Supporting Constructive Video-based Learning: Requirements Elicitation from Exploratory Studies

Antonija Mitrovic[1], Vania Dimitrova[2,3], Lydia Lau[2], Amali Weerasinghe[4], Moffat Mathews[1]

[1]Intelligent Computer Tutoring Group, University of Canterbury, New Zealand
[2]School of Computing, University of Leeds, UK
[3]Leeds Institute of Medical Education, University of Leeds, UK
[4]School of Computer Science, University of Adelaide, Australia

Abstract. Although videos are a highly popular digital medium for learning, video watching can be a passive activity and results in limited learning. This calls for interactive means to support engagement and active video watching. However, there is limited insight into what engagement challenges have to be overcome and what intelligent features are needed. This paper presents an empirical way to elicit requirements for innovative functionality to support constructive video-based learning. We present two user studies with an active video watching system instantiated for soft skill learning (pitch presentations). Based on the studies, we identify whether learning is happening and what kind of interaction contributes to learning, what difficulties participants face and how these can be overcome with additional intelligent support. Our findings show that participants who engaged in constructive learning have improved their conceptual understanding of presentation skills, while those who exhibited more passive ways of learning have not improved as much as constructive learners. Analysis of participants' profiles and experiences led to requirements for intelligent support with active video watching. Based on this, we propose intelligent nudging in the form of signposting and prompts to further promote constructive learning.

Keywords: Video-based Learning, Intelligent support; Requirements elicitation, Experimental studies, Soft skill learning

1 Introduction

Videos have become the main means for content production and consumption for the millennials and iGeneration. Video-based learning [27] is used in a wide spectrum of instructional settings, ranging from flipped classrooms [15], online learning and MOOCS [10,23] to informal learning using YouTube. However, watching videos is inherently a passive form of learning; in order to learn effectively, students need to engage with video content [3-6,13,20,27]. Engagement with videos can be facilitated by embedding interactive activities, such as quizzes and assessment problems [8,12,14, 24], or by providing environments for collaborative annotation of videos [3]. Although such strategies increase engagement, they require substantial effort from the teacher during video production, or sophisticated learning environments.

adfa, p. 1, 2011.

Our approach is to support engagement via interactive notetaking, tapping into learners' familiarity with commenting on videos in social networking sites. For example, in CourseMapper [3], learners can annotate videos, discuss and vote/rate annotations. However, in video annotation environments students annotate videos freely, and there is no explicit support for personalisation. Our approach differs in that we channel support for interaction with important elements of videos via aspects, i.e. micro-scaffolds that direct students' attention on skill-related concepts and foster reflection.

We developed the Active Video Watching (AVW) system [16,18]. AVW is aimed at soft skills learning (such as communicating, negotiating, collaborating, critical thinking), which are crucial for employability in the knowledge economy [26]. Videos can be a powerful method for soft skills training [2,5,6], where learning requires contextualisation in personal experience and ability to see different perspectives. We conducted two studies using the AVW platform to learn about giving pitch presentations. The findings can inform further improvements of the AVW platform (similar to [22,25]), and future enhancements with intelligent nudging features to improve learning.

The paper is structured as follows. Section 2 presents AVW and the operationalisation of the ICAP framework for active video watching. The experimental design is presented in Section 3, followed by findings and elicited requirements in Section 4. Section 5 discusses possible nudging features.

2 Operationalisation of the ICAP framework for AVW

ICAP Framework. Educators agree that engagement is crucial for effective learning [4,19,27]. In a classroom, the teacher can form judgments about students' levels of engagement. However, engagement in online learning (including learning from videos) is often low, and overt actions students perform are the only source of information about their engagement. The ICAP Framework [4] classifies overt learner behaviours into four type of learning modes, corresponding to different levels of cognitive engagement: Interactive, Constructive, Active and Passive. Passive learners are simply receiving information, without performing any additional actions; they might be observing a lecture, reading a book or watching a video, but do not engage further. Active learners do exhibit additional actions, such as note taking, but those actions simply replicate provided information; for example, writing down lecturer's statements, or rewinding the video to watch important parts multiple times. In the constructive mode, the learner generates new information that was not explicitly taught; e.g. summary of points, a concept map, or a self-explanation. In the interactive mode, learners engage in discussions with their peers, which allow them to compare and contrast their opinions, and jointly generate solutions to problems. Chi and Wylie [4] provide evidence that as students become more engaged, starting from the passive mode to the interactive mode, the learning effectiveness increases; i.e. Passive < Active < Constructive < Interactive.

AVW platform. AVW is a controlled video watching environment designed for self-study. It can be customised by the teacher who defines a list of aspects that serve as scaffolds for learning with videos. The choice of aspects should direct the student's attention on skill-related concepts and foster reflection.

