Multimedia Annotations for Practical Collaborative Reasoning

Manuel Cebrián-de-la-Serna¹, María Jesús Gallego-Arrufat² and Violeta Cebrián-Robles³

¹Departamento de Didáctica y Organización Escolar, Facultad de CC. de la Educación, Málaga University, Spain
²Departamento de Didáctica y Organización Escolar, Facultad de CC. de la Educación, Granada University, Spain
³Departamento de Didáctica, Organización Escolar y Métodos de Investigación, Facultad de Educación y Trabajo Social, Vigo University, Spain

ABSTRACT

University education requires students to be trained both at university and at external internship centres. Because of Covid-19, the availability of multimedia resources and examples of practical contexts has become vital. Multimedia annotation can help students reflect on the professional world, collaborating and interacting with colleagues online. This study aims to encourage collaborative practical thinking by using new video annotation technologies. 274 students participated in an experiment of task design focusing on the analysis of a technology-based, award-winning educational innovation project. With mixed research design, qualitative and quantitative data exported from the video annotation platform used was collected and analysed. The results show differences in the quality and quantity of the answers: in the tasks with broad Folksonomy they are more numerous but more dispersed in their analysis, and vice versa. The quality of the answers given with narrow Folksonomy is also higher in both texts and videos modes. Producing multimedia annotations is a practical way to encourage students to practise reflective reasoning about the professional reality.

Keywords PRESERVICE TEACHER EDUCATION, TASK ANALYSIS, EDUCATIONAL RESOURCES, MULTIMEDIA ANNOTATION, REFLECTIVE TEACHING

1 INTRODUCTION

The process of skill acquisition by university students requires an optimization of the relationship between practice and theory in academic and professional contexts, using multimedia resources and in online environments. These should mirror the professional reality as closely as possible even in situations where virtual audio-visual strategies need to be used. Understanding how to deliver this practice means analysing electronic tasks (Kandari, Qattan, & M, 2020). This can be done through multimedia annotation that helps interpret the message.

It is necessary to focus on both methodologies and technologies that allow the analysis of these multimedia messages. Together, they have more of an impact (Becker, 2010). An
annotation is a portion of information that is associated with a piece of original content in order to explain something about that content or to add more information (Gayoso-Cabada, Sarasa-Cabezauelo, & Sierra-Rodriguez, 2019). In academic settings, annotations can be used by students and teachers to tag and highlight texts, images, songs, websites, videos and other resources (Novak, Razzouk, & Johnson, 2012; Sauli, Cattaneo, & van der Meij, 2018; Smith, Blankinship, & Lackner, 2000; Zhu, Chen, Avadhanam, Shui, & Zhang, 2020). This study uses multimedia annotations (hereafter MA) to select and break down these messages to provide reasonings in the shape of commentaries and interpretations as well as sharing them through social tags. This is defined as "the collective action of users associating tags to resources they have created and experimented with" (Lau, Lee, & Singh, 2015).

Although their origin dates back to papyrus annotations in ancient Greece (Muellner, 2015) they have experienced a boom with the development of MA software, especially image and video annotation software, with promising results in terms of user usability and effectiveness in video analysis (Bianco, Ciocca, Napoletano, & Schettini, 2015; Chen, Chen, Xu, March, & Benford, 2007; Monedero, Cebrián, & Desenne, 2015).

Many studies on shared MAs have been carried out in the academic field (Colasante, 2011; Dias-Pereira-Dos-Santos, Loke, & Martinez-Maldonado, 2018; McFadden, Ellis, Anwar, & Roehrig, 2014; Novak et al., 2012; Paradis & Fendt, 2016; Pérez-Torregrosa, Diaz-Martín, & Ibáñez-Cubillas, 2017; Sauli et al., 2018; Su, Yang, Hwang, & Zhang, 2010; Sydnor, 2016). At present, before and during the COVID-19 pandemic, they have been used successfully for the development of tasks in online teaching programs (Zhu et al., 2020).

The research focuses on the analysis of the relationship between theory and practice through students’ reasoning and argumentation by reading, viewing, reflecting and writing commentaries in MAs. To begin with, the argumentations were analysed in order to develop reflective practices in academic and professional settings (D. Cebrián, Pérez, & Cebrián, 2017; Liu & Stapleton, 2014; Nussbaum, Sinatra, & Poliquin, 2008). Evidence shows that "scaffolding computer-mediated discussions can improve the quality of argumentation in students’ writing" (Özçinar, 2015).

