



International Journal of Lean Six Sigma

Six Sigma learning evaluation model using Bloom's Taxonomy

Gabriela Fonseca Amorim, Pedro Paulo Balestrassi, Rapinder Sawhney, Mariângela de Oliveira-Abans, Diogo Leonardo Ferreira da Silva,

Article information:

To cite this document:

Gabriela Fonseca Amorim, Pedro Paulo Balestrassi, Rapinder Sawhney, Mariângela de Oliveira-Abans, Diogo Leonardo Ferreira da Silva, (2018) "Six Sigma learning evaluation model using Bloom's Taxonomy", International Journal of Lean Six Sigma, Vol. 9 Issue: 1, pp.156-174, <https://doi.org/10.1108/IJLSS-01-2017-0006>

Permanent link to this document:

<https://doi.org/10.1108/IJLSS-01-2017-0006>

Downloaded on: 24 January 2019, At: 03:43 (PT)

References: this document contains references to 41 other documents.

To copy this document: permissions@emeraldinsight.com

The fulltext of this document has been downloaded 465 times since 2018*

Users who downloaded this article also downloaded:

(2018), "Annotated methodological review of Lean Six Sigma", International Journal of Lean Six Sigma, Vol. 9 Iss 1 pp. 2-49 <<https://doi.org/10.1108/IJLSS-04-2017-0028>><https://doi.org/10.1108/IJLSS-04-2017-0028>

(2018), "Development of Six Sigma methodology to improve grinding processes: A change management approach", International Journal of Lean Six Sigma, Vol. 9 Iss 1 pp. 50-63 <<https://doi.org/10.1108/IJLSS-11-2016-0074>><https://doi.org/10.1108/IJLSS-11-2016-0074>

Access to this document was granted through an Emerald subscription provided by emerald-srm:478341 []

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

*Related content and download information correct at time of download.

Six Sigma learning evaluation model using Bloom's Taxonomy

Gabriela Fonseca Amorim

*Institute of Production and Management Engineering,
Universidade Federal de Itajuba, Itajuba, Brazil and Department of Industrial and
Systems Engineering, University of Tennessee, Knoxville, Tennessee, USA*

Pedro Paulo Balestrassi

*Institute of Production and Management Engineering,
Universidade Federal de Itajuba, Itajuba, Brazil*

Rapinder Sawhney

*Department of Industrial and Systems Engineering,
University of Tennessee, Knoxville, Tennessee, USA*

Mariângela de Oliveira-Abans

*Institute of Production and Management Engineering,
Universidade Federal de Itajuba, Itajuba, Brazil and Laboratorio Nacional de
Astrofísica, Itajuba, Brazil, and*

Diogo Leonardo Ferreira da Silva

*Institute of Systems Engineering and Information Technology,
Universidade Federal de Itajuba, Itajuba, Brazil and Department of Industrial and
Systems Engineering, University of Tennessee, Knoxville, Tennessee, USA*

Abstract

Purpose – This paper aims to propose a learning evaluation model for Green Belts and Black Belts at the training level. A question bank has been developed on the basis of Bloom's learning classification and applied to a group of employees who were being trained in Six Sigma (SS). Their results were then used to decide on the students' approval and to guide the instructor's plan of teaching for the next classes.

Design/methodology/approach – An action research has been conducted to develop a question bank of 310 questions based on the revised Bloom's Taxonomy, to implement the evaluation model, and to apply it during the SS training.

Findings – The evaluation model has been designed so that the students do not proceed unless they have acquired the conceptual knowledge at each step of the DMAIC (Define, Measure, Analyze, Improve and Control) roadmap. At the end of the evaluation process, the students' results have been analyzed. The number of mistakes in all stages of DMAIC was equal, implying that the training was uniform the entire roadmap. However, the opposite happened in each of the Bloom's Taxonomy levels, showing that some skills need to be better stimulated by the instructor than others.



Research limitations/implications – The learning evaluation model proposed in this paper has been applied to a group of 70 employees who were being trained in SS at a Brazilian aircraft manufacturer. The data have been analyzed using Microsoft Excel[®] and Minitab[®] 17 Statistical Software.

Originality/value – Despite the abundance of courses offering the SS Green Belt and Black Belt certifications, there is no standard evaluation to ensure the training quality. Thus, this paper proposes an innovative learning evaluation model.

Keywords Six Sigma, Training, DMAIC, Learning evaluation model, Revised Bloom's Taxonomy

Paper type Research paper

1. Introduction

For an effective Six Sigma (SS) deployment, it is necessary to invest in quality training. Despite the abundance of courses offering certification, a search regarding the SS trainings at Web of Science and Scopus databases allows to affirm that there is no standard evaluation to ensure the quality of education or even that trained professionals will be able to perform their functions. This is, then, a demand which has not been fulfilled so far.

In a broad educational vision, evaluations provide guidelines, allowing reflection on curricular purposes and enriching the learning process. An interesting and promising approach is to structure them based on Bloom's Taxonomy. The cognitive domain covered in the study by [Bloom *et al.* \(1956\)](#) deals precisely with learning.

This taxonomy was reviewed by some of Bloom's students, updated and published under a new title by [Anderson *et al.* \(2001\)](#). The revised Bloom's Taxonomy (RBT) divides the cognitive domain into six levels: remember, understand, apply, analyze, evaluate and create. Also according to these authors, its structure allows evaluating the validity and coverage of any training, course, curriculum or program.

On the basis of the RBT, this work presents an innovative training evaluation model for Green Belts (GBs) and Black Belts (BBs) that, besides being used as an evaluation tool to certify the education quality, is also suitable for obtaining feedback to better understand the students' training needs.

With this purpose in mind and following the structure of an action research, as proposed by [Coughlan and Coughlan \(2002\)](#), a database with 310 questions has been created and randomly applied to groups being trained in SS at a Brazilian aircraft company. These questions approached each stage of DMAIC (Define, Measure, Analyze, Improve and Control), a popular roadmap adopted in *Lean Six Sigma and Minitab - the Complete Toolbox Guide for All Lean Six Sigma Practitioners* by [Brook \(2010\)](#), a book followed by instructors at training as a guide to implement SS.

The tests have been performed by 70 students and their results are statistically analyzed (see Section 4) to be used not only in a classificatory way, but also as a teaching feedback. The evaluations and the training were developed in Portuguese because most of the employees from the Brazilian companies does not speak English. SS being a technical subject, the translation is literal and thus there is no significant loss of meaning when translated into other languages.

This work is organized as follows: Section 1 introduces and contextualizes the subject. Section 2 presents a literature review about SS, how to measure and evaluate the students at trainings and the RBT. Section 3 details the methodology of the proposed learning evaluation model and how it was applied. Section 4 contains the statistical analysis of the results. Section 5 presents the conclusions and some recommendations.

2. Literature review

2.1 Six Sigma and the roadmap DMAIC

The SS philosophy had been indirectly introduced over the years, but it was officially introduced in 1987 (Folaron, 2003). Bill Smith, an engineer at Motorola, developed the SS program to meet the needs of quality improvement and defect reduction. Bob Galvin, Motorola's CEO at that time, was impressed with the results and decided to apply SS with a focus on manufacturing processes (Montgomery and Woodall, 2008).

