EDUR • Educação em Revista. 2018; 34:e202974
DOI: http://dx.doi.org/10.1590/0102-4698202974

ARTICLE

VALIDATION OF SCIENTIFIC LITERACY ASSESSMENT INSTRUMENT FOR HIGH SCHOOL GRADUATES IN TRAFFIC APPLIED PHYSICS CONTEXT

PATRICK ALVES VIZZOTTO[1 *]
ORCID: https://orcid.org/0000-0003-1613-4858
LUIZ FERNANDO MACKEDANZ[1 **]
ORCID: https://orcid.org/0000-0002-0138-0119

ABSTRACT: We present the validation procedure of an instrument to verify aspects of Scientific Literacy for High School graduates. The questionnaire was divided into questions about concepts of applied Physics in traffic situations and on the profile characters of the physics classes attended by the graduates during High School. Initially, it counts 64 items, that were assessed for their content by teachers in the area and later, through statistical tests, the quantitative results for validation and reliability were estimated. Finally, we remove 19 items and a valid and reliable instrument was obtained with 25 items on traffic applied Physics and 20 items on the Physics Education profile, for use in scientific literacy assessment.
Keywords: Traffic Applied Physics; Physics Teaching; Knowledge retention.

VALIDAÇÃO DE INSTRUMENTO DE AVALIAÇÃO DA ALFABETIZAÇÃO CIENTÍFICA PARA EGRESSOS DO ENSINO MÉDIO NO CONTEXTO DA FÍSICA DO TRÂNSITO

RESUMO: Apresentamos o procedimento de validação de um instrumento para verificação de aspectos da Alfabetização Científica para egressos do Ensino Médio. O questionário dividiu-se em questões sobre conceitos da física aplicada em situações do trânsito e sobre características do perfil das aulas de Física assistidas pelos egressos durante o Ensino Médio. Contando inicialmente com 64 itens, os mesmos foram avaliados quanto ao seu conteúdo por professores da área e posteriormente, através de testes estatísticos, estimou-se os quantitativos de validação e confiabilidade. Ao final, 19 itens foram removidos e obteve-se um instrumento considerado válido e fidedigno com 25 itens sobre física no trânsito e 20 itens sobre o perfil do Ensino de Física, com fins de utilização para avaliação da alfabetização científica.
Palavras-chave: Física aplicada ao trânsito. Ensino de Física. Retenção do conhecimento.

[1] Rio Grande do Sul Federal University, Porto Alegre, RS, Brasil.

[*] Doctorate in the Graduate Program in Science Education at Rio Grande do Sul Federal University – UFRGS. E-mail: < patrick.vizzotto@ufrgs.br > .

[**] Doctor on Science from Rio Grande do Sul Federal University – UFRGS. Associated Professor in Federal University of Rio Grande – FURG. Graduate Program in Science Education – UFRGS.
E-mail: < luismackedanz@furg.br > .

INTRODUCTION

What do the graduates remember about High School classes? How much school knowledge will be used by former alumni to help them in their readings about the world? Does this knowledge enable to understand the dynamics of daily life, so that there is greater possibility that your decisions are critical and responsible? Such questions arise when comparing the aims presented in Public Education Policies documents with the school reality that is commonly observed today.

It is undeniable that the society has its daily life permeated by science and technology, and its immersion in this environment has been exponentially growing, since many of the technological devices used today were not known or accessible to the majority of the population in previous generations. Even so, this exponential growth signals that the future society will enjoy the benefits of science and technology (as well as the problems that these advances can cause).

Whether now or in the future, it is imperative that the school will be one way for Scientific Literacy growth, in order that citizens can understand the dynamics of the world and act responsibly and critically within this reality. We argue that a scientifically literate individual is one who masters the basic knowledge about science and manages to use them to understand everyday phenomena, technological artifacts, and act in society.

These questions and concerns is part of the Scientific Literacy ideology (AULER; DELIZOICOV, 2001). In this sense, we believe that is desired for a scientifically literate individual to possess skills and competencies necessary in order to be a citizen prepared for life in various sectors of society, like work, family, or the handling of electronic devices, or, for example, in order to attest the fate of state-owned resources in its own population. In Santos and Mortimer (2001) view, this can result in decision making for responsible social action by the population.

In order to measure the scientific literacy level of High School students and graduates, some criteria for perform assessment were debated within the science education area, and thus, research instruments were elaborated (AAAS, 1989; MILLER, 1983; LAUGKSCH, SPARGO, 1996).

Studies within the field of Psychology, such as those conducted in the 19th century by Ebbinghaus, also questioned about the degree of retention of knowledges. In particular, there are investigations involving school knowledge, especially Pinto (1989, 1991, 2001), and it is possible to notice a oblivion curve for certain learning induced in the research.

This article describes the process of constructing and validating a questionnaire which aims to evaluate the performance of High School graduates in using their knowledge for interpreting Physics situations, present in a specific daily life: the traffic. In addition to this assessment, we also try to describe the profile of the Physics courses of these graduates, in order to analyze if there is a correlation between this profile and its proficiency in the interpretation of physics situations in traffic.

This work aims to contribute for the literature of the area to present a quantitative way to assess Physics proficiency of High School graduates, based on the use of their knowledges to interpret daily life situations, not focused in concepts memorization and mathematical procedures models, but rather through an analysis of competences and acquired skills level for judge scientific situations applied in everyday life.

This questionnaire has as target audience High School graduates in the following categories: students entering technical and undergraduate courses; pupils from driver training centers; professional drivers and trainers from the traffic area. The validation procedure of the instrument was based on statistical techniques such as difficulty analysis and item discrepancy, internal consistency test and exploratory factorial analysis, obtaining a final, reliable and valid questionnaire with 25 questions of traffic applied Physics and 20 questions about Physics teaching profile.

The article begins by addressing the concept of Scientific Literacy and the productions related to the chosen clipping for this work; in a second moment, we justified the choice of the topic "traffic" as the daily situation used in the questionnaire; later, the construction process and the validation methods are described; and in the last two sections present the pilot test results, its respective discussions and closing remarks.

SCIENTIFIC LITERACY AND THE RETENTION OF KNOWLEDGE OF HIGH SCHOOL GRADUATES

In this article, based on the Sasseron and Carvalho (2011) conception, we adopt the understanding that a scientifically literate citizen is one who mastered the basic knowledge about science and can use it to understand everyday phenomena, technological artifacts and act critically and responsibly in society. Consistent with Auler and Delizoicov (2001), we believe that the contents covered by the Natural Science subjects are considered as ways to understand socially prominent daily topics.

A scientifically literate attitude may enable the individual to acquire a critical view when interpreting scientific policies, which would ensure the monitoring of actions for democratic participation in the decision making and state-owned investments (SANTOS; MORTIMER, 2000; ANGOTTI; AUTH, 2001).

The term "Scientific Literacy" has many interpretations. In general, the term is used as a synonym for "public understanding of science". For Vidor et. al. (2009), this definition relates to what individuals should know about science and technology, although different authors include notions involving individual behaviors, such as intellectual habits and mental abilities that allow the use of scientific knowledge to solve problems and make decisions in everyday situations.

The term was originally used by Paul Hurd (HURD, 1958) , and since then the literature has presented an expressive amount of publications in the area, such as Miller (1983); Fourez (1994); Jenkins (1994) ; Laugksch and Spargo, (1996); Laugksch (2000); Hurd (1998); Gil-Pérez and Vilches-Peña (2001); Lorenzetti (2000); Auler and Delizoicov (2001); Chassot (2000, 2003); Nascimento-Schulze et. al. (2006); Vidor et. al. (2009); Camargo et. al. (2011); Sasseron and Carvalho (2011), in which it can realize slightly different views on the meaning of this term. So:

> It can be seen that at the heart of the discussions of authors who use one term or another, there are the same concerns about the Science Teaching, that is, reasons that guide the planning of this teaching to the construction of practical benefits for people, society and environment (SASSERON, CARVALHO, 2011, p. 60).[1]

Nascimento-Schulze et. al. (2006), by the way, claim that these differences are also adopted according to different purposes for which the concept is used. However, the authors point out that although there are differences in the meaning, there is also a thread that guides the structure of understanding and use of the term : *"it is a desirable goal for all citizens, measurable and useful for the everyday life and strongly linked to the social context*"[2] (p.26). As a foundation for the definition adopted by us, we use the concept postulated by Paulo Freire, when he states that:

> Literacy is more than the simple psychological and mechanical domain of writing and reading techniques. It is the domain of these techniques, in conscious terms. It is to understand what one reads and write what one understands. It is communicating graphically. It's an embodiment. It implies not a visual and mechanical memorization of sentences, words, syllables, torn from an existential universe - dead or half-dead things - but in an attitude of creation and re-creation, implies a self-formation that may result in a man's interfering posture on his context. (FREIRE, 2000, p 118).[3]

Therefore, someone who is scientifically literate will have the skills to understand the dynamics of the world in which he/she is inserted. This includes understanding the nature of science and how it develops; the scientific concepts applied in daily phenomena; and also understand the influence of science and technology on society and the environment.