Students’ thinking must be constructed from the analysis of professional good practices, that is why their presentation methods are an object of research. This is especially the case in written texts and/or videos (Debbag & Fidan, 2020; Hefter & Berthold, 2020; Lee & List, 2019). Texts allow for a more relaxed reading of the message which becomes more contextualized when texts and images are combined to produce argumentation and exchange of meanings (Smith et al., 2000). At the same time, the speed of the video can be changed to emphasize the emotional parts and the maker’s guided narrative (Imran, Cheikh, & Kowsalski, 2016).

Various studies have proposed collaborative annotation of learning resources; however, little research has been carried out on the classification mechanisms used in the annotation tools. The research by Gayoso-Cabada et al. (2019) identifies four mechanisms: classifications based on controlled vocabulary, folksonomies, ontologies and absence of classifica-
tion mechanisms. Annotation classification plays an essential role in the application of MA in education, but the main problem with folksonomies is the open nature of these terms. Collaborative tagging systems, also known as folksonomy, have gained popularity as they easily organize resource content (web pages, images and videos, among others) using open tags. Users can therefore provide information and create a rich and growing corpus of social knowledge that can be used by recommendation technologies (Godoy & Corbellini, 2016).

In this context, the folksonomy’s three-way relationship between users, resources and tags presents new challenges, the goal being to help users by means of recommendation systems (Hsu, 2013; Lau et al., 2015).

As folksonomy research is still quite new, the theoretical perspective and research methods are still being developed. In the university environment multimedia resources are used to illustrate the process, with videos being widely employed and preferred by students; however, the strategies that students use to understand its contents are not known in depth (List, 2018; List & Ballenger, 2019). Moreover, it is also important to find out which is the best design to encourage the critical and reflective learning of these realities and to better understand how the design of these methodologies and tasks that are taught with multimedia resources influence the quality of student annotations.

The purpose of this study is to understand how the type of format chosen and the instructions given in the tuition exercises and tasks influence the quality of the students’ annotations. The idea is also to implement other innovative educational processes through task assignments: guiding students through the annotation process, explaining to them how the process is to be carried out and giving teachers information on how students write and reason about the practical and innovative content. Hence the need to investigate the effect of these MA mechanisms.

This study is part of a broader line of research, that, in this paper, will focus on the analysis of the quality of responses generated by students when they use MAs to comment and reason about an innovative educational project, based on the format or code (text vs. video) and the instructions that are supplied in the task (make narrow annotations –with tags given by the teacher– vs. broad annotations –with no given tags). The objectives of this research are the following: 1. Analysing the levels of quality and quantity of the responses generated by students when they make annotations to define and justify an innovative educational project. 2. Studying to what extent the pre-set tags given by the teacher to guide the task and the format used (text vs. video) affect the variation of the MAs generated by the students.

We thank Professor Kurt E Fendt from MIT’s Digital Humanities Laboratory who allowed us to use the HyperStudio tool at Vidanet Project https://cutt.ly/tmhQmWd

Proyecto para el desarrollo de Open Video Anotaciones -OVA- para la plataforma MOOC de Edx (Monedero-Moya, Cebrian-Robles & Desenne, 2015).
2 MATERIAL AND METHODS

2.1 Participants
274 Educational Science undergraduate students (mean age 20 years) took part in this study. They are studying Educational Technology in Spain and know and understand the purpose of the research, which has been suggested by the teachers (among several others on offer), with which they agree and that does not have a direct impact on their grades. The information was collected over two academic years.

2.2 Design
The study is mixed research –quantitative and qualitative– and the instrument is in the tool itself Coannotation.com which collects all the steps and data generated by users in the proposed activities. Being these data collected and analysed both qualitative and quantitative natura, as shown in Figure 1, above right, where we can also see the button to export the data to excel format for analysis. The replication of this study can be done with the same tool and activity description proposed below. The MAs written by the students were analysed quantitatively and qualitatively through shared messages in an innovative educational project entitled Class of Clans. The project was selected by Fundación Telefónica from the 100 best in the country. Focusing on gamification, it has won awards from SIMO Education and the Spanish Ministry of Education (MEC) for educational innovation and for teaching teams. It is based on the integration of four subjects from the first year of secondary education: Natural Sciences, Social Sciences, Technology and Arts.