In 1988, Motorola developed the curriculum of the SS tools and created courses for practitioners, leading the company to receive the Malcolm Baldrige Quality Award. The excellent results also attracted the attention of other organizations, such as Allied Signal, IBM and General Electric (Aboelimged, 2010). The initiative received major usage across business and industry; first in the USA and then globally (Snee, 2010).

There are different roadmaps for an SS deployment. This work has followed Brook's (2010) guide because it goes beyond traditional books and presents the transition from academic knowledge to real situations in an educational way. Brook adopted DMAIC roadmap combined with the Lean concepts. He also stated that there is, nevertheless, no clear distinction between Lean and SS when it comes to everyday projects. Salah *et al.* (2010) also discussed the integration of these two themes.

The acronym DMAIC is the main SS roadmap. The following summary of each step is based on the guidelines provided by Antony and Bhajji (2002), Eckes (2003), Montgomery and Woodall (2008) and mainly Brook (2010):

- *Define*: Includes definitions of the problem, customer and process. Then, the definitions of the project goals, impact on customer, schedule and a measurable target for the results. Finally, definition of the potential benefits the project could provide.
- *Measure*: The objective is to translate the problem into something measurable. It is important to define what should be measured and to create a data collection plan. Also, it is interesting to develop a measurement system analysis.
- *Analyze*: The objective is to identify the influential factors and causes that determine the behavior of features critical to quality based on the measurements of the previous step.
- *Improve*: The objective is to define and implement adjustments to improve the performance of features CTQ and focus on process optimization, aiming to achieve the technical and financial performance targets previously established for the project.
- *Control*: This step adjusts the management and the control system to prevent recurrence of the initial problem and to maintain the achieved performance. Control measures are implemented to allow corrective actions and to prevent the process from deteriorating.

2.1.1 *A typical Six Sigma team*. Team members and stakeholders at the SS projects were named inspired by martial arts grading as Champion, Master Black Belt, BB and GB. This work is focused on GBs and BBs, which are described below, based on the studies by Pande *et al.* (2000) and Harry and Schroeder (2000):

- *Black Belt*: This includes usually the SS team leader. BBs provide expert assistance with statistical tools, manage changes and projects strategies, revise and clarify the reasons for carrying out the projects to the Executive Leaders and Champions, and they are responsible for training and guidance of GBs. When integrated into the

team, they are primarily responsible for the maintenance and continuity of the SS project. If they are not integrated, they play the role of internal consultants. BBs usually have first a GB certificate.

- *Green Belt*: This includes team members who must be trained to become experts in the SS tools, even not having the same level of experience using statistics or the leadership skills of BBs. GBs apply SS as part of their work but do not focus exclusively on it. The main duties of this function are aiding BBs in data collection and development of experiments for improvement projects, beside leadership in small projects.

2.1.2 Six Sigma training. The SS training and the certification are different for each function. Champions receive an orientation for three to five days. Master Black Belts receive specific training and develop projects for about 200 hours to be instructors or internal consultants. BBs work full time in the management of projects and generally receive around 160 hours of training. GBs work part time in each project and are trained for about 80 hours (Linderman *et al.*, 2003; Schroeder *et al.*, 2008; Aboelmaged, 2010). According to the study by Salah *et al.* (2010), to be certified, BBs and GBs need to achieve the financial goals of the project done during training.

BB training includes learning statistical techniques and practice in applying the tools in real situations. According to the study by Ingle and Roe (2001), this intense training is organization-dependent. For example, at Motorola, the BB is an expert in solving problems and using the SS tools. At General Electric, besides being an expert in the SS tools, the BB receives business management education, becoming able to assume managerial positions. Following numerous discussions on the subject, Hoerl (2001) proposed a model of BB curriculum lasting four weeks.

Montgomery and Woodall (2008) claimed that the SS projects usually take four to six months to be completed and these are selected according to the potential financial impact. In many organizations, the GB training follows the SS implementation, enabling the practitioners to participate in the development of an entire project. In this way, students learn through training and simultaneous implementation of SS in real situations.

During training, both GBs and BBs should be given the chance to assimilate knowledge through its application to real projects. Brook's (2010) *Lean Six Sigma and Minitab - the Complete Toolbox Guide for All Lean Six Sigma Practitioners* works as a guide to Lean SS suitable for technical environments. The evaluation model proposed in this work can be used to guide GBs and BBs training, emphasizing the major difficulties of the group, and to measure the learning level, revealing whether they have enough knowledge to proceed with the next steps of training.

2.2 How to measure and evaluate the training

Any training is successful whenever students reach the expected technical knowledge level. How to measure this knowledge? Gronlund (1974) proposed the idea that the evaluation is a process and, therefore, it is most effective when based on operating principles. Kirkpatrick and Kirkpatrick (2006) proposed a schema to evaluate training programs in four levels: reaction, learning, behavior and results. Three divisions were proposed when it comes to learning: knowledge, skills and attitudes.

"Teaching to the test" is always an issue that comes up among teachers and test administrators (Arends, 2012). Avoiding this concern, Lister *et al.* (2004) conducted experiments in seven countries to test and compare the performance of students through a

standard evaluation tool. This same instrument was improved by Whalley *et al.* (2006) with the RBT and produced even better results.

This taxonomy was originated from discussions during the American Psychological Association convention in 1948 which led Benjamin Bloom to lead a group of educators who accepted the task of sorting the educational objectives (Guskey, 2001). The main intention was to develop a classification of important behaviors in learning process, which culminated in their separation into three main areas to be covered in three different books. These areas coincide with Kirkpatrick and Kirkpatrick's (2006) division for the learning level at training programs assessment:

- *Cognitive*: knowledge-based, domain divided into six levels: knowledge, comprehension, application, analysis, synthesis and evaluation (Bloom *et al.*, 1956).
- *Affective*: attitude-based, domain divided into five levels: receiving, responding, valuing, organization and characterization by value set (Krathwohl *et al.*, 1964).
- *Psychomotor*: skill-based, domain that was never published by the original group.

According to Bloom *et al.* (1956), the cognitive domain is critical to implement the evaluation; therefore a complete and efficient evaluation model for training in SS (DMAIC) should be focused on this area. This domain has been the most exploited by educators since the beginning.

With the education and teaching development over the years, it became necessary to bring new insights to the taxonomy. Since 1995, some experts have studied this taxonomy and, after considerable discussion, eventually decided to develop a second version (Anderson *et al.*, 2001). According to Krathwohl (2002), who participated in both versions, important changes were made as follows.

The names of all categories were changed from nouns to verbs and familiar terms were chosen. The RBT was published by Anderson *et al.* (2001) under the title *A Taxonomy For Learning, Teaching, and Assessing – A Revision of Bloom's Taxonomy of Educational Objectives*. A simplified description of each category of cognitive domain and its representation in terms of the subcategories are shown in Figure 1.

The original Bloom's Taxonomy and the Revised version have already been used to support teaching and learning in different countries, for example, Özcelik *et al.* (1993) in Turkey, Veeravagu *et al.* (2010) in Canada, Gaujard and Verzat (2011) in France and Galhardi and Azevedo (2013) in Brazil.

