

Study of academic plagiarism through multidimensional scaling and network analysis

Estudio del plagio académico mediante escalamiento multidimensional y análisis de redes

<https://doi.org/10.4438/1988-592X-RE-2022-397-548>

José Antonio Sarmiento Campos

<https://orcid.org/0000-0003-3537-1197>

Camilo Isaac Ocampo Gómez

<https://orcid.org/0000-0003-1545-7866>

María Dolores Castro Pais

<https://orcid.org/0000-0003-2002-7262>

Universidad de Vigo

Abstract

This article presents a study conducted through the application of analytical techniques of multidimensional scaling (MDS) and social network analysis (SNA) on the results obtained in a research on plagiarism in academic work. In it, the collection of data was done through the attribution questionnaire for the detection of coincidences in academic works called CUDECO, created and validated to evaluate the relevance that students give to the fact of quoting correctly and determine the causes that lead them to plagiarize. The objectives of the work that we present are three: to study the dimensional model of CUDECO; demonstrate the suitability of analysis triangulation to do so; build a graphic representation of the plagiarism model of the University of Vigo students that enables a general and comprehensive perception of it. The conceptual basis of the study is elaborated from the concepts of plagiarism and triangulation, more specifically analytical triangulation, and its possibilities in the Social Sciences. This permit us to make an introduction to the use of multivariate analytical techniques as an adequate means to validate a survey study. Within them, two applicable procedures are exposed that can be completed with the

analysis of social networks in the situation described (context). To respond to the objectives of the study, two multidimensional scaling procedures are applied: PROXimity SCALing (PROXSCAL) and Scaling by MAyorizing a COMplicated Function (SMACOF). Both procedures are completed with the Analysis of Social Networks (ARS). As the main conclusion, it is established that the suitability of the procedures used to carry out a study of the dimensionality underlying the responses to the items that make up a questionnaire has been demonstrated, in this case, the CUDECO and a better understanding of the causes of academic plagiarism.

Keywords: higher education; analytical triangulation; multidimensional scaling; academic plagiarism

Resumen

En este artículo se presenta un estudio realizado mediante la aplicación de técnicas analíticas multivariadas denominadas escalamiento multidimensional (EMD) y análisis de redes sociales (ARS) a los resultados de una investigación sobre comisión de plagio e integridad académica del alumnado de la Universidad de Vigo (UVigo). La información de partida ha sido recogida a través del Cuestionario para la Detección de Coincidencias en Trabajos Académicos (CUDECO), un instrumento creado y validado para evaluar, tanto la relevancia que el alumnado le da al hecho de citar correctamente, como las causas que lo llevan a cometer plagio. Los objetivos del trabajo son tres: estudiar el modelo dimensional del CUDECO, demostrar la idoneidad de la triangulación de análisis para hacerlo y construir una representación gráfica del modelo de plagio del alumnado de la UVigo que posibilite una percepción general y comprensiva del mismo. La base conceptual del estudio está elaborada a partir de los conceptos de plagio y triangulación, más concretamente de triangulación analítica, y a sus posibilidades en las Ciencias Sociales. Ello permite efectuar una introducción al uso de las técnicas analíticas multivariadas como medio adecuado para validar un estudio de encuesta. Dentro de ellas se exponen dos procedimientos aplicables que pueden completarse con el análisis de redes sociales en la situación que se describe (contexto). Para dar respuesta a los objetivos del estudio se aplican dos procedimientos de escalamiento multidimensional: PROXimity SCALing (PROXSCAL) y Scaling by MAyorizing a COMplicated Function (SMACOF). Se completan ambos procedimientos con el de Análisis de Redes Sociales (ARS). Como principal conclusión se establece que ha quedado demostrada la idoneidad de los procedimientos utilizados para realizar un estudio de la dimensionalidad subyacente a las respuestas de los ítems que componen un cuestionario, en este caso, el CUDECO y una mejor comprensión de las causas del plagio académico.

Palabras clave: educación superior; triangulación analítica; escalamiento multidimensional; plagio académico

Introduction

Plagiarism in academic papers

Plagiarism, like most topics of a pedagogical nature, constitutes a fact of great complexity and, consequently, its study requires the consideration and evaluation of multiple variables (Boillos, 2020; Cebrián-Robles et al., 2018). This is a recently topical phenomenon (Cebrián-Robles, Raposo-Rivas, & Sarmiento-Campos, 2016) in whose explanation it is necessary to take into account intrinsic/extrinsic factors that are in continuous interaction (Amiri, & Razmjoo, 2016; Sureda-Negre, Comas-Forgas, & Oliver-Trobat, 2015).

From a fundamentally hermeneutic approach to the phenomenon of academic honesty in the university, attention should be paid to those papers that, written mostly in the English language, focus on academic literacy and its relationship with plagiarism. In this regard, those of Abasi, & Graves (2008) can be cited, where university policies against plagiarism are criticized, since these can create a formal barrier, sometimes insurmountable, especially in the case of postgraduate students from other countries or centers that do not know the institutional culture of ensuring academic integrity, which generates negative discourses around plagiarism (Moxley, & Archer, 2019). Also of interest is the work of Badenhorst (2018) in which the difficulties presented to postgraduate students when facing academic papers such as the writing of a review are exposed, due to the lack of “a pedagogy that relates citation with the most complex and fluid conceptual and ontological practices that are implicit in academic contexts” (p.121). To an analogous conclusion, although referring to complex forms of plagiarism, come Childers, & Bruton (2016).

Two lines of interest can be identified in the Spanish language. On the one hand, those who emphasize the importance of the willfulness of plagiarism in the papers of university students as a characteristic feature; and, on the other, those who find in plagiarism a conscious behavior admitted by a large number of students. Within the first

group, Boillos (2020) proposes a taxonomy of unconscious plagiarism; Saneleuterio (2017) gives an account of those academic literacy activities in undergraduate students useful to avoid involuntary plagiarism; Vargas-Franco (2019), through a case study, affects the promotion of didactic aspects and gives less weight to punishers.

In the second group, Comas et al. (2011); Duche et al. (2020) found that a high percentage of university students who recognize dishonest academic behavior, although they disapprove it, consider that its use is widespread, known and to some extent accepted. This is a fact that is reinforced at present with the considerable increase in the portals of purchase of academic papers as necessary collaborators in the absolutely dishonest conduct of the current university students (Comas-Forgas, Morey-Lopez, & Sureda-Negre (2021). For our part, we have a certain knowledge of academic plagiarism through the research we have been carrying out over the last three years on its causes (Espiñeira-Bellón et al., 2021; Muñoz-Cantero et al., 2019; Muñoz-Cantero et al., 2021), but we are aware of the limitations of its results. We are interested in further deepening that knowledge. But ... how can we do it from the results we have been obtaining?