AVW offers Personal Space and Social Space (Fig. 1). Initially students watch and comment on videos individually in the Personal Space, using aspects to tag their comments. The system time-stamps comments (i.e. the time elapsed from the start of video). The student can watch videos multiple times, including rewinding or skipping parts. Once the teacher approves comments for sharing, anonymised comments are available to the whole class in the Social Space in which students can browse and rate comments. The students can sort the comments by timestamp or aspect. The teacher defines options for rating to promote deeper reflections. In addition to reading/rating the comments, the students can watch the part of the video associated with a comment.

The AVW platform was instantiated in systems hosted by the Universities of Leeds and Canterbury, respectively. Both instances had identical basic functionality, with the same set of videos and customisation by the teacher. A few small differences include the possibility to add a comment without specifying an aspect in the former instance, while aspects were made mandatory in the latter instance.

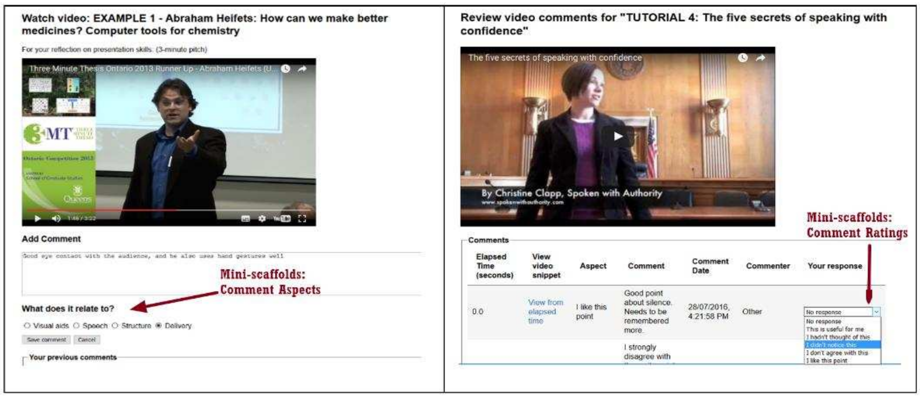

Fig. 1. Adding a comment (Personal Space, left); and rating a comment (Social Space, right).

Operationalising ICAP in AVW. We operationalised the ICAP framework in the context of AVW as follows. Passive Learners are those who watched videos, but have neither manipulated them, or written comments on them. Active Learners are similar in that they do not comment on videos, but manipulate videos (e.g. watching videos multiple times, fast forwarding or rewinding videos). We consider a combined category **Passive/Active Learners (P/AL)** indicating students who watched videos without commenting. **Constructive Learners (CL)** show higher levels of engagement by commenting on videos. Comments, as we will show in Section 4, contain remarks on important events in videos, and contain statements showing reflection and self-explanation. AVW does not currently support collaboration between students, and therefore we do not consider the Interactive mode of ICAP. In addition to P/AL and CL, we have also added another mode to characterise students who do not engage in learning at all, i.e. do not watch videos; we refer to them as **Inactive Learners (IL)**.

3 Experimental Design

Aim. We conducted two user studies with undergraduate (UG) and postgraduate (PG) university students using AVW to support soft skill learning, namely giving pitch presentations. Ethical approvals were obtained from the Universities of Leeds and Canterbury. The main aim was to elicit requirements for intelligent support to improve learning with AVW. We investigated four **research questions**:

- Does AVW support learning (if so, what ICAP behaviour increases knowledge)?
- Do micro-scaffolds help (if so, are there any notable usage patterns)?
- Do the learner profiles differ (if so, what are the important differences)?
- What *is the learners' experience with AVW* (are there any critical difficulties)?

Materials. The videos used in the study were carefully selected from YouTube. Four were tutorials on giving presentations, while the other four were actual recordings of pitch presentations (two TED talks, and two 3-minute PhD pitch presentations). The criteria for selecting the videos were: (i) appropriate content (covering opening, closing, structure, delivery and visual aids; or examples of pitch presentations); (ii) no longer than 10 minutes; (iii) balance of gender for the presenters; (iv) two popular examples and two not so popular (based on the YouTube ratings).

The micro-scaffolds used were related to the target soft skill (Fig. 1). There were three reflective aspects chosen for tutorials: "*I didn't realize I wasn't doing it*" (TA2), "I am rather good at this" (TA3), "I did/saw this in the past" (TA4); these aspects stimulate learners to recall and reflect on their own experiences. There was one additional aspect, "I like this point" (TA1), which allows the learner to externalise learning points. For the example videos, the aspects were: "Delivery" (EA1), "Speech" (EA2), "Structure" (EA3), and "Visual aids" (EA4), corresponding to the concepts covered in the tutorials. Ratings in the Social Space also aimed to promote reflection.