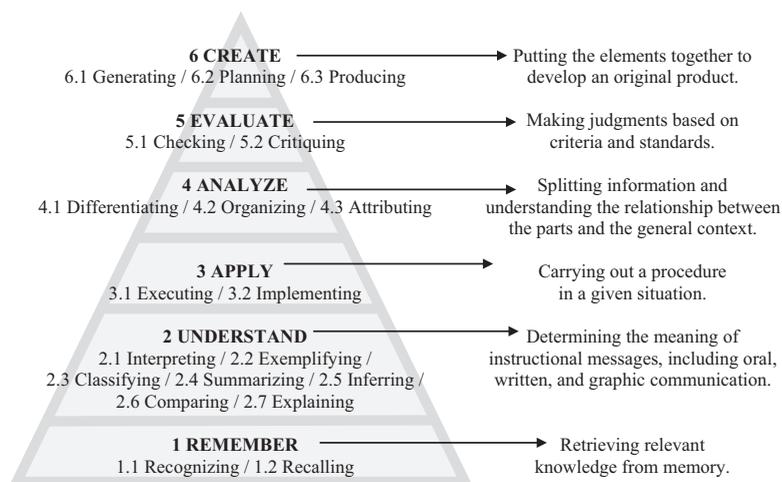
Moreover, the effectiveness of RBT has been confirmed by successful applications in various areas, such as project management (Athanassiou *et al.*, 2003), Web searching (Jansen *et al.*, 2009), computer science (Jesus and Raabe, 2009), decision support systems (Tyran, 2009), sustainability (Pappas *et al.*, 2013), medicine (Phillips *et al.*, 2013), nursing (Krau, 2011) or patient care (Larkin and Burton, 2008).

3. Methodology for an innovative Six Sigma learning evaluation model

The problem emerged from the lack of a structured evaluation model for GBs and BBs under training allied to the need of providing teaching feedback to instructors. To better understand the issues involved, the starting point was to build the scenario based on the actors' speech.

3.1 Surveying the universe of the instructors

An initial data collection for conducting the action research was carried out through interviews with the SS instructors from a group that regularly conduct outsourced courses



Source: Adapted from Anderson *et al.* (2001) and Krathwohl (2002)

Figure 1.
Cognitive domain of
RBT

in private companies. This group was chosen because of the training regularity that would permit a quick and easy access to data of several students in training. At these interviews, it was investigated how the instructors commonly proceed when evaluating the students and what did they expect from an efficient evaluation model.

The instructors said that the student evaluation during the SS trainings was made only through the delivery and presentation of a final project: A real SS project implemented at the workplace. After the presentations, all students were certified as GBs or BBs, but there was no guarantee that they had indeed learned the concepts or could apply them effectively.

Regarding the ideal evaluation model for this context, the instructors pointed closed-ended questions as preferable to avoid subjectivity and 20 questions as a fair amount to be answered in less than one hour. Students should be evaluated in each of the five phases of DMAIC roadmap and very long evaluations would not be appropriate. It would also be interesting to have different tests to be applied simultaneously, but the extra time it demands to prepare and correct is a hindrance to this situation.

The main idea of this work is then to produce an evaluation model with 20 questions randomly chosen from a question bank based on the first four levels of RBT for each of the DMAIC roadmap's phases. The existence of such a question bank with multiple choice-type questions would match the overall expectation of an adequate evaluation model with the possibility of applying different tests without demanding a lot of extra time in preparation and correction.

Among these 20 questions, 18 were defined as closed-ended and randomly chosen from the question bank. The level "evaluate" from RBT was considered equivalent to "analyze" in this assessment context. The number of questions and each one's predefined weight were planned to respect the increasing difficulty, more questions to remember (basic concepts) and less questions to analyze (more demanding), for example.

The two other questions for each DMAIC step, based on the last level of RBT, should be the same open-ended questions for everyone, shared with the students since the first day of training to be delivered at each evaluation day. These two questions are related to each

student's project, and thus the answers are supposed to be always different. [Table I](#) displays the basic plan for the learning evaluation model.

As previously defined by instructors, students must achieve a final grade of 80 points to be approved and to move for the next training phase. The instructors were used to these 80 points and all this work was done to rebuild the evaluation schema, and not the passing score. Coincidentally, adding up the score assigned to each question, the closed-ended questions totaled 79 points and the open-ended questions, 21 points. This was considered as a good proportion by the instructors because the "79 points are not enough to be approved, but very close to it".

A total of 310 questions were developed for the question bank created during this research addressing the entire content of the SS course. This suggested number can and should be extended over the years, according to the training needs.

3.2 Developing the questions based on the proposed evaluation model

For each step of the DMAIC roadmap, 60 closed-ended questions were developed: 32 to "remember", 16 to "understand", 8 to "apply" and 4 to "analyze". Also, two open-ended questions concerning the real SS project implementation were developed to assess the "create" level. There are many online guides for the different RBT levels to assist the questions development process. [Figure 2](#) displays ten examples of the 62 questions developed for the first step of DMAIC roadmap, two of each RBT levels.

The entire training as well as the evaluations was developed in the native language of employees, that is, Portuguese. SS being a technical subject, there was no loss of meaning when translating the questions to English or into other languages. The interested reader is referred to Author (Year)[1] for the complete list of questions.

As a result, an executable file with the evaluation was set for each stage of the DMAIC roadmap. The commercial software QuizMaker[®] was chosen because it is a complete and affordable tool with a friendly interface.

This software randomly selects 18 closed-ended questions from the question bank that compose each student's evaluation. The executable file requires the student's name and email when started. These data and the results are sent to the instructor's email.

At the end, the result is immediately shown on the computer's screen so that students can analyze their performance. [Figure 3](#) shows the message they receive when (a) failed or (b) approved at the assessment. The results, however, are sent only to the instructor.

About one month is devoted to each stage of DMAIC roadmap: classes are held in the first 15 days and followed by the evaluation. The student must deliver the two questions related to the final project at the very evaluation day. The instructors will only evaluate the open-ended questions in case of approval on the closed-ended questions. The passing score is 59, because adding the 21 points from the open-ended questions, students still can achieve

Table I.
Amount of questions
based on each level
of RBT for the
learning evaluation
model

Revised Bloom's Taxonomy	Amount	Individual weight	Total score	Type
Create	2	10.5	21	Open-ended
Evaluate	0	0	0	
Analyze	3	7	21	Closed-ended
Apply	4	5	20	Closed-ended
Understand	5	4	20	Closed-ended
Remember	6	3	18	Closed-ended
	<i>20</i>		<i>100</i>	

