Participatory Media Literacy in Collaborative Video Annotation

Craig D. Howard¹

Accepted: 2 July 2021 / Published online: 14 July 2021
© Association for Educational Communications & Technology 2021

Abstract

Collaborative Video Annotation (CVA) is a kludge where learners annotate video together, experiencing both the video and each other's annotations in a dynamic discussion. Three scenes from small group CVA discussions were selected for analysis from 14 CVA discussions where 8–12 learners interacted via the annotation tool on top of a video. The twenty-second scenes were analyzed for semiotic meaning-making practices and this revealed a variety of participatory media literacy levels among these undergraduates. Topics of discussions were related but not identical, and the selected exemplars showed a range of attention to communicative features of the media. Discussions evolved in dramatically different ways due to the interplay of images, text, and learner choices. Results suggest that converged media require new literacies educators would be wise to explore and wiser still to educate our learners about.

Keywords Participatory media · Video annotation · Media literacy · Semiotic analysis · Preservice teacher education · Computer supported collaborative learning

Introduction

The spring of 2020 accelerated the expansion of teaching and learning mediated by technology, especially in higher education. The trend was already underway. More and more media progressively surround learners, and the percentage of that media that is participatory is ever increasing. Instructors are encouraged to leverage participatory media in their teaching. While traditional media literacy perspectives viewed learners as consumers of media in need of ways to navigate a glut of information overload and adopt critical mindsets (Potter, 2004), new media literacy initiatives target participatory media literacy. This shift in focus aims to enable learners to produce, create, curate, and engage in self-expression in emergent media (Butler et al., 2018). Media literacy scholars have argued developing these skills strengthen social capital (Lantela, 2019), and can allow access to a broader community

(Baleria, 2019). The value of digital literacy in participatory spaces has been established, but precisely how these media function in the learning process has not.

Educators grapple with participatory media and how to best use them for teaching and learning. Many of our designs in online learning attempt to incorporate critical uses of multimedia or visually enhanced experiences. Baylen and D'alba (2015) highlights the participatory aspects of literacy in these media and their relative proliferation, “The rapid ascension of visual technologies as tools for delivering content, facilitating communication, developing critical thinking, and expressing creativity has become a game changer in *encoding and decoding images*. [emphasis added]” (pp. xiii). Learners’ encoding of meaning takes place in digital spaces which ~~that~~ converge media affordances in a myriad of combinations (Herring, 2013). Progressively, these converged spaces contextualize our teaching (Howard, 2012). While there have been calls for more studies of learning in participatory media (Baleria, 2019; Flynn & Lewis, 2015), and entire formal curricula have targeted participatory media literacies (Frohlich & Magolis, 2020), educators may not fully understand the entire range of skills learners may need to adequately participate. That understanding can only develop with a close inspection of how participation unfolds. In this study, learners’ interactions are studied in a specific type of converged and participatory social media, Collaborative Video Annotation (CVA).

Submitted to the special issue [SI: VML] on visual and media literacy in global contexts, Eds. D. Baylen and B. Hokanson

✉ Craig D. Howard
cdh@utk.edu

¹ The University of Tennessee Knoxville, Bailey Education Complex, 1122 Volunteer Blvd., Knoxville, TN 37996, USA

CVA is a kludge; A kludge is a device which has been slightly altered in order to be repurposed for a new use. The instructor repurposed the annotation tool within YouTube so it could be used for asynchronous collaboration among learners. I configured the annotation feature on the video-sharing website (Howard, 2019) to scale up a pre-service teacher observation activity. This repurposing of the tool allowed large numbers of learners to view, and asynchronously discuss, a single example of a teacher teaching. The annotation experience afforded synchronous-like interaction via asynchronous communication where users see their classmates' annotations appear and disappear as if they were interacting in real-time, but in fact, playback must be stopped to participate and (Howard & Myers, 2010). CVA interaction is completely asynchronous. Messages remain on the video timeline for the next student to experience along with the video.

Collaborative video annotation (CVA) offers advantages over in-person interaction, but the experience is time-consuming to create. Asynchronous configurations democratize interaction, allowing all learners the time they need to compose, re-watch and craft their communications. CVA accommodates distance learners and multiple schedules, and affords all of the advantages of asynchronous discussions while still offering visual observations that can provide common ground for learners to discuss what they see. At the same time, the experience is also labor-intensive for an instructor to design. Video must be shot, reviewed, spliced, and curated. Access and grouping of learners in the video-sharing platform must be coordinated, and participation monitored. For this reason, video annotation interaction should be closely studied so its potentials, and potential pitfalls, might be understood. A close analysis of how interaction takes place in CVA offers those who might use annotation as participatory learning media the opportunity to make the most of this pedagogical strategy. In this study, a semiotic analytical procedure exposed how small groups of learners navigated an influx of new affordances in CVA. Exemplars drawn from authentic interactions show how successfully, or unsuccessfully, learners participated in the asynchronous discussions.

Previous Literature

Situating this study among others of media literacy must consider the different foci in the fields' body of research. This study fits into a learner-centered media literacy perspective but does not address the themes of violence, sexuality, health, or the recognition of stereotypes which learners face in mass media that are typically studied. Rather, this study belongs to the body of research that aims to empower learners via a critical perspective of their media participation, as opposed to inoculating learners or a transmission-oriented model of education (Hobbs, 2011). This body of scholarship assumes

a constructivist perspective of learning. Constructivist perspectives embrace a critical lens on participatory media literacy (Kersch & Lesley, 2019; Wright, 2020). The trajectory of these constructivist studies traces back to early studies of new media and multi-literacies explored by the New London Group (1996), and later by Jenkins (2006). These scholars advocated investigations of participatory media that "provide more opportunities for deliberation, discussion, sharing, equity, and participation" (Tugtekin & Koc, 2020, p. 2). Advocates of media literacy aspire to enable learners to successfully navigate and participate in online spaces.