Three surveys were designed to collect data. Survey 1 collected participant's profile (demographic information, background experiences, Motivated Strategies for Learning Questionnaire (MSLQ) [21]); and participants' knowledge of presentations. Survey 2 included the same questions for knowledge of presentations; NASA-TLX instrument [11] to check participants' perception of cognitive load when commenting; Technology Acceptance Model (TAM) [7] to check participants' perceived usefulness of commenting on videos for learning; and questions on usability related to commenting on videos. Survey 3 was similar to Survey 2 but related to rating others' comments.

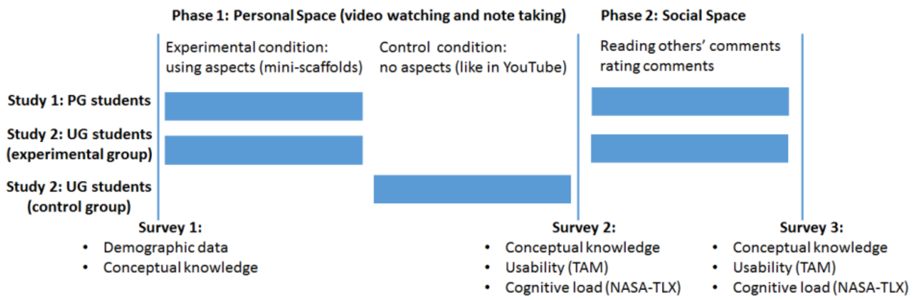

Fig. 2. Outline of experimental set-ups

Procedure. The experimental investigation included two studies (Fig. 2). Study 1 (conducted in March 2016) included PG volunteers recruited via online communities, while Study 2 (conducted in July 2016) included UG engineering students from the University of Canterbury. The goal of Study 1 was to identify whether learning is happening in AVW. The goals of Study 2 were to identify whether the aspects are effective as micro-scaffolds for reflection, and to identify the effect of rating comments in Phase 2 (Social Space) on learning. Hence in Study 2 there were two conditions: experimental (equivalent to Study 1) and control (used AVW without aspects). Both studies were two weeks long. Week 1: After providing informed consent, the participants took Survey 1, watched and commented on the tutorials, then continued with the examples, and completed Survey 2. Week 2: the participants (except those in control condition) rated comments made by other participants and completed Survey 3.

Assessing Conceptual Knowledge. Each survey contained three questions related to participants' knowledge of presentation skills. Participants had one minute per question to write phrases they associated with (i) structure, (ii) delivery and speech, and (iii) visual aids. We developed an ontology of presentations, consisting of three taxonomies related to these areas. Each response was marked by three independent markers, indicating the number of ontology entities associated with the response. The inter-rater reliability was high: the Krippendorff's alpha was for 0.894 for Study 1, and 0.907 for Study 2. The final scores for conceptual understanding were confirmed by a fourth marker using the majority vote, or if that was not possible, re-marking the entries.

Participants. Study 1 started with 48 participants, 38 of whom commented on videos and completed surveys (26 females and 12 males; 3 aged younger than 24, 14 aged 24-29, 10 aged 30-35, 5 aged 36-47, and 6 aged 48 or older; 23 with English as first language, while 15 with Asian/European languages as mother tongue; 28 were PhD students and 10 were Masters). In **Study 2**, 37 participants were randomly assigned to either the experimental group (17 males and 2 females) or control group (13 male and 5 female). The majority of participants (83.8%) were aged 18-23. Sixteen Inactive Learners did not use AVW (although some completed all surveys). The remaining students watched the videos, including 8 Passive/Active Learners (4 control, 4 experimental), and 13 Constructive Learners (6 control, 7 experimental).

4 Findings: Recommendations for Intelligent Support

Did AVW support learning? Table 1 reports the conceptual knowledge scores from Surveys 1-3. Some participants have not completed all surveys, and therefore we provide the actual numbers of participants who have taken each survey in the table. We found evidence of learning: a repeated measures ANOVA on the conceptual knowledge scores for Study 1 revealed a significant effect overall ($F(2,68) = 6.18$, $p = .003$) with the partial eta squared of 0.15 (medium effect). The pairwise comparison shows there was a significant increase from Survey 1 to Survey 3 ($p = .01$). For constructive participants from Study 2, the Friedman test also revealed a significant difference on conceptual understanding scores ($\chi 2(2) = 7.89$, $p = 0.02$). The effect size was large (0.67). There was not enough data to analyse statistical significance of differences for IL and P/AL, but the scores on Survey 3 are lower than earlier scores. Some ILs completed

Surveys 2 and 3 without watching any videos; their conceptual knowledge answers contained the same entries, often using irrelevant concepts.