The elaboration of a questionnaire to measure the level of Scientific Literacy in High School graduates, was boosted by Miller's work (1983), which conceived the act of being scientifically literate is a phenomenon that can be certificated verifying the domain of the individual in three dimensions: Understanding the nature of science; Understanding of scientific contents; and understanding of the impact of science and technology on society.

Based on these definitions, the American Association for the Advancement of Science - AAAS started in the 1980s the Project 2061, which aimed to collaborate for the Scientific, Mathematical and Technological Literacy of US citizens, having as concrete action the publication of Science for All Americans (SFAA) guide. This journal listed notions that students should have at the end of basic education, which included both knowledge and mental skills related to science and mathematics disciplines (AAAS, 1989).

This need to devise ways to measure scientific literacy came from the very imposition of these project development agencies, in order to verify the effectiveness of the investments made (NASCIMENTO-SCHULZE, CAMARGO, WACHELKE, 2006; VIDOR et al., 2011). However, we assert that the fact of quantitatively measuring a construct such as learning, is a topic considered complex and may generate discord between theoreticians from different areas.

In order to evaluate the effect of these initiatives on Basic Education graduating students, Laugksch and Spargo (1996) created and validated an instrument for this purpose, testing the minimum characteristics that a basic school graduate should possess to be considered scientifically literate. This questionnaire was called Test of Basic Scientific Literacy - TBSL and its validation procedure interviewed more than 4000 High School graduates from South Africa. In Brazil, this instrument was translated by Nascimento-Schulze (2006) and is called *Teste de Alfabetização Científica Básica - TACB*. Composed of 110 dichotomous items (true or false), it is divided into three subtests that evaluate the different dimensions of Scientific Literacy postulated by Miller (1983).

However, the developers themselves affirm that it has certain limitations, because it only evaluates the basic aspect of Scientific Literacy, that is, it is estimated only knowledge from interdisciplinary concepts, because the items do not test the individual skill for apply their knowledge in situations of problem solving and decision making :

> TBSL, therefore, is testing only fundamental or basic aspects of scientific literacy, such as knowledge of interdisciplinary concepts; applications of science and the ability to apply knowledge to decision making and problem solving are not being tested here (LAUGHSCH, SPARGO, 1996, p.335).[4]

The questionnaire developed and validated in this article aims to present an option to fill this gap for the Physics area, deepening the act of evaluating the Scientific Literacy of High School graduates, by enabling the interviewee to use their knowledge to interpret everyday situations in a particular context for applied Physics: Traffic. In addition, it is believed to be relevant to investigate how much the school reality may or may not influence the graduate skill for interpret such situations.

Although a citizen can acquire knowledge in various environments he/she can attend, the relevant role of the school as a Scientific Literacy formal space is undeniable. We also understand that the guiding objectives of Brazilian educational system, through its official documents (BRASIL, 1996, 2000, 2017), seek to reach, at the end of basic education, between different skills and competences, a minimum level of Scientific Literacy, even if these purposes are not explicitly referenced in any of these documents.

Based on the school objective of forming a critical citizen, we argue about the extent which school knowledge can effectively be used to carry out a world critical reading and, consequently, increasing the chance that decisions of greater responsibility and criticality are taken by the population. How much of the content taught in school is retained in the cognitive structure of the students, for how long and what the level of retention of school knowledge do the graduates present?

Such questions have already been object of study by authors such as Thorndyke (1977); Tulving (1983); Pinto (1989; 1991; 2001); and Ebbinghaus (2013), which are focused on the observation of the amount of information stored as a function of time. The present work aims to contribute for measurement of the Scientific Literacy level from a different perspective than the so far produced literature in this area: the possibility of the interviewee use their knowledge to analyze if everyday situations are correct or incorrect from the Physical point of view.

It is believed that there is a possibility that these knowledge are to be used in real life situations, which decision-making is required, justifying the need to evaluate graduates in order to conduct research that seek to realize how much school knowledge is retained and can be used in everyday reviews.

Considering the subdivisions proposed by Miller (1983) to explain the Scientific Literacy phenomenon, it is emphasized that this work aims to propose a test that measures the proficiency in Physics, specifically for one of the above mentioned pillars, the "Knowledge of the content of Science". We understand that the other axes proposed by Miller (1983), besides being interdisciplinary, are satisfactorily contemplated in the TBSL.

It is also important to clarify why it has chosen the theme "Traffic" as an everyday situation in the questionnaire. According to Ambev (2017), in the year 2017 there were 39.333 deaths in traffic accidents throughout Brazil. This is equivalent to a hypothetical plane crash per day, killing 107 people in each event, just in the country, for an entire year. What would be considered an unquestionable tragedy in aviation, it may go unnoticed or is naturalized in traffic context, because in this case the death happen in a not concentrated way.

Based on this, and knowing that most of the traffic accidents happen due to human failures, most of the time caused by the lack of education for traffic, according to Negrini-Neto and Kleinubing (2012), it is evident the relevance of the school in the development of a level of individual criticism, which can result in conscious attitudes of cause and effect relationships that any action can trigger.

Within the Physics Teaching, is remarkable the various situations in which scientific concepts can be related to traffic (CHAGAS, 2014; URRUTH; STEFFANI; SILVEIRA, 2015; VIZZOTTO; MACKEDANZ, 2017) , reason why it is believed that this cut of thematic is relevant to attest the levels of understanding and analysis of graduates in Physics . Above all, it is argued that Physics Teaching can contribute to form future drivers and conscious pedestrians.

ELABORATION AND METHODS OF VALIDATION OF THE QUESTIONNAIRE

In this section, we present and discuss our procedure for constructing the research instrument, seeking to address the aspects discussed in the previous section of the text. For this, we will present sequential items, obeying a temporal organization of the process.

CONSTRUCTION OF THE QUESTIONNAIRE

The instrument has two parts, each with a specific objective. In the first one, it is presented items with traffic applied Physics situations, which may be correct or incorrect from the Physical point of view, allowing the interviewee to agree or disagree with the statement contained in each question. In the second one, it is collected information about the characteristics of Physics Teaching during High School.

In order to supply this second goal, we adopted as basis the Significant Learning Theory (AUSUBEL, 2003) to elaborate a scale with the capacity to differentiate, based on the interviewee's perception, characteristics of classes that could enhance or disempower such knowledge acquisition.

The choice of this theory is grounded in our understanding that a significant learning in Physics (and all areas) contents can contribute to individual Scientific Literacy, being therefore important to observe if the graduates' school reality had minimal characteristics for such a process to happen.

The traffic applied Physics items were elaborated based on the phenomena discussed in the handbook used by the drivers' training centers in State of Rio Grande do Sul (OLMA, 2016), as well as in the Physical concepts commonly addressed in the subareas of school Physics.

Such items have options of answer in a dichotomous (true or false) format, and the interviewee can mark a question mark if he/she does not know what to answer, so that this strategy can inhibit the random choice. Items returned with the question mark are considered to be wrong in the analysis process. This strategy was also applied by Laugksch and Spargo (1996) for TBSL respondents, and repeated here by considering a tactics that can be effective in reducing the number of "guesses".

By the way, the items about Physics Teaching profile were based on the literature of the field (AUSUBEL, 2003; TAVARES, 2004; MOREIRA, 2006; DARROZ et al, 2015), being arranged in a Likert scale with five response options: I strongly agree; I agree; neither I agree nor disagree; I disagree; I totally disagree.

The creation of the instrument items obeyed the criteria defined by Pasquali (1998). Each of the subdivisions of the instrument generates a score for each interviewee, representing a value within a scale that represents his/her Physics Teaching profile.

First, we created 62 items of traffic applied Physics and 40 items on Physics Teaching profile. In the first set, 32 are correct and 30 are incorrect. In the second, 20 describe characteristics of classes that can generate significant learning and another 20 with characteristics of classes that may inhibit such learning.

CONTENT VALIDITY

Prior to the application of the pilot test, we sought to verify the content validity of the items. In order to perform this, they were sent to 18 PhD teachers of Physics Teaching area, which we managed here as referees, who were able to analyze the clarity, language, relevance, semantic adequacy and consistency of the issues. This evaluation obtained return from 10 referees, a number that both Alexandre and Coluci (2011) and Pasquali (2017) consider perfectly acceptable for this initial evaluation procedure.

This evaluation happened through a spreadsheet, where the item could be considered as: Not relevant or not representative; Item needs major revision to be representative; Item needs small revision to be representative; and Relevant or representative item. This classification was used to calculate the Content Validity Index (IVC) proposed by Alexandre and Coluci (2011). For the authors, the IVC:

> It measures the proportion or percentage of judges who agree on certain aspects of the instrument and its items. It allows you to initially analyze each item individually and then the instrument as a whole. This method employs a Likert scale with an one-to-four score, in order to evaluate the relevance/representativeness of each item.[5] (ALEXANDRE, COLUCI, 2011, p.3065).