Media literacy educators advocate for critical pedagogies in participatory media. The National Association of Media Literacy Education (NAMLE) asserts that, "the purpose of media literacy education is to help individuals of all ages develop the habits of inquiry and skills of expression that they need to be critical thinkers, effective communicators and active citizens in today's world" (2007, p. 1). Critical pedagogies in participatory media also fit within the NAMLE, 2020a, 2020b definition of media literacy which advocates for pedagogies to support the analysis, evaluation, and creative acts in all communicative media. The critical perspective frames the visually literate individual as a person who brings a critical view to media including user interaction via text, video or audio. Individuals literate in these converged media can competently take part in the development and sharing of knowledge. However, while teacher educators have argued for a framework that includes critical perspectives on media (Kersch & Lesley, 2019) and encourages learners to tell their stories and think critically (Baleria, 2019), platforms' diversity can make the actual practices elusive to guide and support.

Platforms that use video are many and varied. Only through study of participation in isolated designs can we pick apart which features support media literacy and which do not. Scholarship in participatory media literacy should look closer at interaction within media that join video and text to understand how the combination functions in the learning process. Solmaz (2017), in a study of several different participatory media, makes the case that the diversity of affordances among participatory spaces and their associated skills are so great, that each may require different skills to master, and each must be studied individually to be understood. Solmaz (2017) found that among a battery of eleven participatory media configurations, *negotiating* emerged as the most useful, important, and common literacy skill. Solmaz (2017) defined negotiation as interaction that "requires going beyond information dissemination and involves traveling across different communities in one's network and respecting alternative norms" (p. 55). In the context of media that join video and text interactions at the same time, negotiations include navigating the interplay between the two modalities.

Navigating the interplay among images and text requires a new literacy about how the two function together. To

participate effectively, learners must understand how combinations of images and text create meaning. This relationship is not new to studies in media literacy. Sansone (2015) calls these combinations picture-texts and explains, “When the picture-word relationship is used effectively, the pictures reveal aspects of the concept that the text is incapable of explaining, and vice versa” (p. 10). Encoding and decoding picture-texts, or image-texts, is a type of literacy also studied by semioticians. The work of Kress and Van Leeuwen (1996), Kress, (2009), and later work by Mayer (2010), recognized navigating image-texts as a semiotic skill—an ability to signal certain kinds of meanings. In a media landscape that employs these combinations routinely, these image-texts provide exceptional opportunities to explore how to express one’s point, but also opportunities to fail. Digital tools introduce limitless images-text possibilities into learning practices, but the relationships between text and image can be intentional, or unintentional. The relationships between image and text can evidence an ability, or an inability, to navigate the digital world (Hull & Nelson, 2005). Video annotation in online spaces can take this skill set one step further by introducing collaboration into the process. Image-texts are collaboratively created in collaborative video annotation (CVA).

Collaboratively created image-texts in CVA include several learner choices and are highly complex because, unlike static image-texts, new meanings develop with each participant. As one annotates, the resulting image-text changes for the next participant because of the additional commentary and communications which are embedded in the video screen. I have termed this phenomenon a *dynamic imagetext* (Howard, 2012; Howard, 2019). Dynamic imagetexts introduce two specific types of choices for learners—lexical and technical. Learners must choose the written expressions they contribute; what text goes in the box is a lexical choice. Learners must also make technical choices—where in the video timeline to place an annotation, which content from the video or commentary to address, how long to set the duration on screen, and where within the screen real estate to place the annotation box. Both sets of choices combine to create meaning. The combination can dramatically impact the experience of the next entrant into the video annotation space. However, educational researchers have rarely analyzed how video annotation plays out in learning interventions.

The body of literature surrounding video annotation in learning interventions gravitates to preservice teacher education, but analysis of video annotation has also appeared in studies of video viewing behavior. In preservice teacher education, the history of using video annotation to support reflection dates back to when the technology first appeared. Rosaen et al. (2008) and Van Es (2010) documented early investigations of video annotation supporting preservice teacher reflections through guided noticing. Scholars such as Fadde and Sullivan (2013) built on such studies to develop video

annotation for reflective opportunities supported by expert-written annotations on the same video clip. The most widely cited article, (Rich & Hannafin, 2009), provides many detailed discussions about different uses of video annotation in teaching and learning. However, more recently, studies such as Zaier et al. (2020) and Cattaneo et al., 2020 continue to argue for video annotation’s ability to facilitate peer evaluation, academic feedback, and self-evaluation. In behavior analysis, Mirriahi et al. (2016) investigated cluster diagrams of video annotation behavior and determined that viewers reflected one of four behavior types: minimalist, disenchanted, task-focused, or intensive group clusters. Pérez-Torregrosa et al. (2017) reviewed the literature on video annotation and noted that prior to 2017, only 19 studies had been conducted on video annotation in education; they provide an overview of the increase in studies of video annotation in learning interventions from 2006 to 2016. In neither the studies reviewed by Pérez-Torregrosa et al. (2017) nor any other the others mentioned here, has video annotation been interrogated from a semiotic lens. How dynamic imagetexts challenge learners to either learn or remain on the outside of our new media cultures is yet to be explored.

Purpose of the Study

The purpose of this study was to investigate the nuances of learner participation in CVA. A close investigation was conducted to understand the interactive dynamic of CVA as visually-enhanced participatory media. The close investigation required untangling many of the media and expressive choices learners may have made unconsciously. Untangling semiotic choices can reveal the affordances of the media, as well as media affordances which must be taught explicitly so the media can become useful for teaching and learning. Only a close inspection can disclose how affordances function to enable, or inhibit, discussion because learners’ understandings of participatory media are not uniform.

A semiotic approach offers advantages over other types of analysis. Quantitative analyses might confuse reaction to specific video content with learners’ familiarity with the tool. Semiotic analysis of entire scenes among different groups of learners can allow the researcher to appreciate the relative impact of the media affordances among different learners. Semiotic analysis can also lay bare the range of learner ability in navigating the new discursive space. Close investigations of learners’ interactions in holistic scenes on the video timeline reveal variances in participation from a single discursive multimodal prompt. By limiting exemplars to only samples from identical video content areas, the video variable can be mitigated, and thus exemplars might disclose evidence of the range of learning possibilities.

The study was not intended to evaluate the CVA media or the learners. The study was also not intended to test a hypothesis or inform a learning theory. Rather, the purpose of the study was to better understand, through focused exploration, participatory literacy in the CVA design. A theoretical frame was selected that best supports developing a nuanced understanding of dynamic imagetexts in the learning process.