There were no significant differences between CL, P/AL and IL categories on the conceptual knowledge scores from Survey 1, showing that all categories started with similar conceptual knowledge. However, there was a marginally significant difference on the scores for Survey 2 (H = 3.35, p = .09).

Table 1. Comparing conceptual knowledge by category (scores indicate the number of relevant domain concepts mentioned in the participants' conceptual knowledge answers).

	CL Study 1	**CL Study 2**	**P/AL Study 2**	**IL Study 2**
Pre-test before using AVW (Survey 1)	12.89 (6.44) n = 38	13.62 (4.03) n = 13	11.63 (2.97) n = 8	10.63 (4.95) n = 16
Post-test PersonalSpace (Survey 2)	13.74 (6.46) n = 38	17 (4.52) n = 10	11.2 (5.45) n = 5	10.13 (4.82) n = 8
Post-test SocialSpace (Survey 3)	15.86 (6.18) n = 35	18.4 (3.72) n = 5	7.5 (9.19) n = 2	9.5 (6.36) n = 2

When comparing Study 2 participants, there was a significant difference on the conceptual knowledge scores for Survey 2 (H = 7.25, p = .03), with a significant difference between IL and CL (p = .03). We have not compared scores from Survey 3 due to low user numbers. Not all participants engaged in constructive learning, consequently, they did not improve their conceptual knowledge. A large group of participants (43%) from Study 2 have not watched any videos (IL). The percentage of IL in Study 1 is much smaller (20%). We have no data about why ILs have not watched videos. We attribute this to the voluntary nature of the study and demands by other learning activities. In Study 2, 21% of participants watched videos but made no comments (P/AL).

Finding: Only constructive behaviour in both AVW spaces (writing comments and browsing/rating comments) led to increased conceptual understanding. Passive/Active and Inactive behaviour did not lead to increased conceptual understanding.

R1: Further enhance both the Personal Space and the Social Space with intelligent support to foster active video watching that leads to constructive learning behaviour.

Table 2. Comparing control and experimental conditions in Study 2

	Constructive Learners		**Passive/Active Learners**	
	Control (5)	**Exper. (5)**	**Control (3)**	**Exper. (1)**
Pre-test (Survey 1)	13.2 (3.96)	12.2 (2.28)	11 (2.65)	13
Post-test Survey 2	15.8 (2.59)	18.2 (5.98)	12 (6.93)	13
Post-test Survey 3	N/A	18.4 (3.72)	N/A	15

Did micro-scaffolds help? Study 2 was designed to identify the effect of micro-scaffolds (i.e. aspects and ratings) on learning. The participants in the control condition used AVW without micro-scaffolds. Table 2 provides scores for participants who have completed all surveys. The only significant difference on conceptual knowledge scores is for CL from the experimental group ($\chi^2(2) = 7.89$, p = 0.02). The effect size was large (0.667), and the scores from Survey1 and Survey 3 are significantly different (p = .01).

Overall, 1029 comments were generated (790 in Study 1 and 239 in Study 2). There was no significant difference between the average number of comments made by CL from Study 1 (19.58, sd = 13.19) and Study 2 (18.38, sd = 16.59). In Study 2, CL from

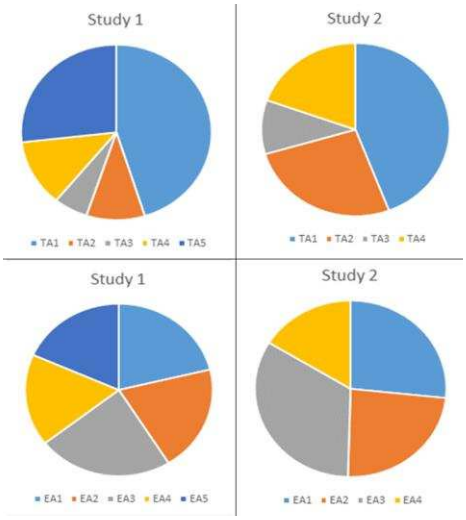

Fig. 3. Percentages of comments based on aspects

the experimental condition made on average 12.86 (sd = 11.65, range = [1,29]) comments, and CL from the control group made an average of 24.83 (sd = 20.13, range = [6,51]). There was no significant difference on the number of comments made by CL from the two conditions. Hence, making comments alone does not contribute to learning; specifying aspects and rating comments is needed. Only 28% of the comments by CL in Study 1 used reflective aspects (TA2, TA3, TA4), while in Study 2 that percentage was 49% (Fig. 3). Study 1 allowed making comments without selecting aspects, and 26.8% of the comments on tutorials (TA5) and 18% of the comments on examples (EA5) were without aspects. As the effect size for CL in Study 1 was medium, versus large effect size in Study 2, the use of aspects and especially reflective aspects lead to increased conceptual knowledge.

