Developing Pre-service Teachers' Professional Vision with Video Interventions: A Divergent Replication

Adrian Simpson, School of Education, Durham University, UK

Naďa Vondrová, Faculty of Education, Charles University, Czech Republic

Much research on teachers' professional vision examines development through video-interventions. Evidence suggests they change focus towards specifics, subject matter and students and away from evaluation. This study was designed to examine whether there was a discernible difference between viewing pre-service teachers' own videos or general videos and if there were differences between disciplines. However, in marked contrast to existing literature, the video intervention led to no important discernible differences between video conditions and few differences between discipline conditions. We discuss possible reasons for this and demonstrate a positive role of divergent replications such as this in understanding the role of video interventions in the development of professional vision.

Keywords: professional vision; video-intervention; replication study; pre-service teachers

1 Introduction

Classroom are complex social environments: multiple events occur simultaneously which are more or less important for the teachers' goals. Managing classroom behaviour, assessing understanding, attending to individuals (while maintaining oversight of all), evaluating motivation and so on all require teachers' attention, interpretation and response. Levels of teacher expertise likely involves differences in viewing and selecting from these environments: developing requires the acquisition and refinement of this 'professional vision'.

Much literature on professional vision focusses on this development. Work based on the seminal studies of Sherin and van Es has shown consistent shifts in focus

following interventions in which participants watch videos of lessons aimed at improving professional vision. These interventions have overwhelmingly shown shifts with participants' comments moving away from the self, classroom management and generalisations; towards more specific comments on subject matter and learners.

This development appears to involve multiple mechanisms. Noticing (selecting elements of a scene) and interpreting (using professional knowledge to make sense of those elements) appear to be interconnected. One cannot reason about what one has not noticed, but equally, existing knowledge, beliefs and orientations influence where one directs attention and how one makes sense of it (Schoenfeld, 2011). Video appears to have great potential to support those interacting mechanisms.

In some cases, interventions allow teachers to view their own videos (allowing reflection on their own approach) or to see videos of colleagues (allowing the opportunity to discuss intentions and alternatives, e.g. Sherin & van Es, 2009; Mitchell & Marin, 2015). In others, teachers watch videos of others, either exemplar lessons or ones selected by the instructor to encourage attention to salient points (e.g. Santagata & Guarino, 2011; Simpson, Vondrová, & Žalská, 2018). These approaches suggest different development pathways and perhaps different disciplines develop professional vision in subtly different ways. The study reported was designed to explore those questions by examining Czech pre-service teachers' shifts of focus in response to different kinds of video, in the context of their disciplines. In particular, it was intended to focus on one main research question and one subsidiary one. First, does the source of a video (coming from their own teaching, or from other teachers) impact on the development of professional vision amongst pre-service teachers on a video intervention programme? Second, does the development of professional vision vary by subject?

However, while the study was carefully designed to explore these questions, the results are at odds with the consistent message from the existing literature. In the light of concerns throughout science and the social sciences about the role of replications, the paper contributes to our understanding of video interventions by suggesting that the divergent replication (along with that of Star, Lynch & Perova, 2011) indicates limitations for the role of the mechanisms by which video interventions are thought to impact on professional vision.

2 Theoretical framework and related literature

2.1 Noticing/professional vision across different subjects

The balance between the generic and domain-specific in professional vision is not always clear. Some elements of viewing a classroom episode might be shared between teachers irrespective of disciplinary background. However, areas to which one attends and interpretations made when viewing, say, mathematics lessons may be different than viewing history lessons.

There are elements of teaching scenes particular to (or disproportionately present in) some subjects. For example, Talanquer, Tomanek, and Novodvorsky (2013) assessed pre-service science teachers' attention to parts of inquiry based teaching (design, analysis of data, generation of conclusions); Uličná (2017) evaluated future English teachers' noticing and reasoning about phenomena related to teaching English and Vondrová and Žalská (2015) focused on pre-service mathematics teachers' noticing of mathematics specific phenomena.

Research in professional vision is understandably skewed across disciplines: those who developed seminal work in this area have tended to examine it within their own discipline. Thus, there is more work in mathematics and science than other areas.

Blomberg, Stürmer, and Seidel (2011) compared pre-service teachers in mathematics/science with those in social sciences/humanities watching the same six clips, three from each subject. They found evidence for different professional vision across groups indicating that subject-specific socializations may result in distinct sets of shared beliefs and values. Participants demonstrated the social sciences/humanities group had stronger professional vision than the mathematics/science group across all the videos (even on mathematics/science clips). The former group saw clips as more authentic and reported higher self-efficacy about evaluating instruction. The authors concluded that professional vision may be generic which can be applied across subjects but may also be affected by subject-specific socialisation within teacher education.