To answer this question, we have carried out the study that is reported in this article. In it, we try to take advantage of the possibilities offered by computer technology and the development of statistics to test a more complex model that, through multivariate analytical techniques, allows us to evaluate possible causal relationships underlying the results obtained so far. The results to which we refer come from the statistical analysis of the data corresponding to the application of a questionnaire to a sample of students at the university (see for this purpose the section Methodology).

Triangulation as a starting concept

When we consult the entry of the term “triangulation” in the Diccionario de la Real Academia Española (RAE), in its 23rd edition, we find a recursive definition of double nature. It is an operation and, at the same time, it is the set of data obtained by that operation or action. It was born in the disciplines of Architecture and Geology, related to the fact of measuring and whose constituent element is the triangle as a paradigm of simplicity,

completeness and security in the results. Also, the term triangulation is used in navigation to locate a position with reference to several points. Its essence is present in the correctness of the measures and in the relationship between those that are known to us. The translation of this concept from the fields of Architecture and Geology, as well as that of the Navy, to scientific research takes place thanks to a process of analogy that finds practical, utilitarian and even ontological similarities.

Triangulation in educational research can be said to generally consist of “the application and combination of several research methodologies in the study of the same phenomenon” (Denzin, 1989, p. 297). This aims to improve both the validity and reliability of the data obtained in the process and of the process itself, and it is here that the dual nature and recursiveness of triangulation are really highlighted. The idea behind triangulation in scientific research is that the convergence of multiple constituent elements of the research process makes it possible to support the conclusion on a stronger basis than can be provided by support in only one of them. If a hypothesis is contrasted from different approaches and is not refuted, it has a greater degree of validity than if it only resists the contrast of one of the approaches.

There are several works that focus on the vicissitudes of triangulation that can be consulted (Arias, 2000; Denzin, 1970; Denzin, 1989; Rodríguez, Pozo, & Gutiérrez, 2006; Thurmond, 2001). All of them contemplate the basic types of triangulation proposed by Denzin (1970) and that would derive from the combination of two or more data sources, researchers, methods and theories. Years later, Kimchi, Polivka, & Stevenson (1991) also consider triangulation in the analysis instrument sector, which would be called analytical triangulation or triangulation in the analysis that, for Rodríguez, Pozo, & Gutiérrez (2006) is a technique of confrontation and a tool for comparing different types of data analysis that, with the same objective, can contribute to validate a survey study and enhance the conclusions derived from it. In a subsequent methodological review of triangulation as a strategic research technique Alzás et al. (2016) consider that when the different methods or instruments of analysis pursue the same objective, as it is the case of this research, we would be facing a genuine process of triangulation, beyond the mere combination or complementation of techniques.

In this sense a very interesting work is the one published by Bright, Heesen, & Zucker (2016) in which they develop a logical-formal model

that concludes with the fact that triangulation provides confirmatory support for its use in improving the reliability of research results and, in particular, it does so even if the researcher is not sure which of the available methods or techniques can be trusted.

Multidimensional scaling

It is necessary to add to the possibilities that, as we have seen, the triangulation of analysis brings to educational research, the capacity, power and versatility of computers and programs that allow to successfully face the simultaneous processing of large amounts of data. In fact, several multivariate analytical techniques emerged from theoretical studies (Bennett, & Hays, 1960; Kruskal, 1964a; Kruskal, 1964b; Shepard, 1962; Torgerson, 1958) have had to wait in time until computer technology facilitated their recovery and promoted their use among the research community. MultiDimensional Scaling (MDS) is one of these multivariate analytical techniques. It can be conceptualized as a procedure that allows the person investigating to determine the representation they perceive with respect to a set of objects or variables (companies, products, ideas or other elements associated with common perceptions), as well as to submit it to judgment (Aldas, & Uriel, 2017; Pituch, & Stevens, 2016). Here the word object has a polysemic nature and refers to any entity that we wish to submit to analysis (Arce, De Francisco, & Arce, 2010).

The purpose of the MDS technique is to transform similarity judgments and general preferences by trade marks or opinions into distances represented in a multidimensional space (Hair et al. 2014; Johnson, & Wichern, 2019). The degree of adjustment between the original data and calculated distances is translated into a numerical indicator called "stress." The closer you get to zero, the better the adjustment. Specific papers such as Real (2001), Guerrero, & Ramírez (2002) or Hahs-Vaughn (2017) can be consulted in order to explore this type of technique.

Social Media analysis

In this study, the Social Media Analysis (SMA) complements the multivariate analytic technique MDS. The SMA is another technique used

to handle and make comprehensible large volumes of information and visualize very complex structures in a simple way. Structures (nodes) and relationships (arcs) can be studied with it. It is based on the theory of graphs (Gaete, & Vasquéz, 2008; Paniagua, 2013) and as a methodological option, is focused on the study of structures composed of elements called actors or nodes and the relationships that occur between them (Sarmiento, Ocampo, & Cid, 2020).

In the analysis of social media, the structure of the relationships in which each node is involved is considered, these nodes are described through their connections (Pavlopoulos et al., 2017).

The application of network theory in the social sciences has had an important boom in recent decades due to its flexibility to analyze relationships and underlying interdependencies in data sets (Sandoval, Morales, & Díaz, 2019).

A network is a set of points that are joined by links (lines or edges) from an association rule that indicates how nodes are related (Mitchell, 2009).

Objectives

In view of the possibilities offered by the technique of analytical triangulation through multidimensional scaling (PROXCAL and SMACOF) and social media analysis (SMA), the following objectives are set out as objectives of the study:

1. To make explicit, to make visible and to study the dimensional model of the Questionnaire for the Detection of Matches in Academic Jobs (CUDECO), based on the analysis of the responses of the students to the 47 items on it.
2. To demonstrate the suitability of the triangulation of analysis in the study of the dimensionality underlying the responses to the items that make up the CUDECO questionnaire, while showing its possible usefulness for other similar papers.
3. To construct a graphical representation of the student plagiarism model of the University of Vigo based on their responses to the items of CUDECO that allows a general and comprehensive perception of the model that the student possesses about academic plagiarism.

Methodology

Context of the study

The present study is part of the project “Study on plagiarism in students of the Galician University System”, developed in academic courses 2018/2019 and 2019/2020. In this study, the Questionnaire for the Detection of Matches in Academic Jobs (CUDECO) (Muñoz-Cantero et al., 2019) was applied to more than 10.000 students of the Galician University System. Of these, 2,664 from the University of Vigo responded to the survey.