Theoretical Frame

Approaching CVA as the production of multimodal imagetexts assumes a semiotic lens. From the semiotic perspective, the researcher must appreciate the total communicative ensemble to understand the process and avoid misconceptions about the use of the media. This is particularly important for novel media such as CVA. As each new learner contributes to the discussion that takes place atop the video, the discussion evolves, and the experience of participation changes. Scholars have made the case that these multimodal texts contain meanings that transcend the contribution of constituent parts, “Multimodality can afford, not just a new way to make meaning, but a different kind of meaning” (Hull & Nelson, 2005, p. 226). Addressing entire annotated scenes as single entities, rather than images and texts in isolation, builds on the work of Kress and Van Leeuwen (1996), Kress (2009), and Jewitt and Henriksen (2017). CVA communications must be understood within the context in which they were created because when learners add text to images, the meaning changes. Essentially, each learner has experienced something slightly different because the CVA experience evolves with participation.

From a semiotic lens, isolated research questions are not comparable across samples due to the dynamic nature of the interaction. Therefore, Jewitt and Henriksen (2017) operationalized a semiotic analysis of multimodal imagetexts into three *related* questions to be considered together of a single communicative ensemble. I have repurposed that operationalization into research questions and followed each with explanations as to how the analytical process might apply to CVA analysis. I present these research questions with the disclaimer that they work in unison, and are not meant to be uniformly applied, but rather explorative in nature.

Research Questions

What are the Semiotic Choices the Interlocutor Made?

Semiotic choices are anything in the communication that appears as a signifier, an intentional move to create meaning. This would include lexical choices that invoke connotations of authority, co-membership, affinity, affiliations, or a host of other communicative strategies. This would also extend to

technological choice if that choice was intended to convey meaning. For example, if a learner chose red as an annotation color to express anger or disagreement with something in the video. Simultaneously, if a learner chose to interpret red in another annotation as signifying meaning, it would be a semiotic choice.

What is the Provenance of the Communications?

The provenance of a communication is a communication’s backstory. A backstory contextualizes the researcher’s understanding by taking into consideration how video action and previous annotations may have impacted communications. Provenance can be thought of as a holistic contextualization. The events, both linguistic and material, that contributed to the present state of affairs inform the analysis. For example, if one interlocutor has established a topic of the discussion, this can be calculated into the analysis of subsequent agreements or disagreements. What might have otherwise been seen as a completely open field of possible student contributions, due to provenance of the scene, may appear to the next student as a dichotomous choice.

What Technical Affordances were Chosen?

How technical modalities are employed can dramatically impact the efficacy of communications. In the case of collaborative video annotation, the selected location of annotations on the screen real estate, location on the video timeline, and duration of the annotation on the video screen are all choices that have dramatic impact beyond one’s own communications. They can support others’ annotations or obstruct others’ views. Location choices can be used advantageously, such as inserting in proximity to the subject of the annotation; this tactic facilitates reference. At the same time, the selected location impede communication. Annotations can obstruct another’s view of video content, or inadvertently obstruct the view of one’s own annotation by placing it in a location that will soon be occupied by another annotation already inserted into the video timeline. Learners make these choices intentionally or unintentionally. Understanding the performed facility of those choices, and how they interact with the lexical choices in question 1, informs our understanding of learners’ media literacy.

Scholars have advocated that communications in similar multimodal products be studied in these dimensions of semiotic choice, provenance, and technical affordance before claims are made about how communication comes together in these spaces (Kress & Van Leeuwen, 1996). Each of these three dimensions impacts the others; thus, the researcher must consider them together. Dynamic imagetexts are fluid and organic. Studies in this theoretical frame aim to put forward possibilities rather than likelihoods or generalizations about

communication. In the case of CVA, the media's uniqueness and dynamic nature are both profound, so the application of Jewitt and Henriksen (2017) operationalization of semiotic analysis is intended to explore, not verify. Semiotic choice, provenance, and media affordance were considered simultaneously to explicate evidence of possibilities. It is from this lens that three, twenty-second selections of learner interaction were analyzed. This theoretical frame qualifies the research question. The research question *How do these learners make meaning in this space?* is approached to expose potential semiotic and media practices, and allow the researcher to become aware of communicative pitfalls in this particular media-enhanced learning intervention. This study is not intended to define literacy in collaborative video annotation or any other media, but rather to explore how literacy might emerge, or fail to emerge, in a learning context.

Context of the Data

This study interrogates learner interactions in a converged space that redesigned an observation activity for preservice teachers, mostly fourth-year undergraduates. For a detailed discussion of the collaborative video annotation (CVA) intervention that generated the data used in this study, see Howard and Myers (2010). The design had previously assigned learners to attend disparate live sessions, and convene in face-to-face classes to discuss their teaching observations. There were a number of shortcomings this redesign hoped to resolve, such as avoiding dissimilar experiences providing little common ground for discussion, and a lack of time allowing every learner to contribute to live discussions. In the previous in-person learning activity, the instructor was unable to guide noticing of excellent teaching, or even track who had, or had not, completed the observation (Howard & Myers, 2010). This instantiation of the activity included 141 students in an undergraduate teacher education course about technology integration. The activity employed a converged video-sharing platform and associated textual overlay tool, YouTube and YouTube annotations. All groups' activity ran asynchronously for thirteen days in the spring semester of a school year prior to the COVID-19 pandemic. Figure 1 shows the base video image in the CVA discussion and the three selections of that same scene used for semiotic analysis. Groups A, B, and C depict the completed CVA discussions analyzed in this study.

In the collaborative activity, learners had been randomly grouped into one of fourteen groups of 8–12 learners to annotate the video. Learners had been assigned a reading a week prior to the task to support the discussion. Participants were emailed a task sheet that included directions on how to complete the activity and a link to their group's uploaded copy of the video. The asynchronous activity took place in the final weeks of the course, and learners had been told in the task assignment that participation of six annotations would garner

full credit, regardless of the accuracy of their annotation content. Learners were tasked with identifying and discussing observed excellence in teaching. All learners whose interactions appear in the present study had signed informed consent forms designed by the university IRB protocol officer. Students were also asked to identify themselves in the task sheet, but reporting in this study pseudonymized names to protect identities as dictated by the approved IRB guidance.