RBT	Examples of questions for Define	
Remember	<p>What is the focus of the Define phase?</p> <p>a) <u>Problems</u> b) Main solutions c) Basic solutions d) Hypotheses e) Losses</p>	<p>What is the one-page document that sums up the findings at the end of the Define phase and allows the project to be reviewed by everyone?</p> <p>a) Storyboards <u>b) Project charter</u> c) Checklist d) Summary e) Problem statement</p>
Understand	<p>The Define phase is important to:</p> <p>a) To point out those responsible and define their obligations in solving problems. b) Collect data from the process. c) Define the solutions to the problems identified. <u>d) Understand the problem before investing time and money on the project.</u> e) Provide resources to start the project.</p>	<p>What is the main purpose of developing the House of Quality?</p> <p>a) <u>Correlate customer needs with process requirements.</u> b) Define the focus of the Six Sigma project. c) Study possible improvements to the process. d) Establish priorities and define responsible for each one. e) Analyze the losses in the production process.</p>
Apply	<p>In a Kano Diagram, a gourmet biscuit offered to customer free of charge to accompany a cup of cappuccino currently is:</p> <p>a) A performance attribute (more is better) b) A basic attribute (must have) c) A good investment d) A mandatory characteristic <u>e) A delighter, an excitement attribute</u></p>	<p>For the production process of a cell phone, which signals would represent the interaction between the following characteristics on the roof of the Quality House (correlation matrix)?</p> <p>1. Touch screen and screen size; 2. Battery type and appliance size; 3. Screen size and size of the device</p> <p>a) 1. - / 2. + / 3. + b) 1. - / 2. + / 3. - <u>c) 1. + / 2. - / 3. -</u> d) 1. - / 2. - / 3. - e) 1. + / 2. - / 3. +</p>
Analyze	<p>Look at the image of the Kano Diagram and discuss which of the options is correct regarding the performance characteristics of a product:</p>  <p>a) Curve 2 corresponds to the basic characteristics (must have) of a product because the better, the more satisfied consumers will be. <u>b) Curve 1 shows the delights. Even if it is not done perfectly, the customer will be satisfied because it is something that was not expected of the product.</u> c) The product's performance is shown in curve 1. The better, the more satisfied the customer stays, because that's what he expects of the product. d) In curve 3 are shown the product delights. The attractions should be perfect to make the customer satisfied. e) In curve 2 are shown the Attractive attributes (delighter). The more, the greater the customer satisfaction.</p>	<p>Regarding the Goal statement, which of the following is NOT within the appropriate standard for Six Sigma projects?</p> <p>a) By the end of 2013, reduce operating costs in line 2 by 15%, falling below competitors. b) Reduce the failure rate in the quality test of products from line 1 from 15% to X% by the end of November. c) Increase production of line 1 from 1000 to 1500 a day without changing the amount of rework by the end of this two-month period. d) By the end of the first half of this year, reduce payment delays from 25% to 10%. <u>e) Reduce the percentage of delay of national deliveries by the end of this year.</u></p>
Create	<p>Develop a Critical to Quality Tree (CTQ Tree) of the project you are involved.</p>	<p>Create a Problem Statement common to your work environment, estimating numeric values when necessary, and develop the project charter of your Six Sigma project.</p>

Figure 2. Examples of questions at all RBT levels for the DMAIC first step (define)

the 80 points needed to be approved. If failed, the student still has 15 days until the beginning of the next DMAIC stage to achieve the passing score. Students that fail successively are advised to start a new training.

3.3 Application of the evaluation model and results

To validate the proposed model, this work was applied to groups being trained in SS (DMAIC) at a Brazilian aircraft company. This company, as well as the outsourced training company, the instructors' and the students' names are not identified in this work for reasons of business confidentiality.

The tests were performed by 70 students, three GB classes with 20 students each and a BB class with 10 students. The results of the first 18 closed-ended questions were first analyzed to check whether the student had enough knowledge to perform the level of creation (Questions 19 and 20) and then move on to the next stage of the DMAIC roadmap.

Only students who achieved a final grade of 80 points could continue the training. Students were able to perform each evaluation three times in case of failures and the results considered here show the grades after all attempts (when necessary). Two types of analysis were made for the same data set: results by student (Table II) and results by question (Table III).

For the first type of analysis, the number of correct answers in each RBT level and the scores at the open-ended questions were filled by the instructor in the gray columns of Table II. This table partially illustrates the results for the first phase of the DMAIC roadmap, Define, and a similar table was built for each other phase. This data layout allows

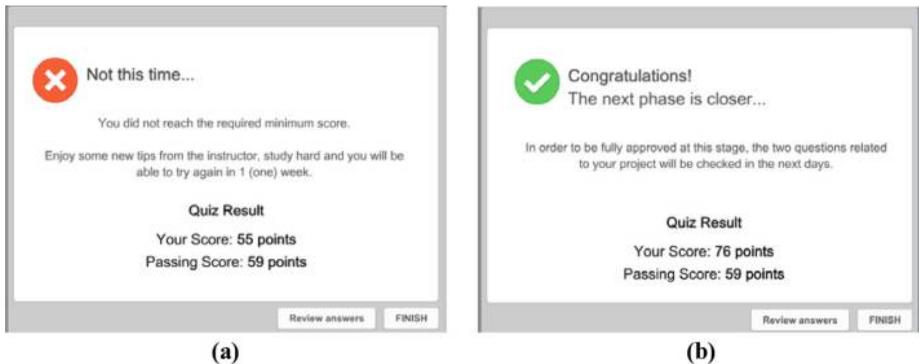


Figure 3. Message of being (a) failed or (b) approved at the closed-ended questions

	Student	Remember	Understand	Apply	Analyze	SCORE 1	Right	Wrong	Situation 1	Create	SCORE 2	Situation 2
	1	5	5	3	3	71	16	2	ap1*	19	90	Approved
	2	6	4	4	3	75	17	1	ap1	20	95	Approved
	3	6	5	4	3	79	18	0	ap1	21	100	Approved

	68	6	4	1	3	60	14	4	ap2**	20	80	Approved
	69	6	3	4	3	71	16	2	ap1	19	90	Approved
	70	6	5	3	3	74	17	1	ap1	19	93	Approved

Notes: *, ap1 = approved at the first attempt; **, ap2 = approved at the second attempt

Table II. Individual students' results at define phase of DMAIC

a deeper and richer analysis of the situation if compared with the usual case when only the final scores to simply pass or fail the students are given.

For the second type of analysis, the results were also classified according to the number of mistakes in each one of the 300 closed-ended questions, as partially shown in Table III. The columns highlighted in gray were filled by the instructor with the number of times each question randomly appeared in the tests, and with the number of mistakes the students committed. The last column refers to the feedback that was calculated to be used as an extra tool (see Section 4.3).

4. Statistical analysis of the results

All statistical analyses have been made using Microsoft Excel[®] and Minitab[®] 17 Statistical Software.

4.1 Student group as a whole

An overview of all the students' results is shown in Figure 4.

It can be seen in Figure 4 that the final grades of 64 students were between 80 and 100, high enough to be approved in all five DMAIC stages. Only six students failed at the three chances and left the course midway. Two students left in the first phase, define (light blue), one in the measure phase (orange) and three in the improve phase (yellow).

Hariharan (2014) categorized the SS projects into three broad types: quality improvement, revenue enhancing and cost saving. Figure 5 provides an overview of the student's results partially shown in Table II with a stack bar chart relating type of project, conclusion and training class for GBs and BBs.

A quick look at Figure 5 shows the following:

- Most projects of GBs deal with the improvement of quality and those of BBs are more concerned with increasing profits.
- There was a significant number of accomplished projects in both groups.