For each above mentioned alternative, it was assigned an one-to-four score, respectively. After the return of the referees, the number of "note 4" evaluations (Relevant or representative Item) was counted for each of the items. Then, the resulting sum is divided by the number of evaluators and thus arrived at the aforementioned ratio.

Therefore, the formula for evaluating them individually consisted of:

$$IVC = \frac{Nr}{Nt}$$

where "Nr" and "Nt" mean the number of "notes 4" scores and the total number of referees, respectively. This index generated a number from 0 to 1 for each item. For the pilot test, only those that generated a value equal to or greater than 0.9 were used.

After the selection of the items considered representative, corrections were suggested by the referees in order to adjust aspects of writing in items that had such a need.

From the initial set for traffic applied Physics (62 items), 38 were part of the pilot test, being 20 items with correct situations and 18 incorrect. From the Physics Teaching profile initial set (40 items), 26 were part of the pilot test, 14 items of characteristics that support significant learning and 12 with characteristics that may inhibit such learning.

VALIDITY AND RELIABILITY OF ITEMS

As long as the traffic applied Physics items response format is different from the items that describe the Physics Teaching profile, assessment procedures of the validity and reliability were conducted separately. After the application of the pilot test, the questionnaires were tabulated in a spreadsheet and later tested in the statistical software Statistical Package for the Social Sciences - SPSS, version 23 for Windows.

It is understood that "*an instrument is valid in the extent to which it measures what it is proposed to measure. For example, a valid instrument to measure reading ability should really measure this characteristic and not others, such as prior knowledge*[6] (MARTINS, 2006, p.5). In its turn, reliability is attested if measurements are similar when applied to the same individual at different times. According Moreira and Rosa (2007) a reliable measure will produce equal results in successive applications for the same subject.

According to Figueiredo et. al. (2008), there are several ways to measure reliability (also called trustworthiness), but the difficulty of reproducing the same application conditions for two distinct moments in educational tests is recognized. Thus, for Moreira and Rosa (2007), one way of calculating instrument reliability is to determine its internal consistency. This indicates whether all sub-parts of le measure the same characteristic. According Souza et. al. (2017), most researchers measures the internal consistency of a questionnaire through Cronbach's alpha coefficient. It demonstrates the degree of variance between items on a scale. Thus, "*the smaller the sum of the variance of the items, more consistent the instrument is considered*[7] (p.651). This calculation generates a coefficient between 0 and 1, and values close to 1 mean greater internal consistency.

However, for instruments with dichotomous responses (right and wrong, yes or no, etc.), as in the case of traffic applied Physics items, it was indicated using the Kuder- Richardson test (PASQUALI, 2017). This test has the same function as the Cronbach test,[8] but it is specific to analyze this type of instrument e internal consistency.

DIFFICULTY AND DISCRIMINATION INDEXES

For the traffic applied Physics issues, the items difficulty and discrimination indexes must be analyzed. The difficulty index calculates the proportion of correct answers, that is, the ratio between the number of respondents who answered the item correctly and the total number of respondents. Such coefficient has a value between 0 and 1, where 0 would mean that nobody hit the item and 1, they all hit the same. It is expected that most of the items will be of medium level which, according to Vilarinho (2015), would be characterized by difficulty indexes between 0.3 and 0.7.

The discrimination index is intended to measure the item capacity to differentiate respondents with greater ability from those with lower ability. According to Vilarinho (2015), this calculation considers the number of correct answers of 27% of the participants with the best individual performance and compares with the score of 27% who obtained the lowest scores for individual success. This measure, according to Vilarinho (2015), "*corresponds to the difference between the hit percentage of the first group and the second group. The greater the difference, the greater is the item breakdown. It is expected that in an educational evaluation, the discrimination power of the item is greater than 0.40*"[9] (p.27).

The calculating formula for this index consists of:

$$Dc = \frac{(Ns - Ni)}{N},$$

where " N " is the amount corresponding to 27% of the sample ; "Ns" corresponds to the number of hits in the upper group ; and "Ni", the number of hits from the lower group.

If the two groups present the same number of hits for an item, the index will be null, not discriminating between them. The opposite occurs if the index is 1, meaning that the group of respondents with better performance hit the issues that the lowest performing group missed. On the other hand, negative values means that the group with the lowest performance hit such item such that the opposite group missed (MOREIRA; ROSA, 2007).

Rabelo (2013) considered that a discrimination index less than 0.20 evaluates the item as deficient, which should be rejected; between 0.20 and 0.30, considers a marginal item, subject to reprocessing; between 0.30 and 0.40, a good item, but subject to improvement; and values greater than 0.40 evaluates the item as good. These measures will aid in the analysis of which items of traffic applied Physics must be removed or reformulated to compose the final questionnaire, before calculating its internal consistency.

EXPLORATORY FACTORIAL ANALYSIS

In questionnaires answered within Likert scales, we use the exploratory factorial analysis in order to observe in more detail the latent factors that can encompass the variables measured in the test.

For Scoares et. al. (2009), the purpose of this type of analysis is to describe a number of initial variables from a smaller number of hypothetical variables. That is, it consists of a multivariate analysis in which all the variables for the theoretical interpretation of the data set are used at the same time. This allows us to check how many factors the object actually measures. For authors, "*using the factorial analysis also allows us to know the degree of association between each factor for each variable, as well as how the set of factors explains the general variability of the original data*" [10] (p.911).

According to Pasquali (2017), the factorial analysis determines, for each question, a factorial load which indicates the covariance between this and the factor, and the closer to 100% is the covariance, the greater the validity of the item. Still according to the author, the factorial analysis:

> It consists of verifying that a series of items can ideally be reduced to a single dimension or variable, which it calls a factor, with which all variables in the series are related. This being the case, we conclude that the items are one-dimensional, that is, they are measuring the same thing, which is what the one-dimensionality principle seeks. The relation that each item has to the factor is expressed through covariance or correlation; this relationship is called a factorial load. Series items that have high factor load are one-dimensional items because they measure the same factor, while items with a load close to 0 are odd items, so they should be discarded because they are not measuring the same thing as the others; these items sin against unidimensionality and, therefore, cannot be analyzed together with others (PASQUALI, 2017, p.116).[11]

Before carrying out the factorial analysis procedure, it is necessary to verify if collected data are adequate to be analyzed through this method. In order to perform this, we measure two other coefficients: the Kaiser-Meyer-Olkin (KMO) and the Bartlett sphericity test.

According to Scoares et al. (2009), the KMO coefficient confirms the adequacy of the sample to perform a factorial analysis. The test indicates the proportion of variance that is common in the analyzed issues. This coefficient also presents values between 0 and 1, and when close to 1 it indicates a satisfactory adequacy to the factorial analysis model. However, if the coefficient presents values lower than 0.6, the application of this method is not indicated for the presented data set.

Finally, the second test of data suitability is Bartlett's sphericity. It measures the level of significance of the analysis, that is, it checks whether all correlations within the correlation matrix are significant. Its significance measure should be below 0.05.

PILOT TEST RESULTS

The pilot test was applied to 214 High School graduates during the first half of 2018. The spaces where we collect this audience were a Drivers' Training Center, a Federal University and a Federal

Institute, both located in Rio Grande do Sul cities. It is important to point out that the students of the regular teaching institutions had not yet received formal instruction within the courses, since the data collection was performed at the beginning of the semester, as their initial activity. Therefore, it is believed that the knowledge used to answer the questions was acquired in Basic Education or in other spaces, whether formal or non-formal.

The application consisted of explaining research objectives, including the non-mandatory participation, that is, at any time during the test, they could give up or request that their data not be used. This anonymous test collected some information in order to produce a description of the responding individuals. A maximum time for completion of the test was not stipulated, but it was observed that the average filling time was 15 minutes.

Of the 214 interviewees, 104 were female (48.6%) and 110 male (51.4%). Besides, 134 of them (62.6%) did not have the National Driver's License - CNH. Almost half (43.4%) had ages between 18 and 20 years, a fact fully understandable due to the environments in which the interviews were applied, since at this age, the admission to Higher or Technical Education coincides with the search for the first Driver's License coincides.

It was noted that although the majority of the sample was composed by young people, there were also atypical numbers of respondents over 40 years old. It is believed that this fact can help for the heterogeneity of the answers of the questionnaire, contributing to its reliability.

For the type of school in which graduates completed High School, we had 162 (75.7%) coming from state public schools, 27 (12.6%) from federal public schools, 21 (9.8%) from private schools. The remaining 4 respondents (1.9%) finished High School through ENEM test.

As the information on their ages might suggest, most of the participants completed High School in the last 4 years prior to the application of the instrument. Nevertheless, a total of 7 participants were observed, who had completed their studies even before the publication of the Education Guidelines and Bases Law - LDB (BRASIL, 1996). Of the total number of interviewees, only 32 of them (15%) were University graduates and the rest of the group - 182 (85%) - were only High School graduates.