The video itself was 7:39 long (see Fig. 1), contained three minutes of directions on how to annotate video and how the video annotation platform functioned. Collaborative video annotation (CVA) allows annotations to overlap. The directions also explained how to select annotation color and screen duration. The tutorial intended to make learners aware that annotation content or video content, could be obscured. It is important to note here that the sequence of the annotation overlap is not chronological; rather, the most recently inserted annotation on the video timeline is foregrounded. Thus, one can obstruct the view of another annotation by inserting one's annotation atop annotations already visible on the screen. Similarly, one's annotations can be buried by others. This explanation was followed by approximately four- and one-half minutes of a veteran teacher teaching. The reading had posited that excellent teaching was effective, efficient, and engaging uses of technology in the act of teaching (Merrill, 2009).

In computer-mediated communication (CMC) research, best practices dictate the reader be offered a descriptive overview of participation in the socio-technical context of the space to orient themselves to the scope of the data (Herring, 2013). Therefore, descriptive statistics of the number of total annotations, total words, average posts per learner, average words per learner, and average words per message appear in Table 1. Table 1 reports the three selected groups' participation and compares their participation frequencies with the total averages from the fourteen groups. It is important to note group size because group size impacts the tenor and types of discussion which take place in any given space (Korenman & Wyatt, 1996; Howard, 2012).

Table 1 also orients the reader to the intensity of the participation in the space. While only six annotations were required to complete the task, none of the fourteen groups averaged less than seven. Compared to asynchronous discussions used in other media such as discussion forums, this is approximately three times the average amount of participation (Hara et al., 2000; Howard, 2012). Annotations appeared in clusters on the video timeline leading to the purposive sampling detailed below.

Data Selection

All fourteen annotated videos were scanned for a single video segment that might provide comparable exemplars of media literacy among learners. The researcher reasoned that

Fig. 1 The base video copied and uploaded for multiple video-annotated discussions, shown in A, B, and C instances of CVA at the same moment in the video

exemplars must contain participation in an identical video timeline section; otherwise, video content could make the collaborative video annotation (CVA) exemplars so dissimilar that analysis might be more tied to video content than learner behavior. Thus, the researcher searched for a singular selection of the video among these fourteen groups to control for video content. The purposive sample would include one scene where learners clustered annotations, one where annotations occupied the scene but did not overlap extensively, and one that would include only a small number of annotations. This purposive sampling was designed to expose the range of learner communicative behaviors under the different learning conditions in CVA.

Certain conditions also eliminated some group discussions from the study. Some groups had less than two annotations in

the 20-s scene. A single annotation does not provide enough discourse to explore the nature of CVA. Additionally, groups video annotation segments that contained interactions from a learner who had not signed the university-approved informed consent were also excluded from the study. In this process, half of the fourteen groups were immediately eliminated. In the remaining seven groups, three exemplars contained interaction on this 20-s scene.

A tool within the media afforded locating the scenes to be analyzed. YouTube's graphical depiction of annotation locations on the video timeline displayed high-activity points shared among groups. Using this tool, the researcher located active 20-s video segments. The process is visually represented in Fig. 2. A single activity point on all three videos led to selecting three scenes to analyze the discussion's real-time

Table 1 Participation frequencies intended to orient the reader to the communicative context, comparing selected groups to the total of all 14. * denotes averages used for totals to present a normalized comparison

Group	Group size	Total annotations	Words	Average annotations per learner	<i>sd</i>	Average words per learner	<i>sd</i>	Average words per annotation	<i>sd</i>
A	10	84	1957	8.4	1.8	195	23	23.3	12.5
B	9	79	2515	8.9	2.8	277	31	31.8	14.7
C	10	69	2119	7.0	1.7	212	30	30.3	19.9
ALL 14	*10.7 sd 2.6	*82 sd 8.42	*2524 sd 372	8.7	2.9	239	43	28.5	18.9

Fig. 2 Graphical depiction of annotations' presence on the video timeline for Group B. The selected area shows high participation locations in the video at approximately at the 3:10 mark. Notice not all annotations appear on the video timeline for the same length of time, denoted by a speech bubble visible inside an annotation identifier

development. In Fig. 2, a yellow outline shows where the researcher focused in the video timeline for the three groups. By looking at multiple groups' participation through the tool, with different videos opened simultaneously in multiple browser windows, the researcher could scan the video timelines to select appropriate samples. Graphical speech bubbles appearing inside the annotation identifiers identified annotations that had been extended on the video timeline, an indicator of the presence of intentional technological affordance choices.

Analytical Procedures

The researcher re-watched each annotation scene several times for each of the three groups. In this 20-s clip, there is talk and movement taking place in the video. The audio contained an interaction between the teacher and a student offscreen. That interchange went as follows:

Teacher: If you connect a TED video to your classroom, it would totally act as a resource. So you could use it in your classroom, of course, but are there any other ideas about how you might want to add it to a lesson? Why would I do this?

Offscreen student reply: It's a way of getting information apart from watching the news or something.

Teacher: Ideally, you are intrinsically motivated to learn something on your own when interacting with more high definition materials like a TED video.

Learners' annotations had previously been timestamped (Howard, 2012). This enabled the researcher to map the sequence of annotations' appearance in real-time, versus their appearance on the video timeline to recreate the experience of each learner entering the space. Once the researcher had developed an understanding of the video segment's chronological

development, the three questions operationalized by Jewitt and Henriksen (2017) were applied to the three scenes.

Results and Discussion

Overview of the Analysis

All three cases contained slightly different discussion topics despite identical locations on the video timeline. While each discussion was tied to video content, none followed the same trajectory of development. The purposive sampling technique garnered scenes with different amounts of participation. Four annotations constituted A, six in Group B, but only two made up Group C. Learners' ability to manipulate the space for their own learning varied with participation in each scene. These scenarios together suggest that for these learners, the busier the screen real estate, the more challenging the media becomes. Semiotic choices were not limited to a single communicative strategy. For example, not all annotations adhered to the discussion topic— a non sequitur appears in Group A. Learner agency was prominent in all three samples; The three exemplars' interaction suggests that learner contributions were as influential in generating engaging on-task discussion as the video itself. The following three sections report my analysis using Jewitt and Henriksen (2017) operationalization; the three questions asked in unison provided a nuanced description of the development of the three dynamic imagetexts.