To test whether the highest graded students performed equally in terms of accomplishing their projects, the completion of the project was assigned the value 1 and the non-completion was assigned the value 0. A binary logistic regression was then carried out (Table IV).

On the basis of these results, the null hypothesis of the final score not influencing the conclusion of the project can be rejected at a 95 per cent confidence level and p -value of 0.003. It is reasonable to state that students with higher scores in the evaluation completed more projects within the training time.

Moreover, it came to the authors' knowledge that although some projects have not been concluded by the end of the training program, those continued to be carried on. It is likely that many of those might have been completed after that period.

Question	DMAIC	Bloom	Content	Use	Mistakes	Feedback
1	D	Remember	Overview define	13	0	0
2	D	Remember	Overview define	12	2	1
3	D	Remember	Overview define	15	0	0
...
308	C	Analyze	1. Statistical process control	54	23	3
309	C	Analyze	1. Statistical process control	56	16	2
310	C	Analyze	7. Hypothesis test	51	3	0

Table III.
Students' results organized by question

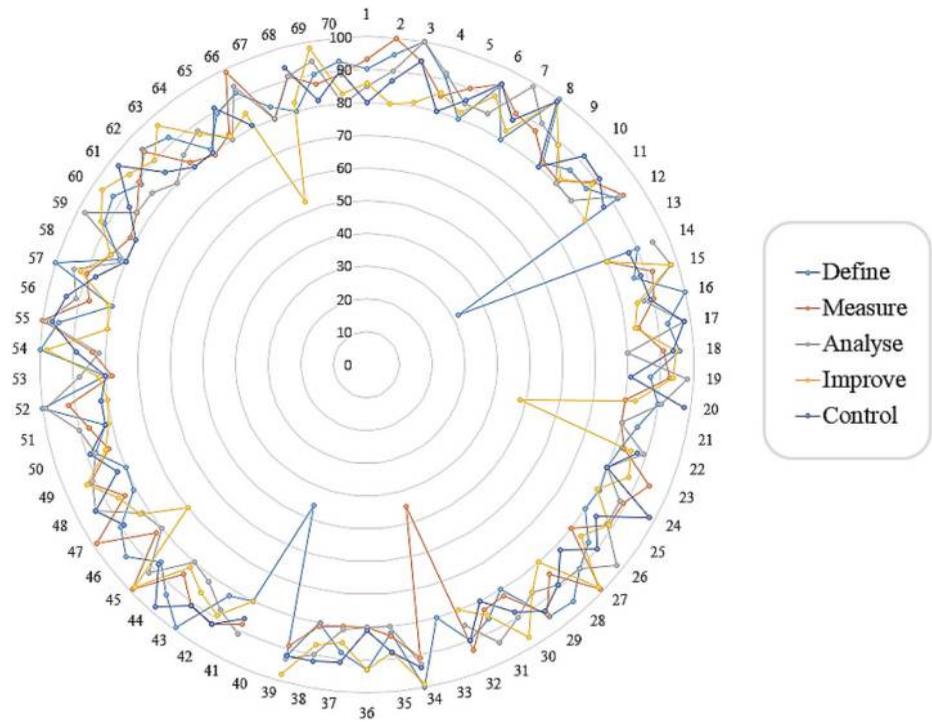


Figure 4.
Overview of the 70
student's grades

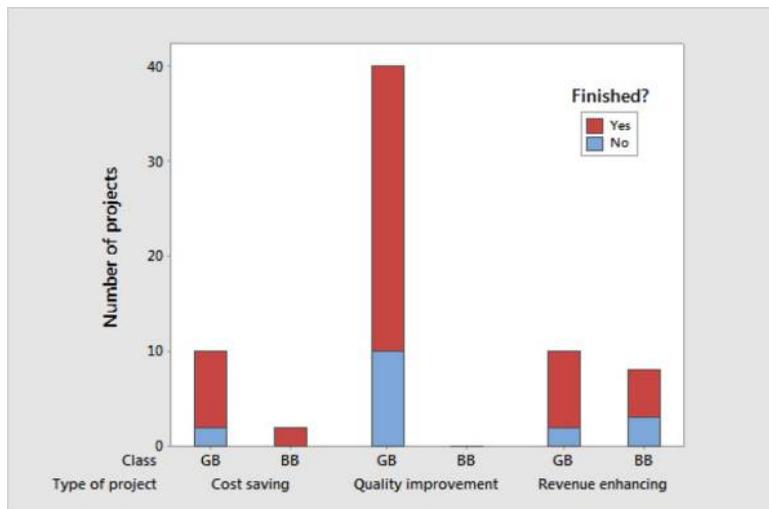


Figure 5.
Number of
accomplished
projects considering
project category and
type of training

4.2 Results organized by question

Figure 6 displays the box plot of all data, which revealed that the distribution of the questions that students missed is suggestive of stratification and trending. Following points can be concluded:

- The number and variances of mistakes in the different stages of DMAIC seem to exhibit the same tendencies.
- The same does not occur for different levels of RBT.

A proper statistical analysis is due to verify whether these results are correctly interpreted at the boxplot (see Table V for an ANOVA of the mistakes). This table provides the degree of freedom (DF) of each factor, the adjusted sums of squares (Adj. SS), the adjusted means squares (Adj. MS), the F -statistic from the adjusted means squares and its p -value.

Predictor	Coefficient	SE coefficient	Z	p -value
Constant	-45.0310	15.5189	-2.90	0.004
Final score	0.536471	0.180423	2.97	0.003

Table IV.
Logistic binary regression

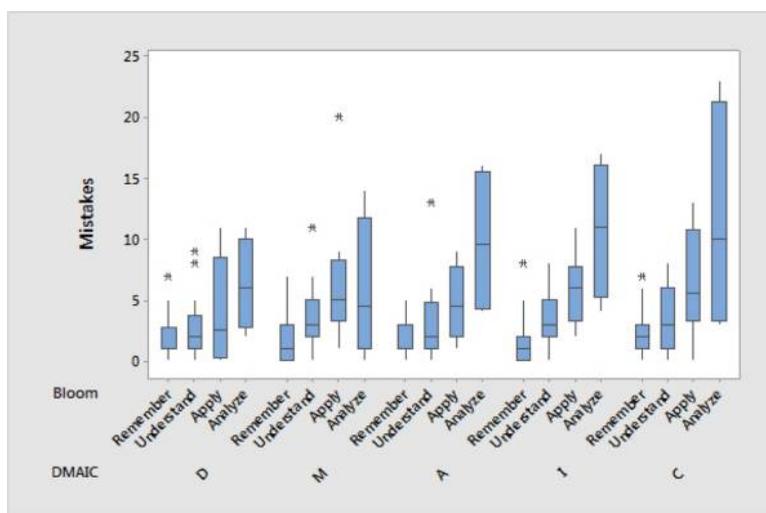


Figure 6.
Boxplot of mistakes for the different levels as of the RBT and for each stage of DMAIC roadmap

Source	DF	Adj. SS	Adj. MS	F	p -value
DMAIC	4	46.19	11.55	1.39	0.237
RBT	3	1,138.71	379.57	45.70	0.000
Error	292	2,425.27	8.31		
Lack-of-fit	12	118.55	9.88	1.20	0.283
Pure error	280	2,306.72	8.24		
Total	299	3,610.17			

Table V.
ANOVA table as given by Minitab® 17 statistical software

The lack-of-fit and the pure error are important to check the adequacy of the fitted model because an incorrect or under-specified model can lead to misleading conclusions. The F -statistic for this test is given by dividing Adj. MS for lack-of-fit by the Adj. MS for pure error. In this case, a p -value more than 0.05 for the lack-of-fit test indicates that the linear model adequately fits the response surface.