However, a consistent pattern to the development of professional vision appears to cross subject boundaries: research with pre-service teachers shows they tend to focus more on the teacher than students, to focus more on pedagogy than subject and subject didactics, to evaluate rather than interpret, and to make general claims rather than reference concrete events, compared to experienced teachers (e.g. Sonmez & Hakverdi-can, 2012 [science and technology]; Pavlasová, 2017 [biology]; Uličná, Stará, & Novotná, 2017 [elementary English, social studies and art]).

Perhaps the generic/specific distinction in professional vision reflects generic/specific teaching skills: despite finding positive correlations between them Steffensky et al. (2015) concluded that the professional vision of classroom management (as a generic aspect of teaching) and of learning support (which might require discipline specific pedagogical knowledge) are distinct constructs.

2.2 Video in the development of professional vision

Whether discipline specific or generic, how teachers develop professional vision, particularly whether specific activities lead to further development, is key. Reflection on

observation would be expected to be crucial to this, but impracticalities of live observations, or combining developing vision with the practice of teaching suggest the need to look for alternatives. Video appears to allow a relatively authentic presentation of complex classroom situations (Borko et al., 2008), enables observers' immersion in the situation (Goldman, 2007) permitting them to make links to their teaching and prior knowledge. Videos also provide participants with more control and time, permit repeated viewing and allow sharing and discussion. These advantages have led to a rich literature on 'video clubs' (Sherin & van Es, 2005, 2009; van Es & Sherin, 2010) or video-based interventions for the development of professional vision.

In these interventions, the approach varies with the instructors' learning goals. Two strategies can be distinguished in literature, based on either situated cognition or cognitive learning theories. The former suggests 'learning should be rooted in authentic activity; that learning occurs within a community of individuals engaged in inquiry and practice; that more knowledgeable 'masters' guide or scaffold the learning of novices; and that expertise is often distributed across individuals' (Whitcomb, 2003, p. 538). 'Video thus represents a complex example from which learners can collectively derive principles or rules' (Blomberg et al., 2014, p. 447). This leads to particular patterns of video intervention: e.g. Blomberg et al. encouraged open reflections on videos (used as discussion prompts), information was provided more implicitly (without direct information about the nature of noticing) and cooperative learning came from group working.

Courses based on cognitive learning theory use videos as illustrations of previously taught principles and rules, structured with prompts and guidelines to avoid overloading working memory (e.g. Wouters, Tabbers, & Paas, 2007). For these

interventions, selection, segmentation and manipulation of videos is important to focus attention.

Theoretical positions influence whether explicit frameworks for more structured analysis of videos - implicitly directing attention – are included. There are many such frameworks including:

- Mathematical Quality of Instruction framework (Mitchell & Marin, 2015);
- observation categories (classroom environment, management, tasks, mathematical content and communication) (Star & Strickland, 2008);
- Video Analysis Support Tool, highlighting aspects to be noticed and drawing attention to the need for evidence and interpretation (Sherin & van Es, 2005);
- Lesson Analysis Framework (Santagata et al., 2007; Santagata & Yeh, 2014);
- Mathematically Significant Pedagogical Opportunities to Build on Student Thinking (Stockero, Rupnow, & Pascoe, 2017);
- Science Content Storyline Lens and Student Thinking Lens (Roth et al., 2011).

Despite different theoretical grounding, teaching approaches, scaffolding structures and contexts, studies tend to provide a consistent message: courses using video apparently result in increased attention to the components of and relationships between instruction and student learning.

Not every study shows every element of development: contrast an original study (Star & Strickland, 2008) and its replication (Star et al., 2011). However, there is a strong general tendency across most studies: becoming less descriptive and giving more explicit interpretation (e.g. Santagata et al., 2007); being more focused on students (or teacher-student interaction) than on teachers (e.g. Stockero et al., 2017); showing deeper awareness of aspects of teaching and learning (e.g. Ling Wong et al., 2006; Roth

McDuffie et al., 2014); showing more selective attention and being more detailed and specific (e.g. Santagata & Guarino, 2011; Sonmez & Hakverdi-Can, 2012).

Simpson et al. (2018) analysed studies from different contexts, all using the Sherin and van Es (2009) framework. These demonstrated a clear, consistent shift among participants to focus more on students (rather than teachers) and on subject matter (rather than pedagogy). The only distinction was that some studies showed increasing interpretation (at the expense of description and evaluation) and others showed increasing description (with decreasing evaluation, without increasing interpretation). Increased interpretation appeared associated with the use of teachers' own videos and with using more experienced teacher participants.