The MAs were divided into groups, according to the independent variables of the study: on the one hand, folksonomy (broad vs. narrow) and on the other hand, the message code (video vs. text). In the first group the MAs were analysed according to the task's instructions, i.e. making a free annotation (broad folksonomy) with tags created by the students, or to provide them with tags (narrow folksonomy) in order to produce said annotations. In the second, i.e. the message code, the content of the messages was prepared for dissemination by the authors of the project. The text message contains structured information and complies with the categories of the educational innovation project (Fundación Telefónica, 2016, 68-70). The video message does not use voiceover, but rather recorded images of student actions, group tasks, activities in the school garden... with a very emotive soundtrack and text labels to involve the viewer in the challenge of the game.

2.3 Procedure
The argumentation on the texts and videos created with the MA methodology through online groups discussions, was carried out using new tools such as Coannotation.com (developed in the [3] project), which creates a layer of annotations from a YouTube video. Annotations are exported to Excel to be studied. Google Drive and Annotation Studio (http://uma.annotationstudio.org/) were used for text annotations. The latter is a tool created by MIT’s Digital Humanities Laboratory’ HyperStudio (EEUU) http://hyperstudio.mit.edu (Paradis & Fendt, 2016) and used in the Vidanet project to test the educational institutions...
participating in this study.

The instructions given to the subgroups vary according to the folksonomy, not to the two codes (text and video). For the broad folksonomy the instructions are: “Read and analyze the text and/or watch the video of the Class of Clans project. Next, make four annotations of the most important aspects that the message triggers in you and the reasons why you consider it to be an innovative project, reasoning and justifying in each annotation why it is innovative”. For the narrow folksonomy they are: “Read and analyze the text and/or watch the video of the Class of Clans project. Next, make four annotations of the most important aspects that the message triggers in you and the reasons why you consider it to be an innovative project, discussing and justifying why it is innovative in each annotation”. Always use these four tags: **Problem**: What problem does innovation try to solve? **Solution**: What solution does innovation suggest? **Evidence**: what type of impact does ICT have? **Competence**: What skills does this innovation foster?”

The study sample was chosen for convenience with 274 students from different degrees (Pedagogy and Primary Degree) from the Faculties of Education of the University of Granada and Malaga for two academic years 2016-17 and 2017-18. The students were distributed into subgroups according to the instructions and codes, obtaining a total of 845 MAs (Table 1). However, in the text annotations task fewer annotations were collected because of a logistical problem unrelated to the research.

### Table 1  Annotations according to the two research variables

<table>
<thead>
<tr>
<th>No. of students</th>
<th>Broad annotations</th>
<th>Narrow annotations</th>
<th>Total video</th>
<th>Total text</th>
<th>Total annotations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>video</td>
<td>text</td>
<td>video</td>
<td>text</td>
<td></td>
</tr>
<tr>
<td>274</td>
<td>243</td>
<td>74</td>
<td>461</td>
<td>67</td>
<td>141</td>
</tr>
</tbody>
</table>

#### 2.4 Data Analysis

After the students created the MAs, the researchers exported the annotations to Excel and then to SPSS for quantitative analysis. Qualitative and quantitative analysis are combined. The treatment of the data implies two strategies of analysis: the first one using Coannotation.com’s statistical graph where the annotations (yellow lines) are grouped into "peaks" and "valleys" (Figure 1).