The ANOVA results are the following:

- In DMAIC case, the null hypothesis of equal values cannot be rejected at a 95% confidence level (p -value = 0.237).
- For the RBT, the null hypothesis of equal values can be rejected at 95% confidence (p -value < 0.050), i.e., the number of mistakes on each level of RBT can be considered significantly different – refer to the studies by [Biau et al. \(2010\)](#), [Johnson \(2013\)](#) and [Vidgen and Yasseri \(2016\)](#) for interesting reviews and critique on the use of p -values.

However, the number of mistakes on each one of the DMAIC phases can be considered equal.

Thus, it is fair to say that there is no much difference on the number of mistakes in all stages of DMAIC roadmap, so the training was uniform the entire roadmap, none of the stages was better explained than other. However, there are different amount of mistakes in each RBT level, showing that some skills need to be more stimulated than others.

When a deficiency in the development of some learning levels of the RBT is identified, how can the situation be reversed? Some researchers and instructors in fields related to engineering and management of public and private universities were interviewed to get suggestions on how to stimulate student learning at each RBT level. The results are:

- *Remember*: Encourage detailed reading of previous classes' summaries, the textbook used in training and any related materials focusing repeatedly on the most important details.
- *Understand*: Stimulate the production of theory schematic summaries and explain everything in detail using theoretical and practical examples.
- *Apply*: Solve exercises in class and provide extra exercises for students as homework.
- *Analyze*: Work out different examples in class, explaining the meaning of each variable and possible interactions among them, also highlighting the differences between each approach.
- *Evaluate*: Offer some group work with individual assessments or perform some practical work with multiple choice activities followed by discussions of the correct answers and explanations for the inconsistency of incorrect alternatives.
- *Create*: Provide opportunities for the application of theory in practical and real examples. Propose different job situations addressed in the classroom, present films and arrange for technical visits and extracurricular activities.

4.3 Feedback per question

A quantity named feedback was defined for each closed-ended question as the ratio of the number of mistakes to the total number of times each question appeared in the examinations, further normalized to the interval [0, 5] – see Column 7 of [Table III](#).

Questions for which the answers were classified as 0 or 1 do not point to a training problem as the majority of students got it right; for which answers were classified as 2 or 3 may indicate that there are still some doubts on the subject and it would be worth a review in the next class if there is opportunity and/or for the next training group; answers classified as 4 or 5 must be thoroughly reviewed to the groups because most students probably did not understand the question or concept. If the same question represents a problem for different training classes over time, the instructional material may be confusing and/or the training may require some methodological changes.

Another way to check against feedback is the use of control charts of the mistake percentage. The questions identified as out of control deserve revision by the instructor as well as the approach of the related content in the classroom.

Figure 7 shows the control chart of mistakes for each of the steps at the DMAIC roadmap: define (Panel a), measure (Panel b), analyze (Panel c), improve (Panel d) and control (Panel e). For all five panels, the red lines represent the upper and lower control limits, which are variable because they concern, respectively, the maximum and minimum proportion of errors acceptable for the questions that randomly appeared in different amounts. The blue lines represent the proportion of students' mistakes on each question.

All questions pointed as out of control and also some other noteworthy details on each panel are shown in Table VI.

5. Conclusions and recommendations

A new model for the SS learning assessment using RBT is proposed. It has been designed so that the students do not proceed unless they have acquired the conceptual knowledge at each step of the DMAIC roadmap, demonstrating that they are able to use it in their own projects. Several tools are presented to help instructors in evaluating the instructional material, the teaching methodology and the students themselves.

After each evaluation, the instructor had the opportunity to review individual problems and those of the entire class. At the end, the number of mistakes in all stages of DMAIC was proved equal, implying that the training was uniform the entire roadmap. However, the opposite happened in each RBT level, showing that some skills need to be better stimulated by the instructor than others.

Individual problems were isolated and solved on the basis of the levels of RBT together with attitudes and initiatives that must arise from the student. Exceptional cases find their causes in other factors than those described here: high-grade students who did not finish their projects during the training, those who could not accompany the group and the ones who evaded the course for personal problems are some examples.

In case of collective problems, there are various aspects to be taken into consideration, e.g. instructional material, approach to different themes, previous skill level of the students, instructor's teaching and students' background. To produce a real change in the situation, the instructors are advised to rethink their posture and the way they teach those particular subjects where failure was significant, or even redesign the entire program if the RBT analysis yielded poor results for the whole class.

The implementation of the proposed evaluation model allows a comparison of how the assessment of students in the SS training was made by this outsourced company before and after the introduction of the RBT. A critical analysis of the main differences between the traditional evaluation model and the model structured from the RBT proposed in this paper is shown in Table VII.

This evaluation model can be applied to training assessment in various areas of knowledge. The authors noticed that, at the time of writing this paper, there had been very