Sample

The students surveyed are: 2,383 undergraduate students and 281 graduate students. 1,625 declare to belong to the masculine gender: 1,625; to the female: 1,026; don't know/no reply: 13. Their average age is 21.31 years and the ages range from 17 to 73 years.

They belong to the different branches of knowledge, whose centers and degrees are distributed throughout the 3 campuses of the university, being their distribution the following: Art and Humanities: 402; Engineering and Architecture: 449; Science: 297; Health Sciences: 252; Social and Legal Sciences: 1.021.

Instrument

CUDECO (see annex) is set up by a total of 59 questions, 9 referring to subject identification variables and 47 Likert-type scale items grouped into five dimensions as follows:

- I) 12 items on the importance of teachers, examples: He gives me precise instructions how to prepare the paper; It gives me the basic rules to cite properly.
- II) 8 items on the usefulness of properly quoting, examples: It serves to refer to the original sources; It serves to recognize the merit of the original authors.

- III) 7 items on plagiarism in the different courses of the career, examples: I have delivered some paper done by someone else in a previous course; I have copied excerpts from printed sources and, without citing them, I have incorporated them into my written work.
- IV) 11 items on the causes that produce it, examples: My colleagues do it; access to online material is easy and convenient.
- V) 9 items on plagiarism in academic papers of their classmates, examples: They have copied excerpts of the papers they have submitted in previous courses as a new one; They have submitted a complete paper downloaded from the Internet, without modifying it, as produced on their own.

Procedure

In order to achieve the proposed objectives, the results obtained at the time were taken through the statistical analysis of the data corresponding to the answers provided by the students of the University of Vigo to the 47 items of the CUDECO and, with them, a matrix of 125,208 units of information was generated (2,664 students multiplied by 47 items answered by each of them). In other words, they reduce complexity without undermining the underlying information.

Statistical analysis techniques

To manage and understand so much information, it is necessary to use multivariate techniques, completing them with social network analysis. As stated above, they constitute valid procedures for reducing the number of information units while maintaining the existing information load. In other words, they allow complexity to be reduced without undermining the underlying information.

The following are the three procedures used in this work:

- PROXimity SCALing (PROXSCAL), used as an MDS with metric data, implemented in IBM SPSS 25. PROXSCAL maximizes the

determination of the input environment and the beginning of the algorithm.

- Scaling by MAyorizing a COMplicated Function (SMACOF), such as MDS with ordinal data, which is available as a package in R1. SMACOF follows the iterative processes of procedures such as Alternating Least Squares Scaling (ALSCAL), although improving its algorithm to minimize stress (López-González and Hidalgo, 2010).
- Social Media Analysis (SMA), a technique that allows to manage and make large volumes of information understandable. In this work we will focus on the visualization of the structures that result from the analysis of the data and for this the Gephi program will be used (Bastian, Heymann, & Jacomy, 2009)². It is an open source software for the analysis of graphs and networks that provides easy access to network data, as well as to spatialize, visualize, filter and cluster.

Results

Multidimensional scaling

PROXCAL and SMACOF output data are observed in Tables I and II. In both cases, the data shows that the two-dimensional model obtained has a good fit.

Thus, in Table I, a stress index of .01767 is the first indicator of this, since the higher the adjustment is worse. Similarly, values close to one, both in the dispersion explained (D.A.F.) and in the Tucker congruency coefficient, are also indicators of good adjustment (Arce, de Francisco, and Arce, 2010, pp. 47-48). Thus, a good fit of the model obtained to the distances derived from the coordinate matrix is observed (Real, 2001, p. 26).

¹ Through the link <https://www.r-project.org/>

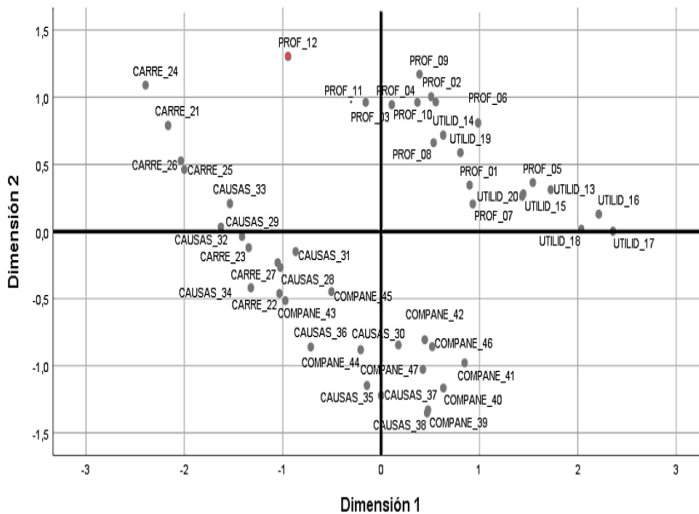
² Accessible in <https://gephi.org/>

TABLE I. Stress measurements and data adjustment using PROXSCAL

Normalized gross stress	.01421
Stress-I	.11919
Stress-II	.26719
S-Stress	.01767
Explained dispersion (D.A.F.)	.98579
Tucker congruence coefficient	.99287

Next, Graphic I, originated as a result of running the PROXSCAL procedure, shows the two-dimensional map of the responses given to the CUDECO items, which takes the values of the two dimensions as coordinates. The coordinate axes have been marked to make the quadrants visible when interpreting the results. Thus, by being able to visualize information and interactions between variables, possibilities are opened for the observer, through his cognitive-perceptive and even intuitive capacity, to perform an exploration and understanding of many units of information at the same time.

GRAPHIC I. Graphical representation of dimensions generated by PROXSCAL



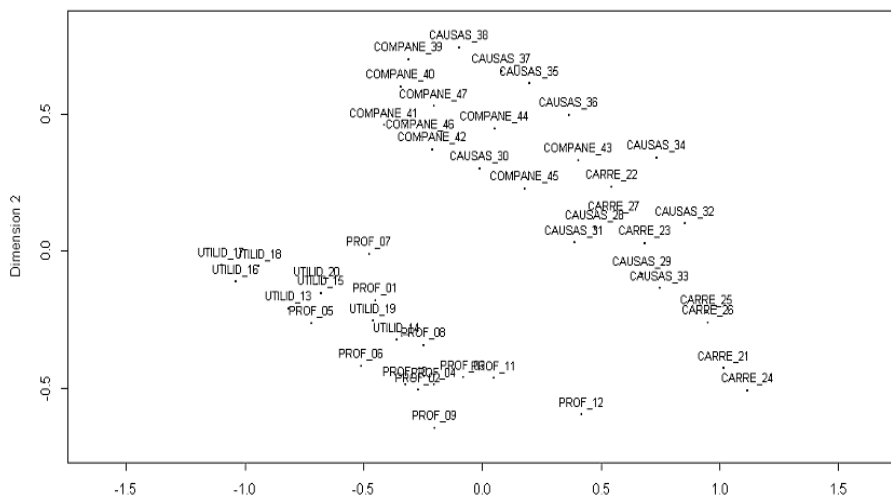
As in Table I, Table II shows that the Stress value is close to 0 and the number of interactions is 29, which also indicates that it is a well-adjusted two-dimensional model.