This quantitative analysis helped focus the attention on certain areas, for example, whether there are more or fewer annotations on the video timeline (which we named "peaks" and "valleys" respectively), as well as the result of the MA quality rating scale. In the second approach, all the annotations, both narrow and broad, are analyzed according to the four initial tags, categorizing the broad ones and adding two more extracted from the data (**Good Practices** – the innovative educational project represents a good practice that can be applied in other situations – and **Others** – different comments not included in any tag). This qualitative analysis of MAs is carried out through content analysis by categories, with the idea of performing a "Q-analysis" or "Connectivity Analysis" (Buendía, Colás, & Hernán-
The quality of the responses is also evaluated with a rating scale of 0 to 3, extracted from an argumentative rubric (D. Cebrián et al., 2017). Value 0 means that the answer has no relation to the tag or question posed in the task in general. Value 1 applies to MAs that contain a mere statement and/or assessment without justification. Value 2 is applied when the student tries to explain and/or justify and Value 3 when there is a broad explanation, reasoned and justified, that denotes a deeper reading and analysis of the message. These levels were not known by the students either before or after the task; what was basically taken into account in their assessment was the coherence with the tag rather than other comments and/or reasoning.

### 3 RESULTS

#### 3.1 Tags in the MAs

First, we analysed the quality and quantity of responses of the students’ annotations on Coannotation.com. The resulting percentage of tag appearance is: Competence (20%), Evidence (16,5%), Problem (16,5%) and Solution (14%). The remaining 33% is highly divided into other headings such as "critical capacity", "cooperation", "didactics", "gamification", "methodology", "teachers", "active" and "reflection".

Figure 2 shows a similar trend in the frequency distribution of the tags, regardless of the folksonomy, although the total is higher in narrow folksonomy (528). Among the most frequent labels we find "solution" (84) and "competence" (79) in the case of broad folksonomy, the latter being slightly higher in narrow folksonomy (155).
There is a difference in the total number of responses according to the folksonomy, 317 in *broad folksonomy* and 528 in *narrow folksonomy*, this difference being greater in the video than in the text. Allegedly this occurs because when the task instructions are to apply the four tags established by the teacher, students are more committed than when they can supply four answers without prefixed tags (Figure 2). The possible reasons to explain why there are differences in the video but not in the text may be that when viewing the video, viewers are somewhat more "caught up" in the storyline, and their attention is possibly more dispersed and/or less concentrated. However, this is not the case when previous instructions are mentioned in the video, i.e. students were more "committed" to the task in the video when they had this instruction than when they did not.

As for the quality of the tags (according to the 0-3 rating), there are significant differences in the mean values of best and worst quality, respectively (tags "Solution" X= 1.45 and "Others" X=0.42), with no significantly different quality values in the rest (Table 2).

<table>
<thead>
<tr>
<th>Tags</th>
<th>Mean</th>
<th>N</th>
<th>Standard Deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>1.27</td>
<td>135</td>
<td>.973</td>
<td>1.00</td>
</tr>
<tr>
<td>Solution</td>
<td>1.45</td>
<td>214</td>
<td>.766</td>
<td>1.00</td>
</tr>
<tr>
<td>Evidence</td>
<td>1.10</td>
<td>156</td>
<td>.844</td>
<td>1.00</td>
</tr>
<tr>
<td>Competence</td>
<td>1.22</td>
<td>234</td>
<td>.731</td>
<td>1.00</td>
</tr>
<tr>
<td>Good practices</td>
<td>1.20</td>
<td>41</td>
<td>.679</td>
<td>1.00</td>
</tr>
<tr>
<td>Others</td>
<td>.42</td>
<td>65</td>
<td>.583</td>
<td>.00</td>
</tr>
<tr>
<td>Total</td>
<td>1.20</td>
<td>845</td>
<td>.831</td>
<td>1.00</td>
</tr>
</tbody>
</table>

The Kruskal–Wallis statistical test shows no differences in the tags "Problem", "Evidence", "Competence" and "Good Practice". On the other hand, there were differences in "Solution" and "Others". The tag "Solution" is the one with the highest frequency of annotations (Table 3) and the tag "Others" does not respond to any of the questions posed, as comments or opinions mostly score a 0 rating.
### Table 3
Differences in the answers according to the quality of the tags "Solution" and "Other" as per the Kruskal-Wallis test

<table>
<thead>
<tr>
<th>Type of tags</th>
<th>N</th>
<th>N Average</th>
<th>Range</th>
<th>Quality of Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem</td>
<td>135</td>
<td>436.67</td>
<td>Chi square</td>
<td>85.696</td>
</tr>
<tr>
<td>Solution</td>
<td>214</td>
<td>495.39</td>
<td>df</td>
<td>5</td>
</tr>
<tr>
<td>Evidence</td>
<td>156</td>
<td>396.44</td>
<td>Asymptotic</td>
<td>.000</td>
</tr>
<tr>
<td>Competence</td>
<td>234</td>
<td>428.29</td>
<td>sig.</td>
<td></td>
</tr>
<tr>
<td>Good practices</td>
<td>41</td>
<td>421.12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Others</td>
<td>65</td>
<td>202.16</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>566</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 3.2 Differences According to Folksonomy

The results do not show significant differences in the quality of responses according to the code in the case of broad folksonomy, but there are significant differences in narrow folksonomy (Table 4) regardless of the code used.