Overall, such data were useful for producing a brief characterization of the participants in the validation process, and showed a heterogeneous but young group who mostly completed their basic studies recently. This allows us to infer that school contents are still recent in their cognitive structures, and the answers to the questionnaires allow us to verify how much they are retained by them.

ANALYSIS OF THE TRAFFIC APPLIED PHYSICS ITEMS

Following, we present the results derived from the validation analysis of the traffic applied Physics questionnaire. First, descriptive data will be exposed, such as the frequency of hits per question and the number of hits. Then, as mentioned previously, we present the Difficulty Index for each item, followed by the Discrimination Index, and finally the results of the Kuder-Richardson test, which calculates the internal consistency of the questionnaire.

The questionnaire on traffic applied Physics consists of 38 items, with answers in a dichotomous format, in which the respondents analyzed the sentences and judged them, agreeing or disagreeing with them. Next, Table 1 presents the number of correct answers for all questions.

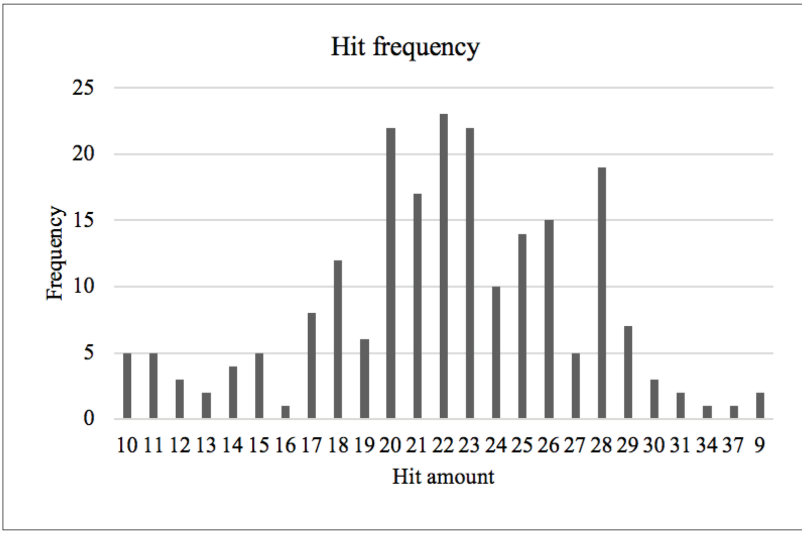
TABLE 1. Hits by item Quantity

Question Number	Number of hits	Question Number	Number of hits	Question Number	Number of hits	Question Number	Number of hits
1	156	11	128	21	81	31	74
2	189	12	124	22	171	32	101
3	79	13	87	23	63	33	100
4	157	14	55	24	114	34	136
5	123	15	154	25	97	35	131
6	160	16	136	26	143	36	129
7	74	17	161	27	109	37	179
8	126	18	86	28	137	38	181
9	88	19	103	29	148		
10	190	20	81	30	129		

Source: the authors.

This sum was the basis for calculating the difficulty index acquired by each item. Items in which most of the participants agreed are considered to be easier, contrary to the ones that the majority missed. In order to complement the presentation of the data, the Graph 1below shows the frequency of hits per respondent.

Graph 1. Individual hit frequency



Source: the authors

In this graph, it can be seen that the most common number of hits was 22, obtained by 23 people (10.7%), followed by 23 and 20, each obtained by a 22-person plot (10.3% of the total group). None of the 214 participants missed all questions, since the minimum number of hits was 9 items, a score acquired by only 5 respondents, making up 2.3% of the total. In turn, only one person scored 37 out of 38 questions.

ITEM DIFFICULTY INDEX

Statistical validation procedure starts by measuring the degree of difficulty of the items. Subdivided into easy, mean and difficult, it is expected that most have an average difficulty index, since a questionnaire mostly easy or difficult does not help the purposes of the test. The following table (Table 2) presents quantitative figures for this measure.

TABLE 2. Item difficulty index

Question Number	Difficulty index	Item type	Question Number	Difficulty index	Item type	Question Number	Difficulty index	Item type
10	0,89	easy	16	0,64	mean	33	0,47	mean
2	0,88	easy	34	0,64	mean	25	0,45	mean
38	0,85	easy	35	0,61	mean	9	0,41	mean
37	0,84	easy	30	0,60	mean	13	0,41	mean
22	0,80	easy	36	0,60	mean	18	0,40	mean
17	0,75	easy	11	0,60	mean	20	0,38	mean
6	0,75	easy	8	0,59	mean	21	0,38	mean
4	0,73	easy	12	0,58	mean	3	0,37	mean
1	0,73	easy	5	0,57	mean	7	0,35	mean
15	0,72	easy	24	0,53	mean	31	0,35	mean
29	0,69	mean	27	0,51	mean	23	0,29	hard
26	0,67	mean	19	0,48	mean	14	0,26	hard
28	0,64	mean	32	0,47	mean			

Source: the authors.

The items are considered of mean difficulty when the index has values between 0.3 and 0.7, with those less than 0.3 being considered difficult and those greater than 0.7, considered easy. As shown in the table, 10 items were considered easy; 2 hard; and 26 of mean difficulty. According to Pasquali (2017), the fact that the questionnaire also has easy or difficult items does not mean that this is a criterion for excluding them, since most are items of mean difficulty. In this way, considering only this criterion, none of the items was considered improper for the validation of the instrument.

ITEM DISCRIMINATION INDEX

This index is key in order to attest that the item differentiates individuals who have more hits from those who hit less. For this, it were observed that 58 individuals (27%) who had better performance in the test, compared with the other 58 (27%) who, in turn, had the lowest performance among all. Returning, if the index has a value close to 0, it means that both groups hit the item equally. On the contrary, if the index is close to 1, it is understood that the best performing individuals answered the questions that the poorer group did not succeed. Items that present negative values should be excluded because it means that the lower performing group scored more than the opposite group.

TABLE 3. Item discrimination index

Question	Index	Question	Index	Question	Index
8	0,569	30	0,431	9	0,241
35	0,517	11	0,414	28	0,241
22	0,500	24	0,397	20	0,190
34	0,500	5	0,379	12	0,172
37	0,500	36	0,379	33	0,172
6	0,483	10	0,345	27	0,155
16	0,483	2	0,328	3	0,138
38	0,483	7	0,310	25	0,138
4	0,466	13	0,310	18	0,121
15	0,466	1	0,293	31	0,103
26	0,466	29	0,293	14	−0,069
17	0,431	32	0,293	23	−0,103
19	0,431	21	0,276		

Source: the authors.

As shown in Table 3, 10 of 38 items were considered non-discriminatory, since the value of the index was less than 0.2, then their removal was suggested. Thus, we emphasize the importance of using more than one benchmark index in the quality analysis of a quantitative research instrument in education, since in a complementary way, what an index may not indicate, as observed in the previous section, another measurement may point out improper items. Thus, items 3, 12, 14, 18, 20, 23, 25, 27, 31 and 33 will be removed from the instrument that will compose the final questionnaire.

KUDER-RICHARDSON TEST FOR INTERNAL CONSISTENCY

In order to complete the validation tests of the traffic applied Physics questionnaire, it is fundamental to measure its internal consistency. This procedure was performed using statistical software Statistical Package for the Social Sciences - SPSS, version 23 for Windows. The Kuder-Richadson test has the same methodological procedure as Cronbach's alpha coefficient, but the latter is not recommended when the answers are of dichotomous nature. As in this questionnaire the answers were tabulated as correct or wrong, the fact of choosing one test over the other is justified.

Firstly, it is important to note that in order to ensure the accuracy of the calculated data, we present the tables generated by the software itself, which occasionally justifies its own formatting and its language in English.

The test was performed initially with 38 items, obtaining a coefficient value of 0.701. The highest internal consistency is associated with the index close to 1, and it is expected to obtain values beyond the minimum recommended. The minimum limit for this value differs from author to author, but there is consensus that values less than 0.6 or 0.7 are not satisfactory, case where is necessary to examine which items could be removed in order to increase its value. The value of 0.701 is already a satisfactory coefficient, however, in the previous section it was observed that 10 items were collaborating for non-discrimination of the test as a whole. Thus, by analyzing the correlation between the items and the test as a whole we seek to observe if, by means of this methodology, those items previously suggested as inappropriate also receive removal recommendation.

For this, we analyzed the "total item correlation", index generated by the software when performing the Kuder Richardson test, as can be seen in Table 4 below.

TABLE 4. Total item correlation for the 38 items.

Question	Total item correlation	Question	Total item correlation	Question	Total item correlation	Question	Total item correlation
37	0,554	17	0,338	24	0,217	12	0,032
38	0,552	26	0,331	29	0,199	27	0,027
22	0,484	8	0,327	7	0,172	3	0,007
10	0,468	15	0,300	32	0,171	20	−0,003
2	0,445	30	0,285	13	0,135	31	−0,019
6	0,427	1	0,277	21	0,122	18	−0,036
4	0,427	11	0,262	28	0,114	14	−0,150
34	0,362	36	0,257	9	0,106	23	−0,227
16	0,352	5	0,225	33	0,044		
35	0,346	19	0,218	25	0,035		

Source: the authors.