Group A

Group A included four annotations, three of which discussed how the teacher leads via her questioning tactics and body positioning. One annotation addressed the media itself and did not continue the topic of the others. All the annotations were positive in tenor. All learners chose different colors for their annotations. Only the final annotation references the affordances of the media. The imagetext developed into a

discussion of teacher behavior with one outlier. Figure 3 provides a screen capture of the interaction to orient the reader to the analysis that follows.

These learners made lexical choices that reveal the perspectives they brought to annotating the video. Learners in Group A, as seen in Fig. 3 and Table 2, demonstrated that they viewed a discussion of excellence or engagement as, in this case, entailing a discussion of respectful teacher behavior. Phrases such as “doesn’t shut down,” “good proximity,” and “asking students about their thoughts,” all refer to the behavior of the teacher in the video. Three of the four annotation writers selected this type of classroom conduct in identifying excellent teaching. The task was open-ended; it had simply asked the learners to identify excellent teaching. These lexical choices evidence the perspectives they brought to how they interpreted the task. There was nothing in the task that asked them to talk about body language or teacher questioning, or even the Socratic method, but all were eluded to in the annotations. For these learners, the teacher’s ability to show respect in the act of leading a class discussion was the excellence in teaching that they chose to discuss. The audio track may have played a role in this topic choice because content from the audio track is eluded to in the discussion.

The first annotation in the scene makes a loose reference to a student’s off-screen voice in the video that impacts the ensuing discussion. A curious interpretation of a student comment sparked discussion. In the audio, a student comments that, *It’s a way of getting information apart from watching*

the news or something. The first learner annotation evidences an assumption that the teacher could have assessed the comment as indicative of the student in the video not paying close attention. The annotation states, *It’s great that the teacher doesn’t completely shut the students down, she rephrases the question to get the students to arrive at the answer she wants*. This observation was drawn in the first annotation and may have inspired these learners to interrogate how the teachers’ response evidenced respectful teaching in light of an errant learner answer.

While there are no direct mentions of the off-screen learner response to the teachers’ questions, the loose connection between the questioning sequence and the learner’s commentary shows that the audio played a role in how learners interpreted the video and the task. The off-screen comment evoked a discussion of respect as shown through teacher questions and body positioning choices. The grey annotation writer infers that this response was not what the teacher was looking for and thus launches the topic. The topical thread continues for three messages culminating in a link between teachers’ question asking and pedagogical effectiveness. Table 2 provides the annotations in chronological order so that the reader can appreciate the progression of the discussion. Chronological order cannot be recognized in screen captures; the timestamp tool allowed for chronological analysis.

The green annotation is an outlier. This learner has selected a different topic entirely. This annotation was inserted eight days after the others as evidenced in Table 2, appearing on

Fig. 3 Group A collaborative video annotation, showing that a student accidentally buried her annotation under another’s annotation

Table 2 Group A's CVA interaction in chronological order, showing the topical connections among annotations, and a divergent topic in the final annotation

Color / Date	Annotation text
Grey Annotation [4/12 19:25]	It's great that the teacher doesn't completely shut the students down, she rephrases the question to get the students to arrive at the answer she wants
Red Annotation [4/12 20:35]	I like that she's asking the students why they think they are doing this, and I like how she is walking around. She has good proximity to the students and keeps moving.
Light Blue Annotation [4/12 21:25]	I definitely agree! asking students questions regarding their thought is the best way to probe into student thinking and create for a very effective lesson
Green Annotation [4/20 18:40]	I think this is a great idea! I never thought of making it [video annotation] anonymous...

April 20th. This is the learner's second time to enter the space and annotate the video, and this might explain why the learner has selected not to engage in discussion about the teacher in the video's behavior. The green annotation writer may have assumed this topic was finalized. We know that she has been in the space before while these annotations were already on the video timeline because she has annotated elsewhere during that timeframe. Therefore, in her first pass she may well have seen these annotations. Seeing no new annotations in her second pass through, she may have decided a new topic was appropriate. The asynchronous nature of the activity may have lessened her desire to interact with the others' selected topic because it afforded her the freedom of a new choice. This learner introduces a new topic and extends the time her annotation is on screen. In so doing, her annotation gets buried under an annotation appearing later on the video timeline. She has extended beyond the default five-second setting, but she does not consider if another annotation is already in place on the video timeline in the location that will soon occupy that screen real estate. The video annotation tool foregrounds the video timeline. This avoids a single annotation occupying all of the screen real estate. In adding her annotations as she did, her annotation is only partially visible.

Not only has the author of the green annotation chosen a location that would later become illegible, but her text evidences a misinterpretation of the functionality of the media as well. She celebrates that the media was anonymous even though it was not. User pseudonyms were visible via cursor roll-over. This analysis demonstrates that the Group A annotations show that learners cognitively engaged in on-task discussion, but in one case, did so with misunderstandings of the media.

The interaction in Group B began with a teacher-inserted annotation attempting to steer the discussion into an analysis of body language and teaching. (See Table 3 for a chronological transcript, and Fig. 4 for a visual of the Group B interaction.) Thereafter, the Group B discussion grew organically into a discussion of how computer labs are configured in schools and how teachers might teach well in those labs. Five separate annotations appeared on the screen during

this 20-s segment of the video, making the scene busy. In this group, most communications revolved around the layout of the lab, arising through learners addressing relationships between the instructor in the video's body positioning and the task's request that learners identify excellent teaching. The imagetext became a composition of overlapping analyses of the video. Figure 4 shows the scene's obscured view, suggesting that space limitations and overlap may have played into the discussion.

Analysis revealed that annotation placement was not out of disregard; but somehow, either through media affordance, design failure, or learner assumptions, the discussion took an erroneous path. Learner's contributions demonstrated how difficult communication can be in such an interactive space. The student who enters an annotation earliest in the video timeline soon has his annotation covered up by others. This results in an image-text where others can see his name, but not his comment. This dynamic suggests that the annotation that covered an earlier one has followed up on the same topic. Topic choices within annotations evidence that learners read the annotations that were visible on the screen when they annotated. As in the case of the red annotation, annotations also evidence that this behavior of stopping the video, reading and annotating, can place an annotation into the video timeline just before another annotation will occupy that screen real estate. Students can effectively obstruct the view of their annotations if they have not checked to see what annotations are coming. In Group B, learners' annotations often obscured each other. The obscured annotations in Fig. 4 are written out in full in Table 3, and depict how the discussion follows one topic through multiple obscured annotations.