The experimental group students from Study 2 provided 332 ratings, with two students providing 150 ratings (73 and 77, respectively). In Study 1, AVW did not log who made ratings. Table 3 shows the split of ratings according to categories. The first three ratings trigger learning, as they show that the participant noticed something new in comments, and we label this class as Trigger Learning. The other class (Induce Opinion) consists of two categories, when participants disagree with comments, or simply state that they like them. Therefore, participants tend to state opinion (although these rating categories were at bottom of the list of categories provided in the interface).

Table 3. Number of ratings on comments – CL engagement in Social Space.

	Rating category	**Study 1**	**Study 2**
Trigger Learning	This is useful for me	349	122
	I hadn't thought of this	260	23
	I didn't notice this	241	30
Induce Opinion	I do not agree with this	213	29
	I like this point	1643	128

Finding: The use of micro-scaffolds (aspects in the Personal Space and ratings in the Social Space) have positive effect on learning. Mandatory use of aspects for all videos and higher usage of reflective aspects in tutorials led to a larger effect size. The use of rating categories was uneven, most referred to state opinion (like/dislike) as opposed to ratings that trigger reflection and further learning.

R2: In the Personal Space, make it mandatory to indicate an aspect when a comment is made. Include intelligent support to encourage students to use a diverse range of aspects, and give preference to aspects that trigger reflection.

R3: In the Social Space, include intelligent support to encourage students to rate comments, and give preference to ratings that trigger reflection.

R4: Include the use of micro-scaffolds in the learner profile that can be used to personalise the intelligent support in both the Personal Space and the Social Space.

Did the profiles of the learner categories differ? Table 4 presents the basic statistics for the different categories of participants from the two studies. In Survey 1, the participants' profiles were collected, using the Likert scale ranging from 1 (lowest) to 5 (highest). There was a significant difference on participants' experience in giving presentations (H = 7.99, p = .046, no significant pairwise differences). There was a significant difference on the use of YouTube (H = 10.14, p = .02), with significant difference between the constructive participants from both studies (H = 17.16, p = .05) which is not surprising, as the participants in Study 1 were older than those in Study 2. There was no significant difference on use of YouTube for learning.

Table 4. Comparing catergories of participants (** and * denote significance at the 0.01 and 0.05 level respectively; Kruskal-Wallis with Bonferroni correction for pairwise comparisons).

	Constructive Study1 (38)	**Construct. Study2 (13)**	**Pass./Active Study2 (8)**	**Inactive Study 2 (16)**
Training	2.16 (.95)	1.77 (.59)	1.5 (.53)	1.81 (.75)
Experience*	2.87 (.78)	2.77 (.59)	2.25 (.46)	2.44 (.73)
YouTube*	3.5 (1.11)	4.38 (.65)	4.13 (.64)	4.19 (.98)
YouTube/learning	2.71 (1.01)	2.85 (.89)	2.62 (1.19)	3.25 (1)
MSLQ Task Value**	4.49 (.38)	3.95 (.4)	3.83 (.53)	4.02 (.45)
MSLQ Self-Efficacy	3.72 (.56)	3.46 (.72)	3.88 (.56)	3.66 (.4)
MSLQ Acad. Control	3.91 (.46)	4.04 (.49)	4.25 (.68)	4.22 (.58)
MSLQ Intrinsic	4.05 (.52)	3.79 (.35)	3.72 (.68)	3.79 (.51)
MSLQ Extrinsic	3.37 (.74)	3.62 (.33)	3.97 (.59)	3.41 (.82)
MSLQ Effort Regul.**	3.81 (.57)	3.92 (2.28)	3.53 (.54)	3.45 (.55)
MSLQ Rehearsal	3.4 (.8)	2.94 (.85)	2.88 (.88)	2.94 (.92)
MSLQ Organization**	3.84 (.94)	3.27 (1.25)	2.38 (1.03)	3.02 (1.07)
MSLQ Elaboration**	4.13 (.54)	3.67 (.49)	3.63 (.74)	3.55 (.75)
MSLQ Self-Regul.**	3.56 (.49)	2.82 (.51)	3.31 (.54)	3.23 (.46)