It is interesting to note that this test also highlighted the same 10 items as low correlation for the questionnaire as a whole (smaller values than 0.1). That is, they do not relate significantly to the measuring instrument, some even relate negatively. These factors corroborate the removal of these 10 items. Although qualitatively, in order to reduce the final questionnaire to 25 items, we choose remove another 3 (items 9, 2 1 and 28) which also showed a non-satisfactory correlation for the instrument (correlations lower than 0.13).

At the end of this removal procedure, the coefficient of internal consistency was calculated again, in which it is observed that the value increased from 0.701 to 0.803, demonstrating that the decision to remove such items was correct, since it raised the index , initially classified as acceptable, to very good, according to Hair et al. (2006). The Table 5 below shows a comparison of the results before and after the reduction treatment carried out.

TABLE 5. Statistics of reliability

38-item questionnaire			25-item questionnaire		
Cronbach's alpha	Cronbach's alpha for standard items	Number of items	Cronbach's alpha	Cronbach's alpha for standard items	Number of items
0,701	0,726	38	0,803	0,822	25

Source: the authors.

Thus, the analysis procedure confirms that the 25 items of traffic applied Physics constitute a statistically consistent instrument, when its components have discrimination power among respondents with higher and lower performance, being suitable for use in data collection that aims to differentiate opposing performance groups.

ANALYSIS OF QUESTIONS ABOUT THE PHYSICS TEACHING PROFILE

These questions aim to characterize the Physics Teaching profile of schools in which the graduates studied, and they were analyzed by means of exploratory factorial analysis. It has the objective of validating the construct as a whole, besides seeking to reduce all the instrument variables to factors that group them into a set of common correlations.

In order to be successful in this methodology analysis, the data must obey some requirements. The first one is that the number of samples would be greater than 50 (HAIR et al., 2006), being ideal that this number extrapolate 100 cases to assure the significance of the results. For this analysis, the number exceeded the minimum oriented, presenting 214 samples. Another important factor in determining the required minimum number of respondents is the relationship with the number of questions in the instrument. Five respondents for each question on the instrument are recommended. Thus, since the questionnaire traffic applied Physics had 38 questions it would require a minimum of 190 people to analyze it. For the school profile questionnaire, with 26 questions, the number of 214 respondents exceeds the minimum required.

Initially the Cronbach's alpha was calculated to verify the initial internal consistency. The value obtained for the 26-question instrument was 0.770, showing an acceptable value, but with the possibility of improvement. For this, the correlation of the individual items with the bulk was analyzed, in order to verify which questions were more correlated with the others and which were not collaborating for the reliability of the instrument. In the Table 7 below, it is noted that 6 items had a small correlation value (<0.10), since they could disrupt the factorial analysis, it is recommended to their removal.

TABLE 7. Total item correlation

Question	Total item correlation	Question	Total item correlation	Question	Total item correlation
4	0,57	10	0,45	1	0,28
16	0,56	12	0,45	20	0,26
24	0,55	13	0,37	6	0,02
9	0,54	18	0,36	19	0,01
3	0,51	2	0,33	15	−0,02
25	0,51	26	0,32	7	−0,18
14	0,51	5	0,30	11	−0,24
21	0,49	23	0,29	17	−0,35
8	0,47	22	0,29		

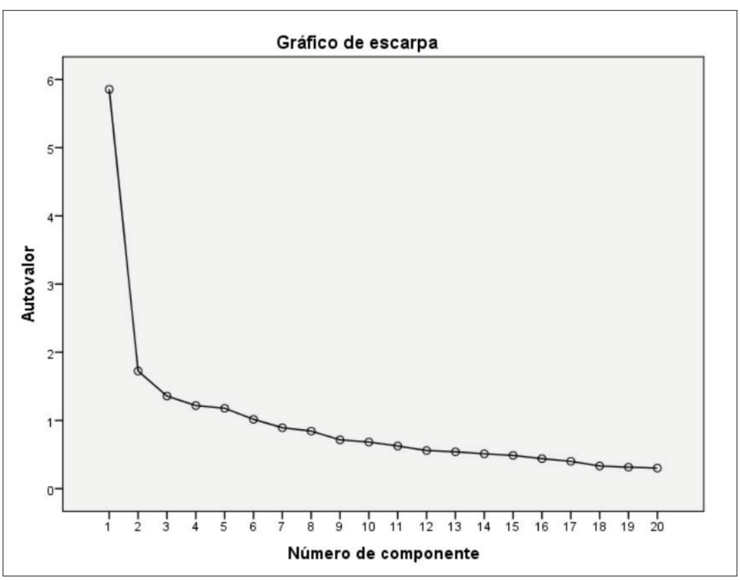
Source: the authors.

After these items removal, we proceeded with the factorial analysis, first for the correlations between all the items. Most of them exceeded the value 0.30, minimum score expected for software estimates factors that relate the variables of the instrument.

The next step was to carry out the adequacy tests of the sample to factorial analysis. For this, the Kaiser-Meyer-Olkin (KMO) tests and the Bartlett Sphericity Test (BTS) were performed. The sample KMO value was 0.854, being this value higher than the expected minimum of 0.60 to consider the sample adequate for this methodology (HAIR et al., 2006). Likewise, the BTS test was statistically significant because its p-value was <0.0000. Considering the adequacy, we performed the analysis, in order to decide the number of factors that would be extracted.

The method used to extract the components was Principal Component Analysis, which generated 6 factors that together explain 62% of the total variance of the instrument. According to the Kaiser criterion, it is suggested to consider the components that have eigenvalues greater than 1, however, a second analysis can be performed through the Scale Chart, also known as the "ScreePlot" Chart, in order to confirm the choice. This graph is a diagram showing the eigenvalues and number of factors in order of extraction. According to Hair et al. (2006), this graph analysis criterion should consider retaining the components that are above the inflection point.

GRAPH 2. Scale Chart for components eigenvalues



Source: the authors.

Based on the observation of this graph, it can be noted that after the second component, the others have similar inclinations, and it is possible to interpret that the first two could cover the grouping of all items of the questionnaire (FIGUEIREDO FILHO; SILVA JUNIOR, 2010).

Next, we present the commonality analysis, which shows the proportion of the variance that is explained for each item from the extracted components. Commonly, it is accepted for each item, communalities with values equal or superior to 0.5. In order to illustrate the relevance of commonalities, observing question 1, which has a value 0.666, what means that the extracted components explain 66.66% of their variance. Given this analysis, it was noted that no variables need to be removed, since none of them has a value below 0.50, as can be seen in the following table:

TABLE 8. Commonalities

Question	Initial	Extraction	Question	Initial	Extraction
1	1,000	0,666	14	1,000	0,553
2	1,000	0,606	16	1,000	0,688
3	1,000	0,599	18	1,000	0,748
4	1,000	0,678	20	1,000	0,557
5	1,000	0,652	21	1,000	0,685
8	1,000	0,525	22	1,000	0,570
9	1,000	0,624	23	1,000	0,590
10	1,000	0,708	24	1,000	0,725
12	1,000	0,584	25	1,000	0,671
13	1,000	0,629	26	1,000	0,695
Extraction method: Main Components analysis					

Source: the authors.

After this procedure, the factorial loads of each variable were analyzed in relation to the extracted components. This analysis demonstrated that components 1 and 2 comprised most of the items of the instrument, making the other 4 components remain grouping in few questions. Thus, a qualitative analysis of the same was carried out in order to reallocate them in the first two components. This decision, together with the exclusion of the 6 items, was accurate, since the internal consistency increased from 0.770 to 0.864, composing a reduced instrument to 20 items in total. After completing this stage, the following consisted of a substantial analysis of the components and their items, with the aim of defining similar aspects between them.

The first component was called "Characteristics that can favor Significant Learning in Physics classes". Here, we find the items highlighting didactic, methodological and evaluative actions that, from Ausubel (2003) perspective, could help in the learning process. This component grouped a total of 14 items.

The second component encompassed the items that indicated characteristics that could disadvantage the Significant Learning in Physics classrooms. Here, again, methodological, evaluative and affective (in respect of teachers) aspects composed this group. Most of the items eliminated by the factor analysis process were from this grouping, which is why this component was finished with only 6 items in total. Nevertheless, the internal consistency value was considered satisfactory, as shown in the sequence. In order to analyze the items in the Table 9 below, the questionnaire is attached at the end of the manuscript.

TABLE 9. Component items and internal consistency for each grouping.