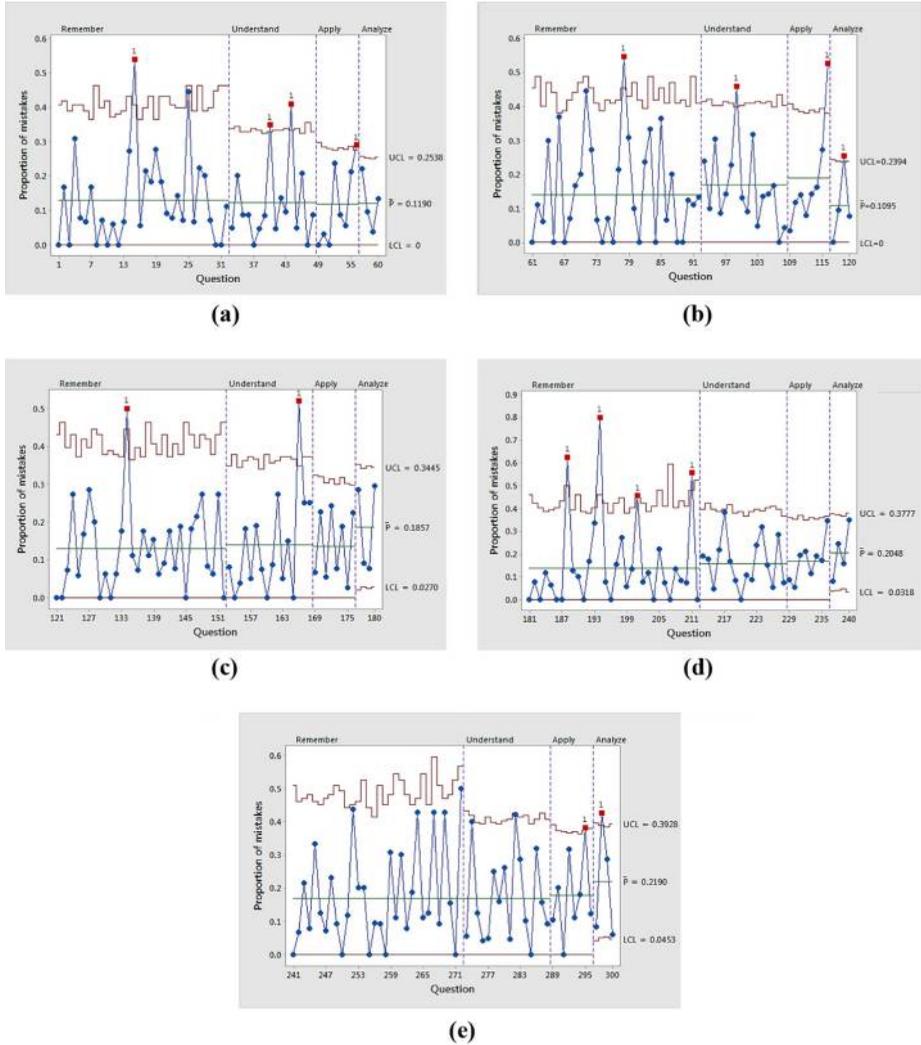


Figure 7. Control chart of the students' mistakes at the (a) define, (b) measure, (c) analyze, (d) improve and (e) control phases

Note: Red dots are points out of control

few publications using RBT to assess learning in the areas of SS or general topics of production engineering, a fact that reveals a gap in the literature, with many possibilities for the development of new research.

Importantly, this work is, in itself, the basic framework to be repeated or replicated in other subjects. As a continuation of this research, a software code has been developed to facilitate the instructor's hard work with the storage of students' grades in spreadsheets. This is essential to make full use of this new evaluation model mainly as a means of providing teaching feedback.

Panel	Stage	Question no.	Subject	RBT level
(a)	Define	15	House of quality	Remember
		40	Kano analysis	Understand
		44	House of quality	Understand
		56	Stakeholders analysis	Apply
(b)	Measure	18	Gage R&R	Remember
		39	Sampling	Understand
		56	Sigma shift	Understand
		59	Capability analysis	Analyze
(c)	Analyze	134	Letter measles	Remember
		166	Hypothesis testing	Understand
(d)	Improve	188	Chain letters and billboards	Remember
		194	C-paired comparisons	Remember
		201	FMEA	Remember
		211	Pilot studies	Remember
(e)	Control	295	Statistical control	Apply
		298	Statistical control	Analyze

Table VI.
Map of the outliers in
the control charts of
the mistakes'
occurrence (Figure 7)

Traditional evaluation	Proposed evaluation model
It depended on the instructor's experience with the training	It depends on the instructor's experience with the RBT
The correction procedure was subjective because it depended on the instructor's judgment, thus, it could be uneven and/or unfair	Uniformity and fairness correcting 79% of the evaluation
The instructor spent too much time correcting examinations	The correction is only performed by the instructor when the student reaches the last level of the learning domain
The student developed the SS project even without knowing the theory very well	The project is only evaluated when the student reaches a certain level of theoretical knowledge
It did not provide feedback on teaching	It provides statistical based feedback on teaching

Table VII.
Differences between
the traditional
learning evaluation
and the model
proposed here

Note

1. Hidden thesis to keep the review blind.

References

- Aboelmaged, M.G. (2010), "Six Sigma quality: a structured review and implications for future research", *International Journal of Quality & Reliability Management*, Vol. 27 No. 3, pp. 268-317.
- Anderson, L.W., Krathwohl, D.R., Airasian, P.W., Cruikshank, K.A., Mayer, R.E., Pintrich, P.R., Rath, J. and Wittrock, M.C. (2001), *A Taxonomy for Learning, Teaching, and Assessing - a Revision of Bloom's Taxonomy of Educational Objectives*, Longman, New York.
- Antony, J. and Bhajji, M. (2002), *Key Ingredients for a Successful Six Sigma Program*, University of Warwick, Warwick.
- Arends, R.I. (2012), *Learning to Teach*, 9th ed., McGraw Hill, New York, NY.
- Athanassiou, N., McNett, J.M. and Harvey, C. (2003), "Critical thinking in the management classroom: Bloom's Taxonomy as a learning tool", *Journal of Management Education*, Vol. 27 No. 5, pp. 533-555.
- Biau, D.J., Jolles, B.M. and Porcher, R. (2010), "*p* value and the theory of hypothesis testing: an explanation for new researchers", *Clinical Orthopaedics and Related Research*.

- Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., and, and Krathwohl, D.R. (1956), "Taxonomy of educational objectives - the classification of educational goals", *Handbook I: Cognitive Domain*, Vol. 16, Longman, New York, NY, p. 207.
- Brook, Q. (2010), *Lean Six Sigma and Minitab - the Complete Toolbox Guide for All Lean Six Sigma Practitioners*, 3rd ed., OPEX Resources, London.
- Coughlan, P. and Coghlan, D. (2002), "Action research for operations management", *International Journal of Operations & Production Management*, Vol. 22 No. 2, pp. 220-240.
- Eckes, G. (2003), *Six Sigma for Everyone*, John Wiley & Sons, Hoboken, NJ.
- Folaron, J. (2003), "The evolution of Six Sigma", *Six Sigma Forum Magazine*, pp. 38-44.
- Galhardi, A.C. and Azevedo, M.M. (2013), "Avaliações de aprendizagem: o uso da Taxonomia de Bloom", VIII Wokshop de Pós-Graduação E Pesquisa Do Centro Paula Souza, *São Paulo*, pp. 237-247.
- Gaujard, C., Verzat, C. (2011), "Former à la créativité . . . un pari insensé?", *Entreprendre & Innover*, Vols 11/12 No. 3, pp. 137-146.
- Gronlund, N.E. (1974), *A Elaboração de Testes de Aproveitamento Escolar*, Editora Pedagógica e Universitária, São Paulo.
- Guskey, T.R. (2001), *Benjamin S. Bloom's Contributions to Curriculum, Instruction and School Learning*, Lexington, KY.
- Hariharan, A. (2014), *Continuous Permanent Improvement*, ASQ Quality Press, Vol. 1.
- Harry, M. and Schroeder, R. (2000), *Six Sigma - the Breakthrough Management Strategy Revolutionizing the World's Top Corporations*, 1st ed., Doubleday, New York, NY.
- Hoerl, R.W. (2001), "Six Sigma black belts: what do they need to know?", *Journal of Quality Technology*, Vol. 33 No. 4, pp. 391-406.
- Ingle, S. and Roe, W. (2001), "Six Sigma black belt implementation", *The TQM Magazine*, Vol. 13 No. 4, pp. 273-280.
- Jansen, B.J., Booth, D. and Smith, B. (2009), "Using the taxonomy of cognitive learning to model online searching", *Information Processing & Management*, Vol. 45 No. 6, pp. 643-663.
- Jesus, E.A.D. and Raabe, A.L.A. (2009), "Interpretações da Taxonomia de Bloom no contexto da programação introdutória", XX Simpósio Brasileiro de Informática Na Educação.
- Johnson, V.E. (2013), "Revised standards for statistical evidence", *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 110 No. 48, pp. 19313-19317.
- Kirkpatrick, D.L. and Kirkpatrick, J.D. (2006), *Evaluating Training Programs - the Four Levels*, 3rd ed., Berrett-Koehler Publishers, San Francisco, CA.
- Krathwohl, D.R. (2002), "A revision of Bloom's Taxonomy: an overview", *Theory into Practice*, Vol. 41 No. 4, pp. 212-218.
- Krathwohl, D.R., Bloom, B.S. and Masia, B.B. (1964), *Taxonomy of Educational Objectives, the Classification of Educational Goals. Handbook II: Affective Domain*, David McKay Company, New York, NY.
- Krau, S.D. (2011), "Creating educational objectives for patient education using the new Bloom's Taxonomy", *The Nursing Clinics of North America*, Vol. 46 No. 3, pp. 299-312.
- Larkin, B.G. and Burton, K.J. (2008), "Evaluating a case study using Bloom's Taxonomy of education", *AORN Journal*, Vol. 88 No. 3, pp. 390-402.
- Linderman, K., Schroeder, R.G., Zaheer, S. and Choo, A.S. (2003), "Six Sigma: a goal-theoretic perspective", *Journal of Operations Management*, Vol. 21 No. 2, pp. 193-203.
- Lister, R., Adams, E.S., Fitzgerald, S., Fone, W., Hamer, J., Lindholm, M., McCartney, R., Mostrom, J.E., Sanders, K., Seppala, O., Simon, B. and Thomas, L. (2004), "A multi-national study of reading and tracing skills in novice programmers", *Working Group Reports from ITiCSE on Innovation and Technology in Computer Science Education*, Vol. 36, New York, NY, pp. 119-150.