### Table 4
Analysis according to broad vs. narrow folksonomy

<table>
<thead>
<tr>
<th>Folksonomy</th>
<th>Quality of the Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad</td>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon W</td>
</tr>
<tr>
<td></td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td>(Bilateral) asymptotic sig.</td>
</tr>
<tr>
<td>Narrow</td>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td></td>
<td>Wilcoxon W</td>
</tr>
<tr>
<td></td>
<td>Z</td>
</tr>
<tr>
<td></td>
<td>(Bilateral) asymptotic sig.</td>
</tr>
</tbody>
</table>

In contrast, there are no significant differences in the quality of student responses when the folksonomy variable is used, regardless of the code, text or video (Table 5).

### Table 5
Differences according to folksonomy

<table>
<thead>
<tr>
<th>Quality of the Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
</tr>
<tr>
<td>Wilcoxon W</td>
</tr>
<tr>
<td>Z</td>
</tr>
<tr>
<td>(Bilateral) asymptotic sig.</td>
</tr>
</tbody>
</table>

There is not enough evidence to state that there are significant differences in terms of folksonomy in the video and in the text separately and together. However, the quality of the answers is greater overall when the folksonomy is narrow. As shown in Figure 3, the quality of MAs in the text is higher in narrow folksonomy (left graphic red line ATE) and the same happens in the video (right graphic red line AVE).
3.3 Differences According to the Code

The results show that there are significant differences in the quality of the answers depending on the code used (text vs. video), the quality of the annotations in the text being higher than in the video, regardless of the folksonomy. In both cases the median is the same, i.e. 1 (Table 6). This may be because the text message is more structured and responds to the issues raised in an educational innovation project (tags and questions provided by teachers). While some students had the textual information that they could potentially use when analysing the video, this did not happen in the majority of the cases, Therefore, it cannot be said that prior text instructions had an influence on the result.

In the case of narrow folksonomy, the use of text obtains answers with a better rating, with a median in both cases of 1 (Table 6).

<table>
<thead>
<tr>
<th>Folksongomy</th>
<th>Format</th>
<th>Mean</th>
<th>N</th>
<th>Standard deviation</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad Text</td>
<td>1.23</td>
<td>74</td>
<td>.930</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>1.23</td>
<td>243</td>
<td>.816</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.23</td>
<td>317</td>
<td>.842</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Narrow Text</td>
<td>1.43</td>
<td>67</td>
<td>.701</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Video</td>
<td>1.15</td>
<td>461</td>
<td>.835</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>1.19</td>
<td>528</td>
<td>.824</td>
<td>1.00</td>
<td></td>
</tr>
</tbody>
</table>

3.4 Quality and Quantity of the MAs

In general, it can be seen the commitment when creating annotations is with the teacher or the task, but not with learning; that is to say, the quality of the answers given as a whole and by separate formats is different mainly due to the fact that the text code is more structured and has a prefixed tag. It can therefore be stated that narrow folksonomy and the text commit and focus the analysis in terms of number and quality of responses.
According to the results obtained, there is a greater quality response in text annotations in both folksonomies, which confirms the previous result that videos "use up" more concentration and it is more difficult to maintain an own and parallel thought, which is therefore more "dispersed". Thus, once the students have watched the video, they make their annotations in those images or sequences that produced an impact, and that were altogether fewer and of lower quality than those required by narrow folksonomy.

Likewise, and reinforcing the above, it is interesting to observe that the text annotations made by the students without given tags contain more words and more explanation as they are, in some cases, backed up with quotations and references to reinforce their reasoning. The total number of words in the four subgroups is: AVA 7828 AVE 11577 ATA 2540 ATE 9819. The chi-square indicates that there is also a 99% probability that there is an association between the variables.