Survey 1 contained also 46 questions from the MSLQ, with the Likert scale of 1 (Not at all true of me) to 5 (Very true of me). The MSLQ questions were summarised into ten scales reported in Table 4. The scores for Task value are significantly different (H = 22.39, p < .05), with CL from Study 1 having higher response than a) A/PL (H = 25.73, p = .011), b) CL from Study 2 (H = 24.89, p = .002), and c) IL (H = 20.137, p = .009) respectively. There was a significant difference on Effort regulation (H = 14.6, p = .002), with Study 2 CLs providing higher scores in comparison to IL (H = 18.92, p = .02) and A/PL (H = 21.829, p = .05). For Organisation (H = 15.52, p = .001), again the PG students scored significantly higher than A/PL (H = 27.88, p = .005) and IL (H = 17.97, p = .03). Similarly, there were significant differences for Elaboration (H = 14.1, p = .003), with PG participants scoring higher than IL (H = 19.48, p = .015), and for Self-regulation (H = 21.35, p = 0.), with PG participants scoring higher than CL from Study 2 (H = 30.68, p = 0). These findings show that PG students generally have better strategies for learning in comparison to P/AL and IL. The only significant differences for the CLs from the two studies were on Task value and Self-regulation.

Finding: Students who are more experienced in the target soft skill are more likely to exhibit constructive learning behaviour. There were differences in MLSQ scales.

R5: Include past experience and MSLQ scales in the learner profile so they can be used to personalise intelligent support in both the Personal Space and the Social Space.

R6: Include different strategies for intelligent support. For Constructive Learners, encourage them to refer to past experience in comment writing and rating. For Passive/Active learners, encourage elaboration, self-regulation, and organisation in comment writing and rating; as well as indicate the task value of active video watching.

R7: Conduct intelligent analysis to further categorise constructive learning in order to identify personalised strategies for this category of learners.

What was the learners' experience with AVW? The participants' perceptions on commenting on the videos (Survey 2) and rating comments (Survey 3) were collected using the NASA-TLX questionnaire on the cognitive workload and the TAM questionnaire measuring perceived usefulness. The participants faced some difficulties.

Cognitive demand. Four NASA-TLX questions measured: how demanding commenting/rating comments was, how much effort was required, how frustrating the activity was, and how well the participant felt he/she performed (Table 5).

Table 5. Average scores for NASA-TLX cognitive load (Likert scale from 1-Low to 20-High) and TAM perceived usefulness (Likert scale from 1-High to 7-Low)

		Constructive Study 1	**Constructive Study 2**	**Passive/Active Study 2**
NASA-TLX Demand	Personal Space	9.89 (4.87)	11.1 (4.95)	10 (7.28)
	Social Space	8.86 (4.84)	9 (4.42)	13.67 (3.21)
NASA-TLX Effort	Personal Space	8.55 (4.21)	8.9 (2.99)	7.4 (5.03)
	Social Space	8.37 (4.89)	7.4 (4.34)	15.67 (.58))
NASA-TLX Frustration	Personal Space	5.79 (4.49)	8.5 (5.06)	5.8 (5.45)
	Social Space	8.63 (6.17)	8.8 (5.36)	5.67 (6.43)
NASA-TLX Performance	Personal Space	12.76 (4.48)	11.5 (5.29)	9.4 (7.7)
	Social Space	10.4 (6.09)	7.6 (3.91)	9.67 (8.5))
TAM Usefulness	Personal Space	3.91 (.38)	3 (.89)	3.68 (1.61)
	Social Space	3.33 (1.77)	4.72 (1.35)	3.87 (6.43)

We do not report the scores for Inactive participants, as they have not interacted with AVW. There were no significant differences between the categories on any of the cognitive load values. The participants found commenting on the videos and rating comments moderately demanding. In relation to demand, 45% of Study 1 participants explicitly noted that commenting on videos prompted thinking, which is evidence of the effectiveness of aspects to support reflection. Seven participants stated they made links with their past experience, e.g.: "I needed to pay proper attention to understand what was explained, to recall my experience, and perceive the usefulness of the tricks and tactics told by the presenter."

The participants from Study 1 found rating comments more frustrating than commenting on videos ($t = 2.89$, $p = .007$), and stated their performance on rating lower than on commenting ($t = 2.14$, $p = .04$). The qualitative feedback on frustration pointed at the large number of comments to be rated, which was time-consuming, as well as the fact that many comments were similar. The participants suggested presenting comments in a structured way, and providing ways to discuss comments with others.

Perceived usefulness. Table 5 reports the average of five TAM questions related to the perceived usefulness of commenting on video in the Personal Space, and to rating comments in the Social Space. The Kruskal-Wallis test revealed a significant difference on Usefulness for commenting on videos (H = 11.54, p = .01), with a pairwise significant difference between the constructive participants from the two studies (p = .013). The constructive participants from Study 2 found commenting on videos more useful than PG students, which can be explained by the fact that UG students had less experience overall with presentations than PG students. There was no significant difference on Usefulness for rating comments across the categories.