- Montgomery, D.C. and Woodall, W.H. (2008), "An overview of Six Sigma", *International Statistical Review*, Vol. 76 No. 3, pp. 329-346.
- Öçzelik, D.A., Aksu, M., Berberoglut, G. and Paykoç, F. (1993), "The use of the taxonomy of educational objectives in Turkey", *Studies in Educational Evaluation*, Vol. 19 No. 1, pp. 25-34.
- Pande, P.S., Neuman, R.P. and Cavanagh, R.R. (2000), *The Six Sigma Way - How GE, Motorola and Other Top Companies Are Honing Their Performance*, McGraw-Hill, New York, NY, available at: <https://doi.org/10.1036/0071376674>
- Pappas, E., Pierrakos, O. and Nagel, R. (2013), "Using Bloom's Taxonomy to teach sustainability in multiple contexts", *Journal of Cleaner Production*, Vol. 48, pp. 54-64.
- Phillips, A.W., Smith, S.G. and Straus, C.M. (2013), "Driving deeper learning by assessment: an adaptation of the revised Bloom's Taxonomy for medical imaging in gross anatomy", *Medical Student Education*, Vol. 20 No. 6, pp. 784-789.
- Salah, S., Rahim, A. and Carretero, J.A. (2010), "The integration of Six Sigma and lean management", *International Journal of Lean Six Sigma*, Vol. 1 No. 3, pp. 249-274.
- Schroeder, R.G., Linderman, K., Liedtke, C. and Choo, A.S. (2008), "Six Sigma: definition and underlying theory", *Journal of Operations Management*, Vol. 26 No. 4, pp. 536-554.
- Snee, R.D. (2010), "Lean Six Sigma – getting better all the time", *International Journal of Lean Six Sigma*, Vol. 1 No. 1, pp. 9-29.
- Tyran, C.K. (2009), "Designing the spreadsheet-based decision support systems course: an application of Bloom's Taxonomy", *Journal of Business Research*, Vol. 63 No. 2, pp. 207-216.
- Veeravagu, J., Muthusamy, C., Marimuthu, R. and Michael, A.S. (2010), "Using Bloom's Taxonomy to gauge students' reading comprehension performance", *Canadian Social Science*, Vol. 6 No. 3, pp. 205-212.
- Vidgen, B. and Yasseri, T. (2016), "p-values: misunderstood and misused", *Frontiers in Physics*, Vol. 4, pp. 10-14.
- Whalley, J.L., Lister, R., Thompson, E., Clear, T., Robbins, P. and Prasad, C. (2006), "An Australasian study of reading and comprehension skills in novice programmers, using the Bloom and SOLO taxonomies", *8th Australasian Conference on Computing Education*, Vol. 52, *Australian Computer Society, Darlinghurst*.

About the authors

Gabriela Fonseca Amorim holds a degree in control and automation engineering and a master's degree in industrial engineering, both from Universidade Federal de Itajubá – UNIFEI, Brazil. She was a Research Scholar at the Université de Technologie de Compiègne – UTC (2009), France, and at the University of Tennessee at Knoxville – UTK (2015/2016), USA. She is a Black Belt in Six Sigma and is currently a PhD candidate at the Industrial Engineering Program at UNIFEI. Her research interests include statistics, Six Sigma and design for Six Sigma. Gabriela Fonseca Amorim is the corresponding author and can be contacted at: amorimgabi@yahoo.com.br

Pedro Paulo Balestrassi received DSc degree in industrial engineering from Universidade Federal de Santa Catarina (Brazil) in 2000. Since 1994, he has worked at Universidade Federal de Itajubá (Brazil) as a Professor at the Industrial Engineering Department. In 1998/1999, he was a visitor at Texas A&M University (USA); in 2005/2006, he was a visiting professor at the University of Texas (USA); and in 2010/2011, he was a visiting professor at the University of Tennessee (USA). He is a Master in Black Belt in Six Sigma and his areas of interest include time series forecasting, artificial neural networks and statistics.

Rapinder Sawhney is a Heath Fellow in Business and Engineering at the University of Tennessee at Knoxville (UTK), USA. He is a PhD in Engineering Science and Mechanics, an MS in Industrial Engineering and a BS in Industrial Engineering, all three by UTK. His research interests include designing efficient and reliable systems, using concepts of natural interaction for work design and defining the value of information in supply chains.

Mariângela de Oliveira-Abans graduated in physics from Universidade de São Paulo – USP, Brazil (1977), and holds a master’s degree in astronomy from USP (1985). She is a Researcher at the Laboratório Nacional de Astrofísica, Brazil, and presently acts as a Brazilian liaison at Gemini Observatory’s Public Information and Outreach Network, Brazilian liaison at the SOAR Telescope’s Public Affairs and Educational Outreach Network and as a Brazilian Manager at the Canada–France–Hawaii Telescope. She is currently pursuing her PhD at the Instituto de Engenharia de Produção e Gestão of the Universidade Federal de Itajubá – IEPG/UNIFEI, in the area of quality.

Diogo Leonardo Ferreira da Silva is a control and automation engineer (2010), a MSc in Electrical Engineering (2013), and a PhD in Electrical Engineering (2013), all three from Universidade Federal de Itajubá – UNIFEI, Brazil. In 2015/2016, he worked as a Research Scholar at the University of Tennessee at Knoxville, USA. Since 2017 he is a professor at UNIFEI and his main areas of interest include control systems, automation and control of processes and electronic circuits.