The MAs show that the type of response with the tag "Solution" responds fundamentally to the purpose that the students thought the innovative educational project had as a whole, i.e. an effective, appropriate and original solution for the use of technologies, with comments such as "it is innovative because it breaks with the structure and spaces of the traditional classroom" (#1281) Others found it innovative from the curricular point of view: “This is a multidisciplinary project, since it includes four subjects of the curriculum” (#1293).

On the other hand, the annotations on Problems are surprising because of the low number of answers (only 6 in wide folksonomy), as if the students were less concerned about the problems of using technologies in the implementation of an educational innovation project. The issue of the reasons for innovation and technology only seldom appears in the students’ notes: “It is often very difficult to work with so many students at once and to get everyone’s attention and get them involved. That is why I think this project is innovative, as making 51 students find it attractive and eager to get involved, is rather difficult” (#1314).

As for the quality of the answers rated from 0 to 3, the level does not exceed the average in all the tags. Few values reach the highest rating of 3, the reason being that they usually describe the fact without interpreting it, they do not explain what skill is being developed or why it is innovative, for instance: “In this image we can see how the children are creating an urban garden. That is, in addition to learning mathematical knowledge, language, etc. They are also learning through environmental pedagogy” (#6556). In other instances, the descriptions are general: “It is a very innovative method, as it is a group research project where all the students of an ESO (secondary school) course participate, and it is all done through an online game” (#6480). In a way, they explain the impact of ICT only by describing how the participants use it in the project, but very few MAs provide details of the skills that can be developed or the possibilities they actually have.

4 DISCUSSION AND CONCLUSIONS

The analysis of the students’ MAs when they reason collaboratively on an innovative educational project allows discovery of if the type of mode (text vs video) and folksonomy (broad vs narrow) influences the quality of their responses. It has been demonstrated that there
are significant differences in the quality and quantity of the responses, with the quality of the annotations in texts being somewhat higher, regardless of the folksonomy. These differences are mainly due to the use of texts and of narrow folksonomy (tags that the annotations need to follow), which produces a greater number and better quality of responses. Similar results have been found in other studies where texts required less time and effort than videos (Hefter & Berthold, 2020). There are also other studies that point to advantages of texts over videos (Lee & List, 2019; List & Ballenger, 2019). However, generating and creating video messages is more motivating than using texts for tasks such as the creation of diaries by the students (Debbag & Fidan, 2020). Therefore, further research is needed on the functionalities of each code when using MAs.

The differences found in this study may be due, firstly, to the fact that the text code contains a more structured presentation of the elements of an innovation project, so the answer can be sought according to the tags, with a greater number of words and reasoning. Second, narrow folksonomy means that students are more likely to respond according to the tags supplied. A similar process as to when a question cannot be left answered, students fill in all the tags because they are there to be filled. Thirdly, the video message is more exciting and evocative about the program, focusing on the innovative solution and the narrative of the video itself, factors that lead to a greater dispersion and subjectivity. This is why the tag "Solution" receives more answers, followed by "Competence". In short, and in line with the idea analysed by Becker (2010) on the combination of methodology and technology, in folksonomy, the audio-visual narrative of the video message, which is more open and subjective, is combined in folksonomy with an open assignment and design. The more structured text format with closed task assignment guides (narrow folksonomy) yields a higher number of annotations. Possibly, for different, more creative skills, broad folksonomy and video annotations might be more beneficial. In any case, this study has found that students have a greater, more focused and motivated commitment to the task when they share their interpretations, annotations and tags with others, similar to the findings by (Qarabash, Heslop, Kharrufa, Balaam, & Devlin, 2019). Also, collaborative tagging activities with controlled vocabulary in the annotation are effective for the personalization of learning resources (Lau et al., 2015). However, although important steps have been taken in technological development and customization of learning (Hsu, 2013), more software development research is still needed to provide more personalized teaching and a closer relationship between user folksonomy and technological tagging.