The PG participants found rating comments marginally significantly more useful than commenting on videos (t = 1.95, p = .06), while the constructive participants from Study 2 ranked them in the opposite way (W = 10, p = .07). The participants were positive about the functionality provided by AVW, and stated that commenting on videos focused attention on important parts of videos, kept them alert and active, and reinforced learning. The majority of participants stated that rating comments supports learning by sharing understanding (when comments are in agreement) and also seeing points from a different perspective. However, 20% of participants did not find rating comments useful; some stated that others' comments were not of good quality, and that presenting comments in a different way (e.g. summary) would be more beneficial.

Finding: Writing comments was cognitively demanding, as participants needed to identify appropriate places in the video and to reflect on past experience. Participants found rating comments relatively frustrating; feedback showed that this was caused by: (i) overwhelming quantity of comments to read and rate; (ii) reading comments of low *quality; (iii) seeing many comments similar to one's own;* and (iv) lack of structure.

R8: In the Personal Space, add interaction means to aid the reflection process; add means to encourage users to write high quality comments to be used in the Social Space.

R9*: In the Social Space, direct learners' attention to high quality comments and to* comments that show different perspectives; provide a structure to browse comments.

5 Discussion: Towards Intelligent Nudging

Following the requirements (R1 to R9) in the previous section, we identify future enhancements of AVW with intelligent nudging to promote constructive video-based learning. Intelligent nudges are personalised interventions aimed to influence user behaviour towards constructive learning without limiting users' personal choices for engaging in AVW. Following previous research for using nudges in learning environments [22], we consider two types nudges – signpostings and prompts.

Learner model. Use explicit profiling by asking students about their experience in the target skill and MSQL scales (R5) and implicit profiling from the interaction logs, including number of comments, use of aspects and ratings (R4). Machine learning can be used to further characterise constructive learning (R7), including clustering of CL and prediction model to identify students likely to be P/AL.

Signposting can be added to both AVW spaces, including: (1) showing 'high *attention'* parts of the video which attracted comments by previous participants to encourage making comments (R1), to facilitate reflection (R8), and to promote rating comments (R3); (2) encourage indication of aspects (R2) and use of ratings (R3) by showing *'focused attention'* parts in the video where comments/ratings predominantly refer to one specific aspect/rating and 'diverse attention' parts where a range of aspects are used. The former can prompt the use of a specific aspect/rating while the latter can potentially show multiple perspectives critical for soft skill learning; (3) using open student models to aid students' awareness of their engagement together with open social student models [1,9,17] to allow social comparison to motivate participation (R2), foster meta-cognitive activities (R8), and indicate the quality of comments (R8 and R9).

Prompts can provide contextualised nudging tailored to the learner's profile and engagement behaviour, including: (1) 'other students made good comments about this part' - encouraging a participant who has not commented on a part of a video that attracted attention of other students (R1) and may suggest possible aspects that others have used (R2); (2) 'can you relate to your past *experience'* - encourage students to refer to past experiences by using the corresponding aspects/ratings (R6) and suggest what other people have said about their past experience; (3) 'have you thought about' – diversify the use of aspects and ratings when the learner tends to use only a fraction of aspects/ratings (R2, R3); (4) 'you may find this useful' – P/AL can be motivated with tips for organisation, self-regulation, and elaboration, and suggestions how AVW can help with these (R6); (5) 'well done' – provide positive feedback to recognise both good quality comments (R8) and use of a variety of reflective aspects and ratings (R2, R3). This requires a deeper analysis of comments, employing the developed ontology to provide words/entities for textual and semantic analysis.

Conclusions. The findings from our studies show that when learners engage in commenting on videos and rating others' comments, their conceptual understanding of the target soft skills increases. We reported a number of ways for further enhancements of AVW, using intelligent nudges. Future plans include enhancing AVW and performing more studies focusing on various soft skills.

Acknowledgments. We would like to thank our participants and Kim Rutter, the instructor of ENEL301. This research was supported by the EU-FP7-ICT-257184 ImREAL grant, a teaching development grant from the University of Canterbury, and a regional grant from Ako Aotearoa.

References

1. Bull, S., Kay, J. (2016). SMILI☺: a framework for interfaces to learning data in Open Learner Models, learning analytics and related fields. Artificial Intelligence in Education, 26(1), pp. 293-331.
2. Cecez-Kecmanovic, D., Webb, C. (2000) Towards a communicative model of collaborative web-mediated learning. Australasian Journal of Educational Technology, 16(1).
3. Chatti, M. A., Marinov, M., Sabov, O., et al. (2016). Video annotation and analytics in CourseMapper. Smart Learning Environments, 3(1), 10.
4. Chi, M. T., Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. Educational Psychologist, 49(4), pp. 219-243.
5. Conkey, C. A., Bowers, C., Cannon-Bowers, J., & Sanchez, A. (2013) Machinima and Video-Based Soft-Skills Training for Frontline Healthcare Workers. Games for health, 2(1), pp.39-43.