2.3 Own vs. public videos

There has been some exploration of this first issue: use of teachers' own videos against use of videos of non-participants (hereafter called 'public videos', they mostly involve generally available videos, rather than ones filmed for the study). However, few studies investigate the effect of viewing one's own versus others' teaching, especially with pre-service teachers.

Seidel et al. (2011) showed that practising science teachers watching videos of their own lessons experienced stronger activation, focused more on teacher guidance and scientific inquiry (rather than other components of teaching) but were less critical, identified fewer consequences and had fewer alternatives than those who watched videos of lessons taught by others.

While not assessing professional vision, Zhang et al. (2011) investigated science teachers' reflections on public videos, teachers' own videos and their colleagues' videos. They considered all three useful for enhancing their reflection on practice, with own videos positively associated with the value of seeing oneself from a distance and

getting feedback from colleagues; peer videos with sharing perspectives and public videos with modelling approaches.

Krammer et al.'s (2015) study is one of the few addressing pre-service teachers. The aim was to develop teachers' goal clarity, teacher support and a positive learning climate. One group worked with videos of their own teaching practice, another group worked with videos of teachers unknown to them and a control group did not work with videos at all (using written materials). The study suggested working with one's own videos had a high degree of acceptance and that its effectiveness received somewhat higher ratings than working with videos of others.

This absence of studies looking directly at the influence of different sources of video on professional vision led us to a research question:

Does the source of videos (coming from their own teaching, or from other teachers) impact on the development of professional vision amongst pre-service teachers on a video intervention programme?

In addition, the practicality of data collection meant that the study was conducted with participants from different disciplines. This allowed us to consider a subsidiary question: Does the development of professional vision vary by subject?

3 Methodology

3.1 Participants

The participants were pre-service teachers from one faculty in the Czech Republic. Two groups were future elementary teachers (educating pupils aged 6 to 12) in their 4th year of a 5 year combined bachelor's and master's degree: one focussed on social sciences and one on art. Two groups consisted of future lower and upper secondary teachers (for pupils aged 12 to 19) in the first semester of a two-year master's degree: one

specialising in mathematics, the other in biology. At the time of taking part in the video intervention, the elementary pre-service teachers had completed general education and psychology courses, some subject and subject didactic courses and teaching practice at elementary schools (not specifically focused on teaching social studies or art). The secondary participants were taking general education and psychology courses alongside subject courses. They had undertaken no teaching practice.

At the beginning of the winter semester, following the relevant ethical approval from the Faculty of Education at Charles University, Prague, all students on these programmes were emailed inviting them to participate. Students take an optional module and the video course was offered as a new option (except for the art students where it was integrated into the main programme).

Participants who chose the new module were randomly assigned to intervention groups ('own' or 'public' video) and students were randomly selected from the students in other modules to form a comparison group. Each group consisted of four students (except for the two intervention art groups and the public video biology group which each had five). There were 17 participants in the 'own video' group, 18 in the 'public video' group and 16 in the comparison group.

3.2 Video interventions

The structure of the video interventions was designed to closely mirror the approach in Simpson, Vondrová and Žalská (2018). It was based on situated cognition learning theory: whole lessons and shorter clips were used, with an emphasis on cooperative learning through in-person and offline opportunities for discussion. A virtual learning environment (VLE) module workshop was used: participants wrote answers to a task first and then were given their peers' answers to comment on.

The video interventions spread across the semester depending on each group's time constraints. Each session consisted of three 45-minute lessons led by the usual course tutor (their subject education teacher). Before each session, participants in the intervention groups were asked to view and comment on whole videos; in each session, this was discussed and shorter clips were used.

Where possible, the interventions for the 'public video' and 'own video' groups were designed to be compatible. Thus, questions in the in-session tasks were general enough to be applicable to all subjects, without precluding opportunities to consider subject specific elements and broad enough to be applicable to both video groups. For example, between most sessions the participants were asked: 'What was the goal of the lesson? What were the pupils to learn? Was the goal fulfilled from the point of view of pupils? Why?', 'Which of the teacher's activities led to the fulfilment of the goal and which not? How differently could the teacher proceed? How would it add to the fulfilment of the goal of the lesson?', 'Write down any other comments. What did you notice? What did you consider interesting and why?'.