Component 1: Characters that favor Significative Learning in Physics classes				**Component 2**: Characters that disfavor Significative Learning in Physics classes			
Item number	Factorial load	Item number	Factorial load	Item number	Factorial load	Item number	Factorial load
1	0,801	13	0,456	2	0,561	21	0,691
4	0,543	14	0,565	3	0,667	23	0,702
5	0,738	16	0,755	20	0,585	24	0,559
8	0,426	18	0,810				
9	0,710	22	0,420				
10	0,742	25	0,643				
12	0,574	26	0,742				
Cronbach's alpha	**0,830**			**Cronbach's alpha**	**0,722**		

Source: the authors.

Thus, through the factorial analysis, one can finalize the instrument on the Physics Teaching profile with 20 items, with reliability to differentiate opposing characteristics grouped in both groups. The internal consistency of the final instrument attests to a level within that expected to be valid and reliable.

DISCUSSION OF RESULTS

For the questionnaire on traffic applied Physics, the minimum number of hits that a respondent should present to consider a good performance is 60%. This minimum value comes from the Laugksch and Spargo (1996) adopted criteria, who also considered this percentage in TBSL. That is, for the traffic applied Physics questionnaire, which had 38 questions, the minimum number of correct answers should be at least 23 questions.

As discussed in the section of items analysis in this questionnaire, none of the interviewees missed all questions, since the minimum value computed was 9 hits. On the other hand, only 47.2% of the total hit amounts above the stipulated minimum value.

Although this amount represents almost half of the 214 interviewed, it cannot be considered satisfactory, since the other half was not able to obtain the minimum number of correct answers expected. Once they were all High School graduates, one would expect higher levels of performance. These results are in line with the tests already carried out to measure scientific literacy in the international and national literature (LAUGKSCH, SPARGO, 1996; NASCIMENTO-SCHULZE, 2006; VIDOR et al., 2009, CAMARGO et al., 2011 and RIVAS, 2015).

Regarding the issues that were excluded from the instrument, they presented frequencies of hits that characterized them as non-discriminatory, that is, both groups, of higher and lower performance in the test, hit them practically in a similar way. Of the 13 traffic applied Physics questions that were removed, 11 were considered to be mean difficulty and only two (questions 14 and 23) were considered to be difficult. Thus, these items removal will aid in the ability of the questionnaire to differentiate who presents a satisfactory performance or not.

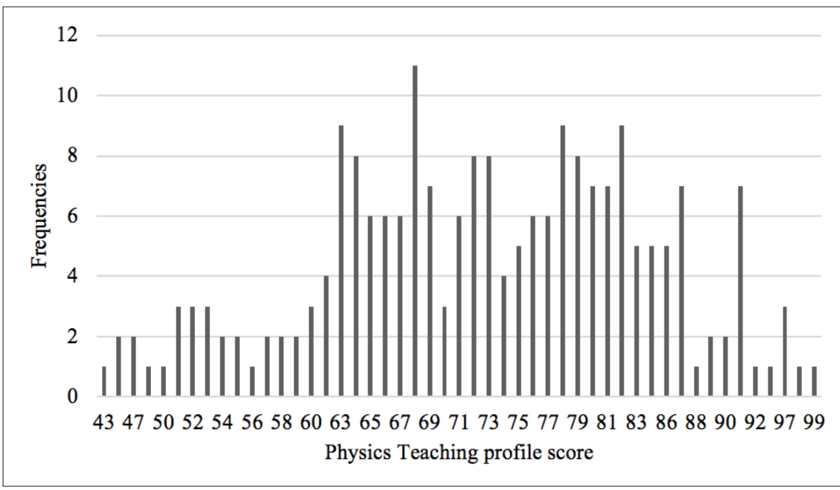
The recommendation that in case of doubt in any question, instead of choosing a random alternative, a question mark should be written on it, it seems appropriate, since not infrequently some alternative was located with this character indicated. This may mean that this method can collaborate so that it does not increase the variance not explained by the internal consistency of the questionnaire.

For Component 1 answers, a score of 5 was given to those who answered “I fully agree”, lowering the score until 1, for “I strongly disagree" answers. For Component 2, the procedure was the reverse, since its alternatives had opposite aspects to the previous ones. In turn, was assigned a value of 5 for those who indicated "I totally disagree", decreasing up to 1 for those who answered "I totally agree".

Therefore, after adding the scores attributed to each item, higher score participants represent those who, according to their perception, attended Physics classes with characteristics that could favor Significant Learning. On the other hand, low score individuals represent students who have attended traditional classes, remaining passive and who currently cannot see the relation between Physics and their daily lives.

Considering 26 items, each with maximum score 5, we have the maximum total score that a participant could present is 130 points and the minimum, 26. The following Graph 4 helps to visualize this pattern.

Graph 4: Frequency of the Physics Teaching Profile



Source: the authors.

The mean score was 73 points, with a standard deviation of 12. The maximum score obtained was 99, reached for only one participant. The minimum score in turn was 43 points. These quantitative figures show that 37.9% of the interviewees obtained a Physics Teaching profile score greater than 60% of the total score (78 points). By this way, it can be characterized that the majority of the interviewees attended classes with characteristics that could disadvantage Significant Learning. Of the 214 participants, 23.8% did not even score half of maximum score (130 points). This reality corroborates Darroz et al. (2015), which showed that even if teachers are aware of methodological actions that may promote meaningful learning, most of them had their pedagogical practice related to traditional conceptions of teaching, considering students' prior knowledge as well as the evaluative process, barriers that make it difficult to work from another perspective.

These data can be considered as a demonstration of the school reality reported by High School graduates, and these conclusions may be related to the performance statistics previously discussed. In order to verify such a hypothesis with a higher confidence interval, a greater number of participants is necessary in order to increase the heterogeneity of the sample.

In observing such characterization differentiations, together with the statistical results already presented, it can be stated that this instrument also performed satisfactorily the task of investigating the Physics classes' profile in which the students attended during most of their High School years.

CONSIDERATIONS

This manuscript aimed to present the validation procedure of a Physics questionnaire for High School graduates. For the purpose of conclusion, it was considered the instrument with satisfactory validity, in which the methodology suggested the removal of some of the items of the original instrument, composing a questionnaire of traffic applied Physics with 25 questions and a questionnaire of the Physics Teaching profile with 20 questions.

Such instrument consists of a suggestion of how ways can be created to question individuals by conditioning them to use not only their memorized knowledge to relate Physical phenomena to their concept or nomenclature, but rather to stimulate them to emit a judgment, based on all the knowledge gained during their lives, analyzing the scientific coherence of present situations within their daily lives.

It is believed that this dynamics of knowledge measurement should be stimulated, not only in traffic situations, but in other contexts, as long as they are significant within the group's realm. Likewise, it is believed that other areas such as Chemistry, Biology and Mathematics could also benefit from this approach because, like Physics, they try to explain knowledge sometimes hermetic to the general public, and understanding it is a need for the development of the criticality, that is expected of a citizen in the full use of his faculties.

The limitations observed in this validation work are the suggestions for future studies: increasing the sample quantity and the heterogeneity of the group, so that the statistical methods can be applied to investigate patterns present in the data. With this work it was possible to note that it is still challenging to measure factors that are not directly measurable, such as Scientific Literacy, for example, but, indirectly, inferences can be made, so that patterns can be observed and deductions inferred.

REFERENCES

ALEXANDRE, N. M. C.; COLUCI, M. Z. O. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. **Ciência & Saúde Coletiva**, v. 16, n 7, p. 3061-3068, 2011.

AMBEV. **Retrato da Segurança viária 2017**. Disponível em:< https://www.ambev.com.br/conteudo/uploads/2017/09/Retrato-da-Seguran%C3%A7a-Vi%C3%A1ria_Ambev_2017.pdf>. Acesso em: jun. 2018.

AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (AAAS):**Science for All Americans**. Washington: AAAS, 1989.

ANGOTTI, J. A. P.; AUTH, M. A. Ciência e tecnologia: implicações sociais e o papel da educação. **Ciência & Educação**, v. 7, n. 1, p. 15-27, 2001.

AULER, D.; DELIZOICOV, D. Alfabetização científico-tecnológica para quê?. **Ensaio Pesquisa em Educação em Ciências**, v. 3, n. 2, p. 122-134, 2001.

AUSUBEL, D. P. **Aquisição e retenção de conhecimentos**: uma perspectiva cognitiva. Lisboa: Plátano, v. 1, 2003.

BRASIL. **Lei de diretrizes e bases da educação nacional**. Brasília: MEC, 1996.

BRASIL. PCN do Ensino Médio: **Parâmetros Curriculares Nacionais**. Ciências da Natureza. Matemática e suas Tecnologias, Ministério da Educação, Secretaria de Educação Básica, 2000.

BRASIL. **Base Nacional Comum Curricular - BNCC**. 2017. Disponível em:< http://basenacionalcomum.mec.gov.br/images/BNCC_20dez_site.pdf>. Acesso em: jun. 2018.

CAMARGO, A. N. B. de; PILAR, F. D.; RIBEIRO, M. E. M.; FANTINEL, M.; RAMOS, M. G. Alfabetização Científica: A evolução ao Longo da formação de Licenciandos Ingressantes, Concluintes e de Professores de Química. **Momento – Diálogos em Educação**, v.20, n.2, p. 19-29, 2011.