TABLE II. Output information of the SMACOF procedure

Call: mds(delta = dis, type = "ordinal")	Model: Symmetric SMACOF
Number of objects:	47
Stress-I value:	.118
Number of iterations:	29

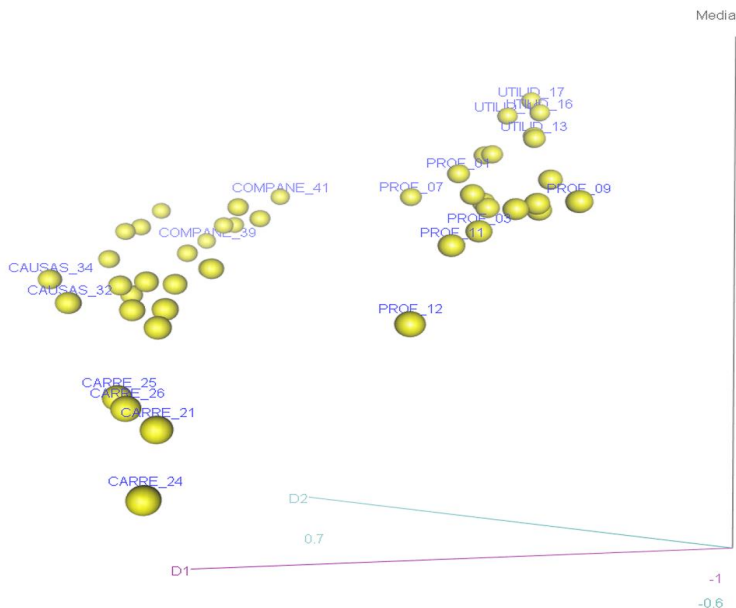
When the SMACOF procedure is executed with the data and a two-dimensional output is requested, the resulting graph obtained is similar to that produced by PROXSCAL (Graphic II).

GRAPHIC II. Graphical representation of dimensions generated by PROXSCAL



If the average value is added as a new coordinate to the SMACOF results, a three-dimensional representation of the results is obtained (see Graphic III).

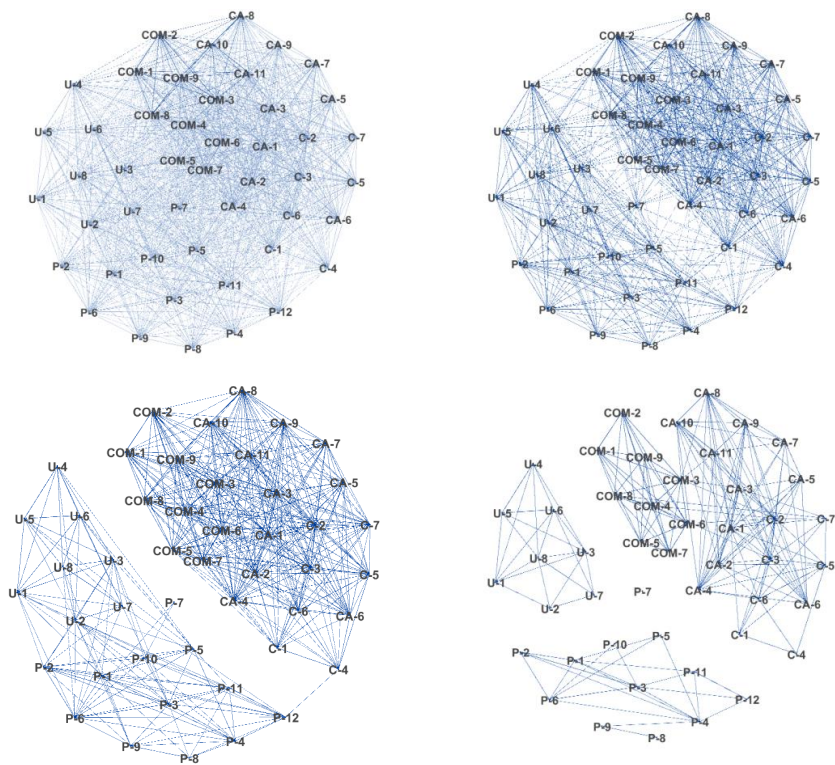
GRAPHIC III. Tridimensional representation



Social Media Analysis (SMA) Social Media Analysis (SMA) requires a set of nodes, in this case the 47 items that make up CUDECO and the relationship between them, using the correlation indexes between those items.

In Graphic IV it can be seen how the intensity of the relationships define groups, the same observed in the MDS results. The first graph shows all relationships, the second only those with correlations $r > .15$, the third $r > .2$ and the last one $r > .3$.

GRAPHIC IV. Data visualization according to the intensity of your relationships



Discussion of results

Although the discussion of results is difficult to be carried out due to the scarce existing research on plagiarism with methodology such as that used in our work, it can be said that works such as those of Soto-Ardila et al. (2019) and Soto-Ardila et al. (2020) are good indicators that we are facing very suitable methodologies to visualize underlying relationships that provide a greater understanding of facts investigated using statistical analysis methodologies.

However, the fact that groups of students who intentionally copy may be identified, as opposed to others whose actions relate more to ignorance and/or lack of information, pointing to the need for teacher assistance to overcome these dishonest behaviors, places us close to the results obtained by Abasi, & Graves (2008); Boillos (2020); Muñoz-Cantero et al. (2021); Saneleuterium (2017).

