mathematics and Science Study (TIMSS). Prepared by the International Association for the Evaluation of Educational Achievement (IEA). Amsterdam, 1995, http://timss.bc.edu/timss1995i/TIMSSPDF/C_items.pdf.

[100] M. Boekaerts, "The adaptable learning process: initiating and maintaining behavioural change," *Applied Psychology*, vol. 41, no. 4, pp. 377–397, 1992.

[101] M. Boekaerts, "Self-regulated learning: bridging the gap between metacognitive and metamotivation theories," *Educational Psychologist*, vol. 30, no. 4, pp. 195–200, 1995.

[102] M. Boekaerts, "Self-regulated learning at the junction of cognition and motivation," *European Psychologist*, vol. 1, no. 2, pp. 100–112, 1996.

[103] M. Segers, F. Dochy, and E. Cascallar, *Optimizing New Modes of Assessment: In Search of Qualities and Standards*, Kluwer Academic, Dordrecht, The Netherlands, 2003.