CHAGAS, C. **A Física no ensino médio através do estudo de fenômenos físicos em um automóvel.** 2014. Dissertação (Mestrado em Ensino de Ciências e Matemática). Universidade Federal do Ceará, Fortaleza. 2014.

CHASSOT, A. **Alfabetização Científica – Questões e desafios para a educação**. Ijuí, Editora da Unijuí, ed. 1, 2000.

CHASSOT, A. Alfabetização científica: uma possibilidade para a inclusão social. **Revista Brasileira de Educação**, n. 22, p. 89-100, 2003.

DARROZ, L. M.; ROSA, C. W. da; GHIGGI, C. M. Método tradicional x aprendizagem significativa: investigação na ação dos professores de Física. **Aprendizagem Significativa em Revista**, v. 5, n. 1, p. 70-85, 2015.

EBBINGHAUS, H. Memory: A contribution to experimental psychology. **Annals of neurosciences**, v. 20, n. 4, p. 155, 2013.

FIGUEIREDO, Vera L. M. de; MATOS, V. L. D. de; PASQUALI, L.; FREIRE, A. P. Propriedades psicométricas dos itens do teste WISC-III. **Psicologia em estudo.** v.13, n.3, p. 585-592. 2008.

FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A. da. Visão além do alcance: uma introdução à análise fatorial. **Opinião pública**, v. 16, n. 1, p. 160-185, 2010.

FOUREZ, G. **Alphabétisation Scientifique et Technique** – Essai sur les finalités de l'enseignement des sciences, Bruxelas: DeBoeck – Wesmael, 1994.

FREIRE, P. **Educação como prática da liberdade**. 24. ed. Rio de Janeiro: Paz e Terra, 2000.

GIL-PÉREZ, D.; VILCHES-PEÑA, A. Una Alfabetización Científica para el siglo XXI: Obstáculos y propuestas de Actuación. **Investigación em la Escuela**, v. 43, n.1, p. 27-37, 2001.

HAIR JÚNIOR, J.; BLACK, W.; BABIN, B.; ANDERSON, R.; TATHAM, R. **Multivariate Data Analysis**. 6ª ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2006.

HURD, P. D. Science literacy: Its meaning for American schools. **EducationalLeadership**, v. 16, p. 13-16, 1958;

HURD, P. D. Scientific Literacy: New Minds for a Changing World, **Science Education**, v. 82, n. 3, p. 407-416, 1998.

JENKINS, E. Scientific literacy. In: HUSEN, T.; POSTLETHWAITE, T. N. (org.). **The international encyclopaedia of education.** Oxford: Pergamon Press, 1994, v. 9. p. 5345.

LAUGKSCH, R.; SPARGO, P. Construction of a paper-and-pencil Test of Basic Scientific Literacy based on selected literacy goals recommended by the American Association for the Advancement of Science. **Public Understanding of Science**, v. 5, n. 1, p. 331-359, 1996.

LAUGKSCH, R. Scientifc Literacy: A Conceptual Overview, **Science Education**, v. 84, n. 1, p. 71-94, 2000.

LORENZETTI, L. **Alfabetização científica no contexto das séries iniciais**. 2000. Dissertação de mestrado. Programa de Pós-Graduação em Educação. Universidade Federal de Santa Catarina. 2000.

MARTINS, G. de A. Sobre confiabillidade e validade. **Revista Brasileira de Gestão de Negócios**, v. 8, n. 20, 2006.

MILLER, J. Scientific literacy: a conceptual and empirical review. Daedalus, v. 112, n.2, p. 29-48, 1983.

MOREIRA, M. A. A teoria da aprendizagem significativa e sua implementação em sala de aula. Brasília: Editora Universidade de Brasília, 2006.

MOREIRA, M. A.; ROSA, P. R. da S. Uma introdução à pesquisa quantitativa em Ensino. Porto Alegre: Ed. dos Autores, 2007.

NASCIMENTO-SCHULZE, C. M. Um estudo sobre Alfabetização Científica com jovens catarinenses. Psicologia: teoria e prática, v. 8, n. 1, p. 95-117, 2006.

NEGRINI NETO, O.; KLEINUBING, R. Dinâmica dos acidentes de trânsito: análises, reconstrução e prevenção. Campinas: Editora Millennium, 2012.

OLMA, M. Consciência sobre rodas: Habilitação do condutor. 20ª ed. Porto Alegre – RS: Águia, 2016.

PASQUALI, L. Princípios de elaboração de escalas psicológicas. Revista Psiquiátrica Clínica, v. 5, n. 25, p. 206-213, 1998.

PASQUALI, L. Psicometria: teoria dos testes na psicologia e na educação. Petrópolis: Editora Vozes, 2017.

PINTO, A. A retenção à longo prazo do conhecimento escolar. **Revista Portuguesa de Pedagogia**. XXIII, 283-295, 1989.

PINTO, A.; OLIVEIRA, T. O que é que os alunos recordam das aulas que leccionámos? **Revista Portuguesa de Pedagogia**, XXV, 87-102, 1991.

PINTO, A. da C. **Memória, cognição e educação:** Implicações mútuas**.** Educação, cognição e desenvolvimento: Textos de psicologia educacional para a formação de professores, 2001.

RABELO, M. L. **Avaliação Educacional**: Fundamentos, Metodologia e Aplicações no Contexto Brasileiro. Rio de Janeiro. SBM, 2013.

RIVAS, M. I. E. **Avaliação do nível de Alfabetização Científica de estudantes de biologia**. Trabalho de conclusão de curso. Graduação em biologia. Universidade Federal do Rio Grande do Sul – UFRGS, 2015;

SANTOS, W. L. P.; MORTIMER, E. F. Uma análise de pressupostos teóricos da abordagem CTS (Ciência-Tecnologia-Sociedade) no contexto da educação brasileira. **Ensaio Pesquisa em educação em ciências**, v. 2, n. 2, 2000.

SANTOS, W. L. P.; MORTIMER, E. F. **Tomada de Decisão para Ação Social Responsável no Ensino de Ciências**, Ciência & Educação, vol. 7, n.1, p. 95-111, 2001.

SASSERON, L. H.; CARVALHO, A. M. P. **Alfabetização Científica**: uma revisão bibliográfica. Investigações em Ensino de Ciências, v.16, n.1, p. 59-77, 2011.

SCOARIS, R. C. de O.; PEREIRA, A. M. T. B.; SANTIN FILHO, O. Elaboração e validação de um instrumento de avaliação de atitudes frente ao uso de história da ciência no ensino de ciências. **Revista Eletrónica de Enseñanza de las Ciencias**, v. 8, n. 3, p. 901-922, 2009.

SOUZA, A. C. de; ALEXANDRE, N. M. C.; GUIRARDELLO, E. de B. Propriedades psicométricas na avaliação de instrumentos: avaliação da confiabilidade e da validade. **Epidemiologia e Serviços de Saúde**, v. 26, p. 649-659, 2017.

TAVARES, R. Aprendizagem significativa. **Revista conceitos**, v. 55, n. 10, 2004.

THORNDYKE, P. Cognitive structures in comprehension and memory of narrative discourse, **Cognitive psychology**, v. 9, n. 1, p. 77-100, 1977.

TULVING, E. **Elements of episodic memory.** New York, Oxford University Press, 1983.

URRUTH, H. G.; STEFFANI, M.; SILVEIRA, F. L. Física e segurança no trânsito: um curso de Física e educação para o trânsito para jovens e adultos. **Revista Polyphonía**, v. 26, n. 2, p. 313-321, 2015.

VIDOR, C. de B.; COSTA, S. S. C. da; SILVA, A. M. M. da; RAMOS, M. G. Avaliação do nível de Alfabetização Científica de professores da educação básica. In: VII Encontro Nacional de Pesquisadores em Ensino de Ciências, 2009. **Anais** do VI ENPEC. Florianópolis: ABRAPEC, 2009. Disponível em: http://www.foco.fae.ufmg.br/viienpec/index.php/enpec/viienpec/paper/view/1047. Acesso em: fev. 2018.

VILARINHO, A. P. L. **Uma proposta de análise de desempenho dos estudantes e de valorização da primeira fase da OBMEP**. Dissertação (Mestrado Profissional em Matemática). Universidade de Brasília, Brasília, 2015

VIZZOTTO, P. A.; MACKEDANZ, L. F. A compreensão da Física aplicada ao trânsito na perspectiva de egressos do ensino médio, alunos de cursos de primeira habilitação. **Revista Brasileira de Ensino de Física**, São Paulo , v. 39, n. 3, e3404, 2017.

NOTAS

[1] Pode-se perceber que no cerne das discussões de autores que usam um termo ou outro, estão as mesmas preocupações com o ensino de Ciências, ou seja, motivos que guiam o planejamento desse ensino para a construção de benefícios práticos para as pessoas, sociedade e meio ambiente.