Despite an instructor attempting to steer the discussions onto a fruitful trajectory, the combination of technological choices and misconceptions drew the discussion elsewhere. The author of the light blue annotation launches the direction of the subsequent discussion by the lexical choice, "set up." (This annotation is not visible in Fig. 4; please see the second annotation in Table 3.) The choice of this term *set up* introduced the computer lab layout as a point of discussion, and with it, a notion that the instructor has *chosen* the layout of the

Table 3 Group B's transcript of overlapping annotations, depicting the organic growth of a CVA discussion

Color / Date	Annotation text
Grey (overlaid, this is a teacher annotation) [4/10 9:06]	Good body language. Yeah, the student in the front seems to be reacting to the body language. Is that effectiveness?
Light Blue [4/12/2010 13:00]	Jami Here. I don't know that there is an effective way to set up a computer lab so that all students are visible and so they are all engaged. Even if they were in rows, the students in the back would not be visible. However, I think that the way the classroom is set up now, that the teacher really does have good [sic] visibility of all students. She only has to walk back and forth to see down the aisles. I think the way it is set up now is the most effective, and allows for efficient teacher movement.
White (above) [4/12 21:30]	I agree with Jami. The way she has the rows set up now are probably the most effective way she could do it. As a teacher you would need to remember to walk down the aisles constantly to make sure students are on task. (Rachel)
Red (left, overlaid) [4/12 22:30]	I think that as a teacher teaching in a computer lab can be challenging... But if you move around the classroom it seems that it keeps the students more engaged, and gives to the opportunity to monitor what you students are viewing at their individual computers... Just a thought! -Nicole
White (below) [4/19/2010 15:05]	I agree with Jami about the set up of the classroom. I feel that any classroom with computers is hard to set up and hard to monitor all students at once.
Red (right, overlaid) [4/19/2010 20:15]	I agree. I don't think that this is the most efficient way for a classroom to be set up, because she can't see all of the computer screens to keep everyone on track

computer lab. There is nothing in previous annotations, comments, or in the video suggesting the teacher had designed the computer lab configuration. However, the following annotation then builds on this misconception, *The way she has the*

rows set up now are probably the most effective way (see Fig. 4, white annotation in the upper left). A red annotation follows with empathy, *teaching in a computer lab can be challenging*, and another white annotation agrees with the original student

Fig. 4 Group B's highly interactive CVA, including extensive overlapping of annotations

message, but seems to credit that post with the empathetic content expressed in the red post. In the end, the lexical choice *set up* is reiterated six times in the group discussion. The final annotation claims agreement, presumably with all the other annotations, but actually expresses the opposite perspective, that the instructor has *not* shown effective teaching by how she set up the lab. This logical incongruence is likely due to many of the previous annotations being obscured. Table 3 places the annotations in chronological order of appearance in the video, but it is important to note that not all of these were visible to learners during the entirety of the 20-s scene.

The communications backstory, provenance in Jewitt and Henriksen (2017) terms, can lend significant insight into the chaotic interaction on the screen. Participation came in bursts, with the final annotation in this scene appearing with only 3:45 left in the two-week asynchronous activity, at 8:15 pm. (See Table 3 for timestamps for all the annotations.) In the rush to complete the task, in a busy visual space, with complex media affordances covering up and concealing interlocutor contributions, this scene evidences the chaos in managing discourse in a complex media-enhanced space that requires careful thought to employ. Intentions to contribute admirably may have been at odds with outside factors that competed for learner attention. The first three annotations appeared in a seven-hour window more than a week before the due date, followed by a seven-day lapse in participation before another short burst of participation right before the entire task was due. The last two annotations appeared in a five-hour time span. Understood from this perspective, the curious path of the discussion is not nearly as surprising.

These contextual factors, combined with unfamiliarity with the media, created a task that was more difficult than it seemed. These preservice teachers brought with them assumptions about teaching that experienced teachers may have long abandoned, such as an assumption that teachers design computer labs. However, the impact of this unguided assumption was exacerbated by media choices. Learners obscured each other's annotations. Subsequent learners were unable to view the entirety of commentary while annotating. While asynchronous communication democratized the communication, the media also allowed a student to hold the floor for an excessively long time. The red annotation was placed into the timeline before the analyzed 20-s scene and remained until four seconds after it ended. The author of the earlier red annotation may have wanted her annotation to be read, but it ended up being obscured. Similarly, the white annotation completely blocked the view of the teacher in the video, obscuring the subject of the discussion. The Group B exemplar presents a case that collaborative video annotation is deceptively difficult, and understanding what it means to be literate in this space may require close analysis. The Group B exemplar contrasts dramatically with Group C.

Group C

Group C contained two annotations that supported each other. Both annotations address a related topic, include similar lexical decisions, and are placed unobtrusively into the video timeline and video real estate. The topical content of these two annotations in Group C were not semantically distanced from the other exemplars, but the discursive strategy was. The first learner to annotate the scene (see annotation in center of Fig. 5) has chosen to use a lexical item from the task itself, “engage.” Like the other exemplars, the discussion is tied to the teacher's behavior, talking about how the teachers use questions. The imagetext became a coordinated analysis of the video. Figure 5 shows an elegant interchange, unencumbered by obscured annotations and tightly focused on the video's action.

The annotation in the center of Fig. 5 appeared on the final day of the activity [4/19 13:20] and was colored green. This learner has chosen to do important linguistic work to facilitate the discussion by transcribing part of the teacher's talk from the video. When a learner pauses the video to annotate, there is no audio to support crafting text. In this case, the center annotation has already laid out the instructor's audio language in text form, making it far easier for the annotation on the left to respond. The responding annotation, inserted 4 h later [4/19 17:00], builds on the insight and makes the same lexical choice. The analytical discussion that takes place between the two learners is possible because one student has done the labor of transcribing the teacher's speech, making it easier for the other to contribute while the video is stopped to annotate. Similarly, both learners made sensitive location choices, not overlapping each other's annotations, nor blocking the view of the speaker in the video. In this case, the learners display a sensitivity to media and lexical choices that afforded a concrete and insightful discussion of the video.