Conclusions

In view of the results obtained, the following can be concluded:

1st. In relation to objective 1, it can be said that the in-depth study of graphic representations improves the knowledge of the model and allows us to visualize and confirm relationships unknown until now:

- If we start from a panoramic view (Graphic 3) and we approach to identify groups we find the following:
 - The general vision shows us an arrangement of the variables in the form of V.
 - The average values of the variables are lower the closer they are to the axis.
 - From the axis (career-variable 24), each variable is arranged according to its value in each of the dimensions.
 - The distribution of values in one of the dimensions is uniform and continuous (teaching staff-variable 12 to usefulness-variable 17), while in the other dimension there is a jump in the values, which is greater when the average of the variables is higher, which generates the aforementioned arrangement in V (from career-variable 24, to peer-variable 41; from career-variable 24, to causes-variable 34).
 - Those items belonging to the utility and teachers factors are placed to one side; while causes and companions are placed to the other side. Those related to the career factor and, to a lesser extent, with causes and teachers are grouped at the axis.
- As we zoom in and get closer, we discover several groupings of variables (Table III) that, starting at the bottom (see Graphic V), would be the following:

- In the first grouping, with the values of the lowest average, those belonging to the career factor that refer to the whole or total copy would appear and, beside them, but alone, we find the item that refers to the lack of coordination of teachers.
- As part of the y axis of the chart, the copy of fragments from printed sources, associated with causes related to ignorance, slight sanctions or social apology (the others also do so), would be placed.
- In the same y axis, but with higher average values, the group of copies of fragments from the web and the notes of the faculty are found, associated with possibility causes.
- Moving to the x axis, but with similar averages, we observe a group that defines a facilitator teacher, who follows the work and provides personalized attention in the classroom, the papers that he requests are creative, adapts the time to the workload and uses anti-plagiarism tools.
- With higher average values, in the same x axis, a group that defines utility as a necessary requirement and scientific nuance in addition to allowing the demonstration of generating new information appears. In this group the teachers who pay individualized attention in tutorials is also present.
- Following, and with the highest average values, a group appears with items solely belonging to the utility factor, which refer to recognizing the merit of authors, referring to the original sources and allowing to support the arguments themselves.
- Finally, the group with higher values of the y axis groups the actions of partial copying by the colleagues. This includes the delivery of papers written by some colleagues from previous years, associated with causes related to work overload, lack of time, ease of doing so and convenience of internet access.

GRAPHIC V. Concepts resulting from groupings of CUDECO items

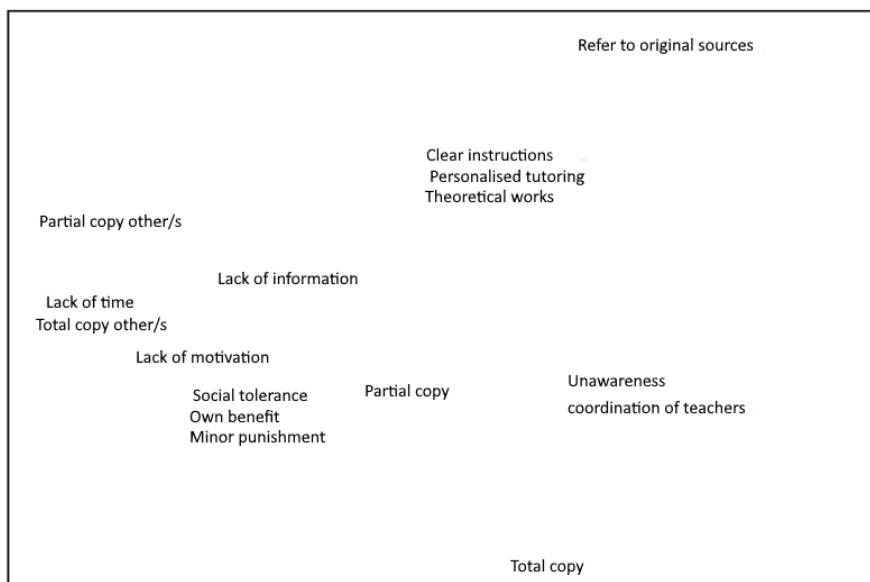


TABLE III. Groupings of variables into more general concepts

Full copy	Career-24	I submitted as my own some complete paper downloaded from the Internet, without modifying it
	Career -21	I submitted a paper written by someone else in a previous course
	Career -26	I did some work entirely from printed sources, without citing the authors
	Career -25	I did some work entirely from fragments literally copied from web pages

Minor sanctions Social tolerance Partial copy Self-leverage	Career -33	Penalties are not serious
	Causes-29	My classmates do it
	Career-23	I copied excerpts from printed sources (books, newspapers, magazine articles, etc.) and, without citing them, incorporated them into the paper I had written.
	Causes-31	It gives me the chance of getting better academic results
	Causes-28	It is a "shortcut" that is generally accepted
	Career-27	I used excerpts from my teachers' notes to elaborate some paper, without mentioning them
	Career-22	I copied fragments of texts from web pages and, without citing the sources, I incorporated them into the paper I had written.
Copy all other/s Lack of information Ignorance Lack of time Lack of motivation	Colleagues-43	They submitted a complete paper downloaded from the Internet, without modifying it, as their own
	Colleagues-45	They did entirely a job from printed sources
	Causes-34	I was unaware that citing sources is mandatory.
	Causes-30	Access to online material is easy and convenient
	Causes-35	Lack of precise instructions on how to write a paper
	Causes-36	Lack of motivation
	Causes-37	Lack of time
Partial copy other/s	Colleagues -39	They submitted some paper written by someone else in a previous course
	Colleagues -40	They copied excerpts from papers submitted in previous courses as a new one
	Colleagues -41	They copied excerpts of texts from web pages and, without citing the sources, incorporated them into the paper they had written.
	Colleagues -42	They copied excerpts from printed sources (books, newspapers, magazine articles, etc.) and incorporated them into the paper they had written without citing the sources.
	Colleagues -44	They submitted a paper composed solely of literal excerpts extracted from online sources.
	Colleagues -46	They used excerpts from their teacher's notes to write a paper, without citing them
	Colleagues -47	They admit as appropriate the "cut and paste" method when submitting a paper
Teaching staff coordination	Teaching staff-12	He coordinates with other teachers to avoid saturating students with work

Ignorance	Causes-32	I was unaware of regulations at my university penalizing this practice
Teaching staff usefulness	Teaching staff-9	He requests creative/innovative paper
	Teaching staff-11	He adapts the workload to the time available for each subject
	Teaching staff-4	He offers individualized attention in the classroom during the preparation of the paper
	Teaching staff -2	He provides me with the basic rules to cite properly
	Teaching staff -6	He evaluates the correct citing of the incorporated documentary sources
	Teaching staff -8	He requires essentially practical papers
	Usefulness-14	It serves to demonstrate that I learned properly
	Usefulness -19	It is used to generate discussions or debates from the mentioned one
	Teaching staff -1	He provides me with precise instructions to prepare the paper
	Usefulness -13	It serves as a necessary requirement in the realization of academic papers
	Usefulness -15	It serves to give a scientific nuance to my paper
	Usefulness -20	It serves to demonstrate that I can generate new information or ideas from the cited one
Theoretical papers Refer to original sources	Teaching staff -7	He requires essentially theoretical papers
	Usefulness -16	It serves to recognize the merit of the original authors
	Usefulness -18	It supports my arguments

2nd. If we take into account objective 2 and the achievements obtained in this regard, it can be said that the suitability of the triangulation of analysis has been demonstrated in the study of the dimensionality underlying the responses of the items that make up the CUDECO questionnaire. Undoubtedly a fact indicative of the importance and usefulness of this test to achieve a better knowledge of the underlying dimensions that may exist in a psycho-pedagogical or socio-pedagogical event.