6. Cronin, M. W., Cronin, K. A. (1992) Recent empirical studies of the pedagogical effects of interactive video instruction in "soft skill" areas. Computing in Higher Education, 3(2), 53.
7. Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, pp. 319-340.
8. Giannakos, M., Sampson, D., Kidziński, Ł. (2016). Introduction to smart learning analytics: foundations and developments in video-based learning. Smart Learning Environments,3(1).
9. Guerra, J., Hosseini, R., Somyurek, S., Brusilovsky, P. (2016). An Intelligent Interface for Learning Content: Combining an Open Learner Model and Social Comparison to Support Self-Regulated Learning and Engagement. Proc. 21st Int. Conf. Intelligent User Interfaces, 152-163.
10. Guo, P., J., Kim, J., Rubin, R. (2014). How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos. Proc. 1st ACM Conf. Learning at Scale, pp. 41-50.
11. Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. In Proc. the Human Factors and Ergonomics Society annual meeting, 50(9), pp. 904-908, Sage Publications.
12. Kleftodimos, A., Evangelidis, G. (2016). Using open source technologies and open internet resources for building an interactive video based learning environment that supports learning analytics. Smart Learning Environments, 3(1), pp. 1-23.
13. Koedinger, K.R., Kim, J., Jia Z., McLaughlin E., Bier, N. (2015). Learning is Not a Spectator Sport: Doing is Better than Watching for Learning from a MOOC Learning at Scale, Proc. 2nd ACM Conf. Learning @ Scale, pp. 111-120.
14. Kovacs, G. (2016). Effects of in-video quizzes on MOOC lecture viewing, Proc. 3rd Learning @ Scale, pp. 31-40.
15. Kurtz, G., Tsimerman, A., Steiner-Lavi, O. (2014) The Flipped-Classroom Approach: The Answer to Future Learning? European J. of Open, Distance and E-learning, 17(2), pp. 172-182.
16. Lau, L., Mitrovic, A., Weerasinghe, A., Dimitrova, V. (2016) Usability of an Active Video Watching System for Soft Skills Training. Proc. 1st Int. Workshop on Intelligent Mentoring Systems, ITS 2016, Zagreb, Croatia.
17. Long, Y., Aleven, V. (2017). Enhancing learning outcomes through self-regulated learning support with an Open Learner Model. User Modeling and User-Adapted Interaction, pp. 1-34.
18. Mitrovic, A., Dimitrova, V., Weerasinghe, A., Lau, L. Reflexive experiential learning using active video watching for soft skills training. In: Chen, W. et al. (Eds.) Proc. 24th Int. Conf. Computers in Education, pp. 192-201. Mumbai, India, Nov 28 – Dec 2 2016. APSCE.
19. Morgan, G., Adams, J. (2009). Pedagogy first: Making web-technologies work for soft skills development in leadership and management education. Interactive Learning Research, 20(2), pp. 129-155.
20. Pardo, A., Mirriahi, N., Dawson, S., Zhao, Y., Zhao, A., Gašević, D. (2015). Identifying learning strategies associated with active use of video annotation software. In Proc. 5th Int. Conf. Learning Analytics and Knowledge, ACM, pp. 255-259.
21. Pintrich, P. R., De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. Journal of educational psychology, 82(1), 33.
22. Thakker, D., Dimitrova, V., Lau, L., Yang-Turner, F., Despotakis, D. (2013). Assisting user browsing over linked data: requirements elicitation with a user study. Int. Conf. Web Engineering, Springer Berlin Heidelberg, pp. 376-383.
23. Vieira, I., Lopes, A. P., Soares, F. (2014). The potential benefits of using videos in higher education. Proc. EDULEARN14 Conference, IATED Publications, pp. 0750-0756.
24. Wachtler, J., Hubmann, M., Zöhrer, H., Ebner, M. (2016). An analysis of the use and effect of questions in interactive learning-videos. Smart Learning Environments, 3(1), 13.
25. Wang, X., Wen, M., Rosé, C. P. (2016). Towards triggering higher-order thinking behaviors in MOOCs. Proc. 6th Int. Conf. Learning Analytics & Knowledge, ACM, pp. 398-407.
26. World Economic Forum Report (2016) What are the 21st-century skills every student needs? https://www.weforum.org/agenda/2016/03/21st-century-skills-future-jobs-students
27. Yousef, A. M. F., Chatti, M. A., Schroeder, U. (2014). The state of video-based learning: A review and future perspectives. Int. J. Adv. Life Sci, 6(3/4), pp. 122-135.