Conclusion, Future Research, and Limitations

From three scenes, this article offered illustrative exemplars of how meanings were made via annotations atop video. The researcher observed and documented varying levels of participatory literacy, but also varying levels of difficulty as high numbers of simultaneous annotations negatively impacted the discussions. The complexity in navigating the space rose with the number of learners participating in a single scene. Designers of converged spaces for learning would be wise to consider such phenomena, as would those interested in understanding what constitutes participatory literacy in new media. To manifest the accomplishments of the “negotiating” that Solmaz (2017) identified of traversing communities, more scaffolding in media awareness may be required. These learners were residential learning classmates, who shared

Fig. 5 Group C screen capture shows a less active but more concrete discussion

undergraduate experiences in the same physical location. Future collaborative spaces may require even more sensitivity to media affordances if learners are drawn from diverse populations.

Oftentimes mediated learning is pursued because a shared geographical space is out of reach. Suppose we imagine the more likely scenario of learners geographically dispersed. In that case, we can imagine just how much more designers of instruction need to be aware of converged spaces' complexities because opportunities for synchronous interaction to resolve miscommunications would likely be fewer. Participatory media literacy aspires to bridge different communities (Baleria, 2019). This study provides cases, with necessary caveats of potential miscommunication, where geographical separation might be a source of sharing instead of an obstacle to learning. We are more alike than we are different; there is a shared experience to being a student. Recognizing conventions that develop among learners using certain media, such as participation just before the deadline as a convention of asynchronous forums, might offer ways to mitigate miscommunication while bridge disparate communities. Facilitating collaborations among those who inhabit very different communities has much to offer (Baleria, 2019; Lantela, 2019), but it will require sensitivity to media choices, lexical choices, and the provenance of our communications.

These exemplars also evidenced some adroit discursive tactics from learners who ostensibly had never communicated in this particular media before. Targeted research into which affordances need scaffolding, and which do not, could inform design, guiding literacy educators to where their effort is most

needed. The cases presented in this study suggest that technological configurations may have far more variability than designers of instruction may recognize. Educational researchers are still deciphering the learning potentials of converged spaces. While topic choice, lexical choice, and facility with the media were not predictable among these exemplars, in all three cases, on-task discussions did take place. Studying identical locations in the video timelines did not result in analyses of even similar interactions, but demonstrated successful meetings took place nevertheless. Stochastic designs such as collaborative video annotation (CVA) offer researchers opportunities to explore how discussions can result in learning performances, but also how learners might be led astray or even have their interactions buried in an inability to communicate and express themselves. Not all learners will bring adroit discursive tactics to mediated communication.

This semiotic analysis made only limited claims. Scenes were presented as exemplars, illustrations of possibility rather than likelihoods of interaction. Future research would benefit from more statistical calculations of larger samples that might place conditional probabilities on interactions to inform designers. Educators interested in advancing participatory media literacy may find utility in investigating other populations besides preservice teachers. Other groups may acquire proficiency in the media differently than this group—the less experienced may pick up proficiency in the media slower, and the more seasoned quicker, but this cannot be told from such a limited sample and such focused inquiry. While limited, this semiotic inquiry still provides literacy educators with a new perspective, an in situ chronological perspective.

Recording and noting the real-time, chronological appearance of each annotation facilitated an analysis of specific learner interaction and brought about observations otherwise impossible. These types of research strategies are necessary to understand how media literacy can impact learners. Calls for this have been made before. Grossmann and McDonald (2008) wrote, “The field of research on teaching still lacks powerful ways of parsing teaching that provide us with the analytic tools to describe, analyze, and improve” (p. 185). As more of our teaching migrates to mediated spaces, understanding learning experiences through such a theoretical lens can bring to light nuances of interactions that designers, scholars, and even the learners themselves could not have envisioned.

Acknowledgments I would like to acknowledge those who gave commentary on early drafts of this study, namely Colin M. Gray, Khosi L. Lunga, and Sam W. Burmester. Adrienne McPeake and Nellie McCollum helped me get up to date on the fast-paced evolution of video annotations. Tiffany Roman graciously allowed me to video record and share her excellent teaching with hundreds (and hundreds) of learners on YouTube. Thank you to you all.

Author Contribution All authors contributed to the study conception and design. Material preparation, data collection, analysis, and all writing was performed by Craig Howard. The first draft of the manuscript was written by Craig Howard. Craig Howard is the sole author of this article.

Declarations

Informed Consent Informed consent was obtained from all individual participants included in the study as per guidelines of the Office of Research Administration.

Conflict of Interest The author declares that they have no conflict of interest.

Research Involving Human Participants and/or Animals All procedures performed involving human participants were in accordance with the ethical standards of the Indiana University Office of Research Administration (IUB Human Subjects Office Exemption Granted for New Protocol Title: *Video Annotations as Communicative Discourse* #1001000991) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

- Baleria, G. (2019). Story sharing in a digital space to counter othering and foster belonging and curiosity among college students. *Journal of Media Literacy Education*, 11(2), 56–78. <https://doi.org/10.23860/JMLE-2019-11-2-4>.
- Baylen, D. M., & D'alba, A. (Eds.). (2015). *Essentials of teaching and integrating visual and media literacy: Visualizing learning*. Springer.
- Butler, A., Fuentes-Bautista, M., & Scharrer, E. (2018). Building a media literacy in higher education: Department approaches, undergraduate certificate, and engaged scholarship. In J. Cubbage (Ed.), *Handbook*