3rd. With respect to objective 3, it has been possible to construct a graphic representation of the plagiarism model of the students of the University of Vigo from their responses to the CUDECO items, which allows a greater understanding of it, suggesting new hypotheses and the corresponding analyses.

As a summary, when interpreting the reality of the phenomenon of plagiarism in university students of UVigo through CUDECO, two different meta-dimensions are identified, in their spatial representation, that group the five dimensions included in the questionnaire.

The first, with high values, combines the usefulness of citation with the actions of teachers to avoid plagiarism behavior. This places us facing the importance of teaching actions in the classroom and office hours to avoid student plagiarism, and even to correct them immediately if they occur.

In this regard, it is possible to propose, in the form of hypotheses, didactic actions that, probably pointing to the usefulness that it has for any student to know how to quote, try to clarify the reasons that exist to do it well, at the same time that they take advantage of the opportunities provided by the teaching-learning process of a subject (from the presentation of the teaching guides to the evaluation by competences). This entails doing so on the basis of the teachers' example and the tutelage of the students in the classroom work, including the use of tools (Turnitin, 2021). The design, implementation and evaluation of organizational measures for teacher coordination and facilities for the management of regulations and guidelines for classroom papers, end of grade or end of master's degree papers, which are normally available in faculties, schools and university library services are required. The handling of some papers as examples and sources of critical comments will also be useful.

The second meta-dimension includes the plagiarism actions carried out by the students throughout their careers, with the lowest values, which contrasts, with high values, with those carried out by their classmates. And in the midst of the actions of plagiarism the causes that lead to its realization appear. This relationship places us before a type of student who interprets plagiarism as a socially tolerated fact and that in university centers, if discovered, is hardly penalized. Today, with the Internet, total and partial copies can be made according to the objective without this being described as unethical, because what is on the netpaper is public. It is something that, especially if it is partial copy, most of the colleagues do, because they need, like them, to get the papers and the subjects forward. In addition, what is on the Internet, being public, copying it and pasting it cannot be considered unethical. The justifications that are based on the lack of coordination of the teachers and the limited time

available to the students, two usual attempts of justification, can probably be added to this.

In view of this position, as a complementary proposal to the previous one, and in view of the revised bibliography, in addition to the basic explanations of what intellectual property law means, it is appropriate to use the university's anti-plagiarism tools, as well as existing regulations and rules which include punishment for the taking over of others' intellectual property for the sake of their own benefit. This will require working with the students, in each subject and in the corresponding office hours devoted to the papers to be carried out, on the meaning of the detected coincidences and the need to mention and refer to the authorship of the material. Some considerations about the integrity of individuals as one of the values that must characterize university life are likely to be required. On the other hand, the poor coordination of teachers due most of the time to causes beyond their control, must require the attention of the decanal and rectoral teams.

Limitations and Prospects

The limitations of this research study can be found in the superficial treatment that has been given to the data, without having delved deeper in the social media analysis (SMA). Hence, in these work the SMA is presented as a supplementary technique of multidimensional scaling (MDS). On the other hand, the fact that the sample is limited to the University of Vigo only and despite its sufficiency and representativeness in the branches of knowledge, could include some bias related to aspects of the aforementioned university.

As for future work that should be done taking into account the intuitions derived from the information represented in the Graphs, it can be said that we are open to the approach of new hypotheses whose verification may require moving from intuitions to analytical processes. In particular, aspects such as the existence of two meta-dimensions isolating in groups, three and two respectively, the initial dimensions of the measuring instrument are indicating that it is probably appropriate to carry out a review of some aspects of the CUDECO, such as the inclusion of a new dimension or dimensions covering the space between the two discovered meta-dimensions. Likewise, in subsequent researches, it would

be necessary to respond to the limitations set out above with a sample of the entire university population of the three universities of the Galician University System and, to the same extent, to deepen the analysis of the data collected, either with MDS or SMA or any other analytical technique that can be used.

References

- Abasi, A., & Graves, B. (2008). Academic literacy and plagiarism: Conversations with international graduate students and disciplinary professors, *Journal of English for Academic Purposes*, 7 (4), pp. 221-233, <https://doi.org/10.1016/j.jeap.2008.10.010>.
- Aldas, J., & Uriel, E. (2017). *Análisis multivariante aplicado con R* (2.^a ed.). Madrid. Paraninfo.
- Alzás, T., Casa, L. M., Luengo, R., Torres, J. L., & Catarreira, S. V. (2016). Revisión metodológica de la triangulación como estrategia de investigación. *Investigación Cualitativa en Ciencias Sociales*, 3, pp. 639-648.
- Amiri, F., & Razmjoo, S. A. (2016). On Iranian EFL undergraduate students'percepción of plagiarism. *Journal of Academic Ethics*, 14, pp. 115-131. doi:<https://doi.org/10.1007/s10805-015-9245-3>
- Arce, C., de Francisco, C. y Arce, I. (2010). Escalamiento multidimensional: concepto y aplicaciones. *Papeles del Psicólogo*, 31(1), pp. 46-56.
- Arias, M. (2000). La triangulación metodológica: sus principios, alcances y limitaciones. *Investigación y Educación en Enfermería*, XVIII (1), pp. 37-57.
- Badenhorst, C. (2018). Citation practices of postgraduate students writing literature reviews. *London Review of Education*, 16(1), pp. 121-135. <https://doi.org/10.18546/LRE.16.1.11>
- Bastian, M., Heymann S., & Jacomy M. (2009). Gephi: un software de código abierto para explorar y manipular redes. *Conferencia Internacional sobre Web y Redes Sociales (ICWSR)*. *Informática*. Recuperado de: <https://medialab.sciencespo.fr/publications/Gephi%20paper.pdf>