Students show a willingness to actively collaborate and get involved in tasks requiring reflection about the video’s messages by analysing and writing shared annotations about the videos and texts. As in other research, the attention of students in the analysis of the messages increases by counteracting the speed of the message and its discourse (Po-Sheng et al., 2018), thus facilitating a more reflective and reasoned reading of the practical and innovative contents. Nonetheless, two important aspects need to be highlighted: In terms of methodology, teachers should follow up on the task, because at the beginning, as has happened in other studies (McFadden et al., 2014), students often do not make comments or ask questions to other students and only focus on their own annotations, which are lower-level
Multimedia Annotations for Practical Collaborative Reasoning

Manuel, Cebrián-de-la-Serna; et al.

reflections (descriptions and explanations) but not in-depth analysis. As for technology, in some cases and as in other studies (Gao, 2013), it is important to provide students with a simple and intuitive interface so they can manage the huge amount of annotations that are generated in the tasks, as well as having data visualization tools (statistical graphs, word clouds...) that allow an overview or a selection of a specific topic within this set of annotations (selective search tool for words or tags). This will be an important factor in the selection of tools in the future, hence, it is worthwhile proposing new comparative studies of the possibilities currently offered by MA systems from a technological point of view, together with new experiments and evaluations of their functionalities (Big data, AI, Machine Learning...), as well as review studies on the different methodologies and research results of the studies in diverse and INCLUSIVE contexts.

The text is more structured and offers more information about impact and innovation, while the video only presents images of student actions, without voiceover or text tags that reinforce the images. Therefore, we cannot generalize the results to other contexts, and investigations on more exciting and less structured texts vs. more structured videos and more explanatory video messages should be carried out. More specifically, it would be interesting to add to the video a soundtrack with the same wording as the text and minimize the text exactly to the size of the video transcript. This would mean two messages with the same text with an audio-visual aid. In any case, more research should be done on how both codes can complement each other in online programs. This is in line the work by García-Martínez, Rigo-Carratalá & Jiménez who seek to improve the reading process with multimedia resources. The results show minimal differences in the final evaluation and the authors state “...a Multimedia methodology to develop the textual comprehension, properly designed, helps the educational process to be more effective because students make fewer mistakes when answering questions. Furthermore, we have seen that at the time of the evaluation, the introduction of a multimedia methodology compared to a textual presentation does not improve (at least in absolute terms), the student's ability to succeed in their responses.” (2017, p. 9). One option is a combined methodological design of the folksonomy variables and multimedia codes depending to the relevance, similar to the study by Dennen, Bagdy, and Cates (2018) which examines student tagging activity within a five-week social bookmarking unit, considering that "these skills [tagging] are important components of information literacy and are used increasingly in professional settings where large quantities of information are being amassed, evaluated, and shared" (2018, p. 117).

In future studies it will be necessary to experiment with the use of methodological design in different tasks and skills. It would also be advisable to study the relationship with internal vs. external motivation variables (e.g. final grades) and their use in broader contexts such as a complete program and different grades (Mirriahi, Joksimović, Gašević, & Dawson, 2018; Mirriahi, Liaqat, Dawson, & Gašević, 2016). The very nature of the competence acquisition process demands it, because there is a profusion of technological innovations and multimedia messages in our social, family and professional life, and students need to consider these messages through a more critical and professional vision. It essential that university training is backed by documentation and multimedia resources of professional practices,
and to create methodologies and activities that allow a greater theoretical-practical relationship with the aim of obtaining a deeper, more critical and reflective effort in the students’ responses. In the current situation, greater flexibility is also needed in the university system in its teaching methods, adopting the methodologies and technologies used in the university environment by configuring an entire multimedia PLE-portfolio (M. Cebrián, 2017).

ACKNOWLEDGEMENTS

Funded by: Ministry of Science and Innovation, Spain
Funder Identifier: http://dx.doi.org/10.13039/501100004837
Award: EDU2013-41974-P

REFERENCES


Smith, B. K., Blankinship, E., & Lackner, T. (2000). Annotation and education. *IEEE Multimedia, 7*(2), 84–89. [https://doi.org/10.1109/93.848437](https://doi.org/10.1109/93.848437)


Sydnor, J. (2016). Using Video to Enhance Reflective Practice: Student Teachers’ Dialogic Examination of Their Own Teaching. *The New Educator, 12*(1), 67–84. [https://doi.org/10.1080/1547688x.2015.1113346](https://doi.org/10.1080/1547688x.2015.1113346)