- of research on media literacy in higher education environments* (pp. 153–171). IGI Global.
- Cattaneo, A., Boldrini, E., & Lubinu, F. (2020). “Take a look at this!”: Video annotation as a means to foster evidence-based and reflective external and self-given feedback: A preliminary study in operation room technician training. *Nurse Education in Practice*, 44, 102770. <https://doi.org/10.1016/j.nepr.2020.102770>.
- Fadde, P., & Sullivan, P. (2013). Using interactive video to develop pre-service teachers’ classroom awareness. *Contemporary issues in technology and teacher education*, 13(2), 156–174. Waynesville, NC USA: Society for Information Technology & Teacher Education. <https://www.learntechlib.org/primary/p/42133/>
- Flynn, J. E., & Lewis, W. (2015). Multimodal composition in teacher education: From consumers to producers. In D. M. Baylen & A. D’Alba (Eds.), *Essentials of teaching and integrating visual and media literacy* (pp. 147–163). Springer.
- Frohlich, D. O., & Magolis, D. (2020). Developing a responsive and adaptable emergent media curriculum. *Journal of Media Literacy Education*, 12(1), 123–131. <https://doi.org/10.23860/JMLE-2020-12-1-9>.
- Grossmann, P., & McDonald, M. (2008). Back to the future: Directions for research in teaching and teacher education. *American Educational Research Journal*, 45(1), 184–205. <https://doi.org/10.3102/0002831207312906>.
- Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science*, 28, 115–152. <https://doi.org/10.1023/A:1003764722829>.
- Herring, S. C. (2013). Discourse in Web 2.0: Familiar, reconfigured, and emergent. In D. Tannen & A. M. Trester (Eds.), *Georgetown University round table on languages and linguistics 2011: Discourse 2.0 language and new media* (pp. 1–25). Georgetown University Press.
- Hobbs, R. (2011). The state of media literacy: A response to Potter. *Journal of Broadcasting & Electronic Media*, 55(3), 419–430.
- Howard, C. D. (2012). *Higher order thinking in collaborative video annotations: Investigating discourse modeling and the staggering of participation*. (publication no. 1287116198) [Doctoral dissertation, Indiana University Bloomington]. ProQuest Dissertations Publishing.
- Howard, C. D. (2019). Critical thinking in collaborative video annotations: Relationships between criticism and higher order thinking. In G. Mariano & J. Figliano (Eds.) *Handbook of research on critical thinking strategies in pre-service learning environments*. (pp. 493–514). IGI Global Information Science Reference Series. <https://doi.org/10.4018/978-1-5225-7823-9.ch024>.
- Howard, C. D., & Myers, R. D. (2010). Creating video-annotated discussions: An asynchronous alternative. *International Journal of Designs for Learning*, 1(1), multimedia. <https://doi.org/10.14434/ijdl.v1i1.853>.
- Hull, G. A., & Nelson, M. E. (2005). Locating the semiotic power of multimodality. *Written Communication*, 22(2), 224–261.
- Jenkins, H. (2006). *Convergence culture*. New York University Press.
- Jewitt, C., & Henriksen, B. (2017). Social semiotic multimodality. In N. M. Klug & H. Stöckl (Eds.), *Handbuch Sprache im multimodalen Kontext / Handbook of language in multimodal contexts* (pp. 145–164). De Gruyter.
- Kersch, D., & Lesley, M. (2019). Hosting and healing: A framework for critical media literacy pedagogy. *Journal of Media Literacy Education*, 11(3), 37–48. <https://doi.org/10.23860/JMLE-2019-11-3-4>.
- Korenman, J., & Wyatt, N. (1996). Group dynamics in an email forum. In S. C. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 225–242). John Benjamins.

- Kress, G. (2009). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Kress, G. R., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Psychology Press.
- Lantela, P. (2019). “So, tell me what kind of a thing it really is”—Finnish older adults making sense of home technology. *Journal of Media Literacy Education*, 11(2), 122–145. <https://doi.org/10.23860/JMLE-2019-11-2-8>.
- Mayer, R. E. (2010). *Multimedia learning*. Cambridge University Press.
- Merrill, M. D. (2009). Finding e3 (effective, efficient, and engaging) instruction. *Educational Technology*, 49(3), 15–26.
- Mirriahi, N., Liaqat, D., Dawson, S., & Gašević, D. (2016). Uncovering student learning profiles with a video annotation tool: Reflective learning with and without instructional norms. *Educational Technology Research and Development*, 64(6), 1083–1106.
- National Association for Media Literacy Education. (2020a). *Media literacy defined*. <https://namle.net/publications/media-literacy-definitions/>
- National Association for Media Literacy Education. (2020b). *Core principles of media literacy education in the United States*. <https://namle.net/wp-content/uploads/2020/06/Namle-Core-Principles-of-MLE-in-the-United-States.pdf>
- Pérez-Torregrosa, A. B., Díaz-Martín, C., & Ibáñez-Cubillas, P. (2017). The use of video annotation tools in teacher training. *Procedia - Social and Behavioral Sciences*, 237, 458–464.
- Potter, W. J. (2004). *Theory of media literacy: A cognitive approach*. Sage Publications.
- Rich, P. J., & Hannafin, M. (2009). Video annotation tools: Technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education*, 60(1), 52–67.
- Rosaen, C. L., Lundeberg, M., Cooper, M., Fritzen, A., & Terpstra, M. (2008). Noticing noticing: How does investigation of video records change how teachers reflect on their experiences? *Journal of Teacher Education*, 59(4), 347–360.
- Sansone, K. L. (2015). Using strategies from graphic design to improve teaching and learning. In D. M. Baylen & A. D’Alba (Eds.), *Essentials of teaching and integrating visual and media literacy* (pp. 3–26). Springer.
- Solmaz, O. (2017). Adapting New Media Literacies to Participatory Spaces: Social Media Literacy Practices of Multilingual Students. *Journal of Media Literacy Education* 9 (1), 36–63
- Tugtekin, E. B., & Koc, M. (2020). Understanding the relationship between new media literacy, communication skills, and democratic tendency: *Model development and testing*. *New Media & Society*, 22(10), 1922–1941. <https://doi.org/10.1177/1461444819887705>
- Van Es, E. A. (2010). A framework for facilitating productive discussions in video clubs. *Educational technology: The magazine for managers of change in education*, 50(1), 8–12.
- Wright, W. T. (2020). Within, without, and amidst: A review of literacy educators’ perceptions of participatory media technologies. *Journal of Media Literacy Education*, 12(2), 1–12. <https://doi.org/10.23860/JMLE-2020-12-2-1>
- Zaier, A., Arslan-Ari, I., & Maina, F. (2020). The use of video annotation tools and informal online discussions to explore preservice teachers’ self- and peer-evaluation of academic feedback. *Sage Journal of Education*, 201(1), 19–27. <https://doi.org/10.1177/0022057420903269online>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.