- Bennett, J. F., & Hays, W. L. (1960). Multidimensional unfolding: Determining the dimensionality of ranked preferente data. *Psychometrika*, 25, pp. 27-43.
- Boillos, M. M. (2020). Las caras del plagio inconsciente en la escritura académica. *Educación XX1*, 23(2), pp. 211-229, doi: 10.5944/educXX1.25658
- Bright, L., Heesen, R., & Zucker, A. (2016). Vindicating methodological triangulation. *Synthese*. doi: 10.1007/s11229-016-1294-7.
- Cebrián-Robles, V., Raposo-Rivas, M., Cebrián-de-la-Serna, M., & Sarmiento-Campos, J. A. (2018). Percepción sobre el plagio académico de estudiantes universitarios españoles. *Educación XX1*, 21(2), 105-129 <https://doi.org/10.5944/educxx1.20062>
- Cebrián-Robles, V., Raposo-Rivas, M., & Sarmiento-Campos, J.A. (2016). ¿Ética o prácticas deshonestas? el plagio en las titulaciones de educación. *Revista de Educación*, 374, pp. 161-186. [10.4438/1988-592X-RE-2016-374-330](https://doi.org/10.4438/1988-592X-RE-2016-374-330)
- Childers, D., & Bruton, S. (2016). "Should It Be Considered Plagiarism?" Student Perceptions of Complex Citation Issues. *Journal of Academic Ethics* 14, pp. 1-17. <https://doi.org/10.1007/s10805-015-9250-6>
- Comas, R., Sureda, J., & Oliver, M. (2011). Prácticas de citación y plagio académico en la elaboración textual del alumnado universitario, en Hernández, M. J., & Fuentes, M. (Coords.) *La red como recurso de información en educación. Revista Teoría de la Educación: Educación y Cultura en la Sociedad de la Información*. 12(1). Universidad de Salamanca, pp. 359-385 http://campus.usal.es/~revistas_trabajo/index.php/revistatesi/article/view/7837/7863
- Comas-Forgas, R., Morey-López, M., & Sureda-Negre, J. (2021). La publicidad en buscadores de las plataformas españolas de compraventa de trabajos académicos: análisis del tráfico, costes y palabras clave. *Revista Española de Documentación Científica*, 44 (3), pp. 1-15. <https://doi.org/10.3989/redc.2021.3.1767>
- Denzin, N. K. (1970). *The research act*. Chicago: Aldine.
- Denzin, N. K. (1989). *Strategies of Multiple Triangulation. The Research Act: A Theoretical Introduction to Sociological Methods*. New York: Mc Graw Hill.
- Duche, A. B., Arias, D., Ramos, T., & Gutiérrez, O. A. (2020). Representaciones sociales de estudiantes universitarios peruanos sobre el plagio en la escritura académica. *Revista Conrado*, 16(72),

- pp. 155- 162. Epub 02 de febrero de 2020. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442020000100155&lng=es&tlng=es.
- Espiñeira-Bellón, E.-M., Muñoz-Cantero, J.-M., Gerpe-Pérez, M.-E., & Castro-Pais, M.-D. (2021). Ciberplagio como soporte digital en la acreditación de trabajos académicos. *Comunicar*, 48. <https://doi.org/10.3916/C68-2021-10>
- Gaete J., & Vasquéz, J. (2008). Conocimiento y estructura en la investigación académica: una aproximación desde el análisis de redes sociales. *REDES. Revista hispana para el análisis de redes sociales*, 14, 5.
- Guerrero, F. M., & Ramírez, J. M. (2002). El análisis de escalamiento multidimensional: una alternativa y un complemento a otras técnicas multivariantes. En *X Jornadas ASEPUMA, Asociación Española de Profesores Universitarios de Matemáticas para la economía y la empresa*, pp. 1-11.
- Hahs-Vaughn, D. L. (2017). *Applied multivariate statistical concepts*. Nueva York: Routledge.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis*. Seventh Edition. Harlow, England: Pearson New Internacional Edition.
- Johnson, R., & Wichern, D. (2019). *Applied multivariate statistic analysis* (6.^a ed.). Nueva Jersey: Pearson
- Kimchi, J, Polivka, B., & Stevenson, J.S. (1991). Triangulation: operational definitions. *Nursing Research*, 40 (6), pp. 364-366.
- Kruskal, J. B. (1964a). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29, pp. 1-27.
- Kruskal, J. B. (1964b). Nonmetric multidimensional scaling: a numerical method. *Psychometrika*, 29, pp. 115- 129.
- López-González, E., & Hidalgo, R. (2010). Escalamiento Multidimensional no Métrico. Un ejemplo con R empleando el algoritmo SMACOF. *ESE: Estudios sobre Educación*, 18, pp. 9-35.
- Mitchell, M. (2009). Complexity: *A Guided Tour*. Reino Unido: Oxford University Press.
- Moxley, K., & Archer, A. (2019). Writing centres as dialogic spaces: Negotiating conflicting discourses around citation and plagiarism. *Stellenbosch Papers in Linguistics Plus*, 57, pp. 25-37 <https://doi.org/10.5842/57-0-808>

- Muñoz-Cantero, J. M., Rebollo-Quintela, N., Mosteiro-García, J., & Ocampo-Gómez, C. I. (2019). Validación del cuestionario de atribuciones para la detección de coincidencias en trabajos académicos. *RELIEVE*, 25(1), art. 4. doi: <http://dx.doi.org/10.7203/relieve.25.1.13599>
- Muñoz-Cantero, J. M., Espiñeira-Bellón, E.M., & Pérez-Crego, M. C., (2021). Medidas para combatir el plagio en los procesos de aprendizaje. *Educación XX1*, 24 (2), pp. 97– 120 <https://doi.org/10.5944/educXX1.28341>
- Paniagua, J. A. (2013). *Curso de Análisis de Redes Sociales: metodología y estudios de caso*. Granada: Editorial Universidad de Granada.
- Pavlopoulos, G., Paez-Espino, D., Kyrpides, N., & Iliopoulos, I. (2017). Empirical Comparison of Visualization Tools for Larger-Scale Network Analysis, *Advances in Bioinformatics*, vol. 2017, Article ID 1278932, 8 pages. <https://doi.org/10.1155/2017/1278932>
- Pérez, J. (2000). La triangulación analítica como recurso para la validación de estudios de encuesta recurrentes e investigaciones de réplica en Educación Superior. *RELIEVE*, v. 12, n. 2. Obtenido el 7 de Enero en http://www.uv.es/RELIEVE/v12n2/RELIEVEv12n2_6.htm.
- Pituch, K. A., & Stevens, J. P. (2016). *Applied multivariate statistics for the social sciences*. (6.^a ed.). Nueva York: Routledge.
- Real Academia Española (R. A. E.) (2014). *Diccionario de la Lengua Española (23.^a edición)*. Madrid: Espasa Calpe.
- Real, J. E. (2001). *Escalamiento multidimensional*. Madrid: Hespérides.
- Rodríguez, C., Pozo, T., & Gutiérrez, J. (2006). La triangulación analítica como recurso para la validación de estudios de encuesta recurrentes e investigaciones de réplica en Educación Superior. *RELIEVE*, 12, 2. Recuperado de: http://www.uv.es/RELIEVE/v12n2/RELIEVEv12n2_6.htm.
- Sandoval, S., Morales, M. A., & Díaz, H. E. (2019). Estrategia de escalamiento en las cadenas globales de valor: el caso del sector aeroespacial en México. *Entre-Ciencias*, 7(20), pp. 35-52. doi: 10.22201/enesl.20078064e.2019.20.68425
- Saneleuterio, E. (2017). A citar se aprende citando: experiencia didáctica en alfabetización académica. *Estudios lambda. Teoría y práctica de la Didáctica en Lengua y Literatura*. 2(1), pp. 143-159. <https://doi.org/10.36799/el.v2i1.44>
- Sarmiento, J. A., Ocampo, C. I., & Cid, A. (2020). El análisis de redes sociales aplicado al currículo prescrito. Un ejemplo en educación primaria.