[2] é uma meta desejável para todos os cidadãos, mensurável e avaliável, útil para a vida cotidiana e ligada fortemente ao contexto social

[3] A alfabetização é mais do que o simples domínio psicológico e mecânico das técnicas de escrever e de ler. É o domínio dessas técnicas, em termos conscientes. É entender o que se lê e escrever o que se entende. É comunicar-se graficamente. É uma incorporação. Implica, não uma memorização visual e mecânica de sentenças, de palavras, de sílabas, desgarradas de um universo existencial – coisas mortas ou semimortas – mas numa atitude de criação e recriação, implica numa autoformação de que possa resultar uma postura interferente do homem sobre seu contexto.

[4] The TBSL is thus testing only fundamental, or basic, aspects of scientific literacy, as knowledge of interdisciplinar concepts, applications of Science, and the ability to apply knowledge for decision making and problem solving are not being tested here

[5] Mede a proporção ou porcentagem de juízes que estão em concordância sobre determinados aspectos do instrumento e de seus itens. Permite inicialmente analisar cada item individualmente e depois o instrumento como um todo. Este método emprega uma escala tipo Likert com

pontuação de um a quatro, a fim de avaliar a relevância/representatividade de cada item.

[6] um instrumento é válido na extensão em que mede aquilo que se propõe medir. Por exemplo, um instrumento válido para medir a capacidade de leitura deve medir realmente esta característica e não outras, como por exemplo, conhecimento prévio

[7] quanto menor a soma da variância dos itens, mais consistente é considerado o instrumento

[8] In SPSS the Kuder-Richardson test is performed using the same procedure in which the Cronbach's alpha is calculated because the software comprises the internal consistency test for dichotomous items as a particular case of Cronbach's alpha test, which calculates the internal consistency for other data formats.

[9] corresponde à diferença entre o percentual de acerto do primeiro grupo e do segundo grupo. Quanto maior foi essa diferença, maior será a discriminação do item. Espera-se que em uma avaliação educacional, que o poder de discriminação do item seja superior a 0.40.

[10] utilizar a análise fatorial permite também com que se saiba o grau de associação entre cada fator a cada variável, assim como o quanto o conjunto de fatores explica a variabilidade geral dos dados originais.

[11] Consiste em verificar se uma série de itens pode ser reduzida idealmente a uma única dimensão ou variável, que ela chama de fator, com o qual todas as variáveis da série estão relacionadas. Sendo este o caso então se conclui que os itens são unidimensionais, isto é, estão medindo a mesma coisa, que é o que o princípio da unidimensionalidade procura. A relação que cada item tem com o fator é expressa através da covariância ou da correlação; esta relação se chama de carga fatorial. Itens da série que têm alta carga no fator são itens unidimensionais, pois medem o mesmo fator, enquanto itens com carga perto de 0 são itens estranhos e, por isso devem ser descartados, porque não estão medindo a mesma coisa que os demais; estes itens pecam contra a unidimensionalidade e, portanto, não podem ser analisados juntamente com os outros.

Submission: 10/06/2018
Approbation: 14/08/2018

Contact:
Patrick Alves Vizzotto
Postgraduate Program in Science Education
Department of Biochemistry
Rua Ramiro Barcelos, 2600, Bairro Rio Branco,
Porto Alegre|RS|Brasil
CEP 90.035-003

ANNEX: FINAL VALIDATED QUESTIONNAIRE

Age: __________.

Sex: ( ) Male ( ) Female.

Driver's license: ( ) Yes ( ) No

High School conclusion year: __________.

Type of school in which you have completed High School:

( ) Private; ( ) State Public; ( ) Federal Public; ( ) ENEM.

School Level:

( ) High School; ( ) University graduate; ( ) Postgraduate studies.

The questions are in the form of statements. Read each sentence carefully and based on your knowledge, mark each of the following items with the letter "C" when you agree to the statement or with the letter "D" when you disagree with it. If you do not really know the answer, mark it with a question mark "?".

	1	On rainy or low temperatures days, air-contained water vapor condensation causes the windows to become fogged.
	2	When a highway is wet, the ideal is to reduce the speed to avoid a possible loss of grip between the tire and the road.
	3	The use of the seat belt, due to the inertial effects of the bodies, acts to avoid the tendency of the body to remain in movement and thus to collide with the structure of the car in sudden braking and collisions.
	4	Due to the non-rigidity of the material of the seat upholstery, passengers in the back seats of the vehicle are exempted from wearing a seat belt if their speed does not exceed 40 km h.
	5	Due to the inertial effects, the Airbag avoids the collision of the body with the structure of the car in brakes and collisions, thus being a seat belt substitute.

	6	Vehicles that also use spherical mirrors, such as external mirrors, provide the driver a greater field of vision when compared to vehicles that use only flat mirrors.
	7	In a collision modern vehicles are safer than the old ones because, among many factors, the materials that make up the bodywork and bumper are more susceptible to kneading in modern vehicles, which helps to dissipate the collision energy.
	8	If an aquaplaning occurs, the best thing to do is to accelerate the car in order to increase grip with the ground.
	9	The total distance required to stop a vehicle consists on the distance that the car travels during the reaction time of the driver and the distance required for braking, these distances are directly proportional to the car's speed.
	10	The phenomenon of aquaplaning only happens when the tires are bald.
	11	A tire is considered bald when it does not have the capacity to drain the water that lies between the tread and the road surface on rainy days, reducing the coefficient of friction between the two materials.
	12	In an accident, older cars are safer compared to today's cars, because the current ones knead more easily through any application of force, leaving the occupant more vulnerable.
	13	Eventually we get shocked by leaning our toes on vehicles due to static electricity being discharged from our body to Earth.
	14	Using gasoline without any percentage of alcohol would provide greater fuel burning power, generating more power than conventional roads could handle to be safe in traffic.
	15	Cars, buses and trucks have different speed limits on the same highway because they have different amounts of movement for the same speed.

	16	When the outside temperature is greater than the internal temperature of a vehicle, the internal fog of your windows occurs.
	17	A car and a truck with masses 1 and 15 tons, respectively, both at 80 km/h have the same energy of movement.
	18	One way of reducing the blind regions in the exterior mirrors of a vehicle is to adjust them almost perpendicularly with its body.
	19	The traction wheels of a vehicle traveling with the fifth gear engaged receive less amount of engine force (torque) than when it is in smaller gear
	20	A vehicle may invade the opposing lane or leave for the shoulder during a curve due to the inertia, the tendency to follow in a straight line tangent to this curve.
	21	Tankers, carrying fuels, walk with a metal chain dragging on the ground in order that it serves as a ground so that no static electricity is stored in their body.
	22	Stopping a vehicle means converting your kinetic energy into other forms of energy, such as sound and thermal energy.
	23	The amount of movement of a car walking at 80 km/h is less than that of a truck at the same speed.
	24	A vehicle becomes lighter when traveling at high speed.
	25	The sound of sirens and fast cars seems distorted as they approach or drift away from an observer because the high-pitched sound propagates in the air faster than a low sound.

About the attended Physics classes during High School, mark each statement with an X in the option that represents your opinion. Check the "CP" option when fully agree; "C" when Agree; "N" when Do not agree, nor disagree; "D" when Disagree; and "DP" when disagreeing fully.

CP	W	N	D	DP		
					1	After the evaluations there were moments to reflect on the mistakes and to understand what their difficulties were.
					2	He felt motivated to learn physics.
					3	There was a large number of formulas needed to be memorized without analyzing their meaning
					4	Upon completion of High School you could fully understand the physical phenomena present in your daily life.
					5	Most of the time, the teacher started a new subject questioning what the students already knew about the subject.
					6	There were no moments for discussions between teacher and students and between the students themselves.
					7	After class on a new physical concept, you could usually observe such phenomena in everyday situations.
					8	In order to teach new content, the teacher presented the historical context from which that knowledge derived.
					9	Generally the questions of tests and exercises were considered only as right or wrong, ignoring or disregarding the development of the resolution of the same.
					10	In the Physics classes there were experimental activities, in which the students manipulated the experiments.

					11	It was common for the teacher to relate Physics subjects to other school subjects.
					12	The way the Physics subjects were approached made the classroom unmotivating.
					13	In Physics classes there were moments of debate between teacher and students, and between the students themselves.
					14	A part of the evaluation was composed by Self-Assessment.
					15	Normally the physical concepts studied were forgotten shortly after the test.
					16	The teacher encouraged students to seek to broaden their knowledge through research, videos, reports and simulations.
					17	In general, the teacher presented some introductory material (scientific texts, reports, videos, films, simulations) before beginning the study of a new concept in order to familiarize the students with the theme.
					18	The Physics contents had little relation with everyday situations.
					19	During the class the teacher allowed the students to talk about the subject being addressed.
					20	In order to teach a new subject, the teacher usually approached more general aspects of the subject to later detail the concept studied.

Thank you for your contribution.