- Profesorado. Revista de Currículum y Formación de Profesorado*, 24(2), pp. 294-323. doi: 10.30827/profesorado.v24i2.14086
- Shepard, R. N. (1962). The analysis of proximities: multidimensional scaling with an unknown distances function (I y II). *Psychometrika*, 27, 125-139, pp. 219-246.
- Soto-Ardila, L. M., Caballero, A., Alzás, T., Veríssimo, S., & Casas, L. M. (2019). Actitudes ante las Matemáticas: Análisis cualitativo del conocimiento grupal mediante software de representación de redes sociales. En Costa, Oliveira, Castro y de Souza (Eds.), *Congreso Ibero-americano en Investigación Cualitativa*, (pp.169-178). Lisboa, Portugal.
- Soto-Ardila, L. M., Caballero, A., Carvalho, J. L., & Casas, L. M. (2020). Nuevo método de análisis cualitativo mediante software para el análisis de redes sociales de la percepción grupal hacia las Matemáticas. *Pixel-Bit. Revista de Medios y Educación*. 58, pp. 27-50. <https://doi.org/10.12795/pixelbit.73356>
- Sureda-Negre, J., Comas-Forgas, R. L., & Oliver-Trobat, M. F. (2015). Plagio académico entre alumnado de secundaria y bachillerato. Diferencias en cuanto al género y la procrastinación. *Comunicar*, 44 (22), pp. 103-111. doi:<https://doi.org/10.3916/C44-2015-11>
- Thurmond, V. A. (2001). The point of triangulation. *Journal of Nursing Scholarship*, 33(3), pp.253–258. doi:10.1111/j.1547-5069.2001.00253.x.
- Torgerson, W. S. (1958). *Theory and methods of scaling*. Nueva York: Wiley.
- Turnitin LLC. (2021). Executive guide: plagiarism spectrum 2.0 [Archivo pdf] https://go.turnitin.com/1/45292/2020-07-24/bmqp61/45292/211420/Originality_Plagiarism_Spectrum_2.0_Administrators_Ebook.pdf?_ga=2.33190142.1555813439.1626942459-1269377846.1626942459
- Vargas-Franco, A. (2019). Apropiación y plagio académico: un estudio de caso sobre una alumna debutante en la escritura en la educación superior. *Íkala, revista de lenguaje y cultura*. 24(1), pp. 155-179. <http://dx.doi.org/10.17533/udea.ikala.v24n01a08>

Contact address: María Dolores Castro Pais, Universidad de Vigo, Facultad de Educación y Trabajo Social y Departamento de Didáctica, Organización Escolar y Métodos de Investigación. As Lagoas s/n, CP, 32004, Ourense. E-mail: maria.dolores.castro.pais@uvigo.es

Anexo

CUDECO variables grouped by dimensions

The teachers...

1. He provides me with precise instructions to prepare the paper
2. He provides me with the basic rules to cite properly
3. Track the evolution of the work taking into account the citation of sources
4. He offers individualized attention in the classroom during the preparation of the paper
5. Offers individualized attention in tutorials during the preparation of the work
6. He evaluates the correct citing of the incorporated documentary sources
7. He requires essentially theoretical papers
8. He requires essentially practical papers
9. He requests creative/innovative paper
10. Handles detection tools for copied fragments in jobs
11. He adapts the workload to the time available for each subject
12. He coordinates with other teachers to avoid saturating students with work

How useful is it for you to quote?

1. It serves as a necessary requirement in the realization of academic papers
2. It serves to demonstrate that I learned properly
3. It serves to give a scientific nuance to my paper
4. It serves to recognize the merit of the original authors
5. It serves to refer to the original sources
6. It supports my arguments
7. It is used to generate discussions or debates from the mentioned one
8. It serves to demonstrate that I can generate new information or ideas from the cited one

Throughout the race career..

1. I submitted a paper written by someone else in a previous course
2. I copied fragments of texts from web pages and, without citing the sources, I incorporated them into the paper I had written
3. I copied excerpts from printed sources (books, newspapers, magazine articles, etc.) and, without citing them, incorporated them into the paper I had written
4. I submitted as my own some complete paper downloaded from the Internet, without modifying it
5. I did some work entirely from fragments literally copied from web pages
6. I did some work entirely from printed sources, without citing the authors
7. I used excerpts from my teachers' notes to elaborate some paper, without mentioning them

Causes that have motivated you to carry out the previous actions

1. It is a "shortcut" that is generally accepted
2. My classmates do it
3. Access to online material is easy and convenient
4. It gives me the chance of getting better academic results
5. I was unaware of regulations at my university penalizing this practice
6. Penalties are not serious
7. I was unaware that citing sources is mandatory
8. Lack of precise instructions on how to write a paper
9. Lack of motivation
10. Lack of time
11. Work overload

I think my colleagues...

1. They submitted some paper written by someone else in a previous course
2. They copied excerpts from papers submitted in previous courses as a new one
3. They copied excerpts of texts from web pages and, without citing the sources, incorporated them into the paper they had written
4. They copied excerpts from printed sources (books, newspapers, magazine articles, etc.) and incorporated them into the paper they had written without citing the sources.
5. They submitted a complete paper downloaded from the Internet, without modifying it, as their own
6. They submitted a paper composed solely of literal excerpts extracted from online sources
7. They did entirely a job from printed sources
8. They used excerpts from their teacher's notes to write a paper, without citing them
9. They admit as appropriate the "cut and paste" method when submitting a paper