The key difference was choice of video: in the 'public video' group, the instructor selected from widely available videos. In the 'own video' group, the instructor chose from the videos of lessons conducted by group participants.

The 'own video' group had a preparatory stage where participants worked with a practicing teacher from a local school to select a topic and plan a lesson which the participants taught with the teacher's class. While the instructor attended these sessions, preparation was undertaken independently. One video was produced by each participant; except for one art group lesson which was taught by two participants. Instructors used these videos in the same way as the 'public video' group, with whole lessons assigned to participants prior to the sessions and clips from the lessons (selected

by instructors after viewing the lessons and reading participants' immediate reflections) used in the sessions. To allow all four videos to be used across three sessions, one session included two videos.

3.3 Measuring Professional Vision Development

Before and after the intervention, participants were asked to watch a public video of a lesson from their subject and write a commentary: 'Write as if you wrote for someone who did not see the lesson and would like to know what it was about and what happened in it. In your comments, deal with issues which you consider important from the point of view of a pupil's learning of content and of a teacher's teaching strategies. You can watch the video several times.' The tasks were completed at home with answers submitted through the VLE. Participants were aware that no module marks were assigned for these tasks.

Two videos from Czech lessons judged by the instructors to have similar levels of general and subject specific phenomena were selected for each subject. To ensure balance, half of the participants in each group were randomly assigned one video at the beginning of the semester and the second at the end. This eliminates effects associated with seeing the same video twice and removes confounding between video choice and intervention effect.

3.4 Analysis of data

Studies in professional vision use different analytic methods. Some use group-level measures (combining individual responses into one, or using a group response, e.g. Stockero, 2008), others use individual measures (e.g. Santagata et al., 2007; Simpson et al., 2018). The former does not allow assessment of within-group variance, resulting in lower statistical power in identifying shift in vision. We adopted the latter approach.

Similarly to Simpson et al. (2018), we used probably the most widely adopted and elaborated framework: that developed by van Es and Sherin (2010). This categorises responses against Actor, Topic, Stance, and Specificity dimensions. Actor can be: Teacher, Student (or students), Curriculum Developer (e.g., a reference to a textbook author), Self (the observer discusses themselves in relation to the video), or Other. Topic can be Classroom Management, Climate (the social environment), Subject, Pedagogy, or Other. Stance can be Describe (a straightforward recounting of an event), Evaluate (a judgment about the event), Explain (accounting for the event, but not supported by theory) and Theorise (explanations including some connection to theory). Specificity can be either Specific (relating to something visible in the video) or General (referring to the whole lesson or some generalisation beyond the videoed events).

Each response was assigned an identifier (by an assistant external to the coding team), so coders were blind to group and stage. It was divided into units of analysis, each representing an observation making sense on its own: this might be a whole sentence or separate phrases (if they contained a shift in attention). Each unit was assigned a code for each dimension using a coding manual detailed in Simpson et al. (2018). This was used by coders (the instructors) to independently code three responses; they then met to agree on codes and interpret the manual. Thereafter, each response was coded by two coders independently, negotiating differences to reach agreement (Table 1).

Table 1: Examples of Coding

Unit of analysis (original/translation)	**Actor**	**Topic**	**Stance**	**Specificity**
Chytřejší žáci tak mohou hlasovat podle jejího výrazu a ne podle svého vlastního úsudku. *More able pupils can thus vote according to [the teacher's] expression and not according to their*	Student	Pedagogy	Explain	General

own consideration.				
Žáci aktivně spolupracují. *The pupils actively cooperate.*	Student	Climate	Evaluate	General
Učitelka se snažila lehce toto téma propojit s jiným předmětem – chemií (chtěla, aby si vzpomněli na krystalizaci). *The teacher wanted to connect the topic with another subject – chemistry (she wanted them to remember crystallization).*	Teacher	Subject	Theorise	Specific

4 Results

4.1 Research question 1

First, we looked for an effect of either group or stage (pre- or post-intervention) on response length, in terms of numbers of units of meaning. Mean response length was 62 units (SD=42), but inspection showed outliers: 4 of the 102 responses exceeded 136 units (1.5 times the interquartile range). The analysis was conducted twice (including and excluding outliers) and there were no substantial differences in outcomes. We therefore report results with outliers included.

We conducted a two-way analysis of variance (ANOVA), with one within-subject variable (stage) and one between-subject variable (group). There was no effect for stage ($F(1,48)=1.58$, $p=0.22$) or group ($F(2,48)=0.08$, $p=0.93$). Figure 1 shows that, if anything, participants wrote slightly less after the intervention.

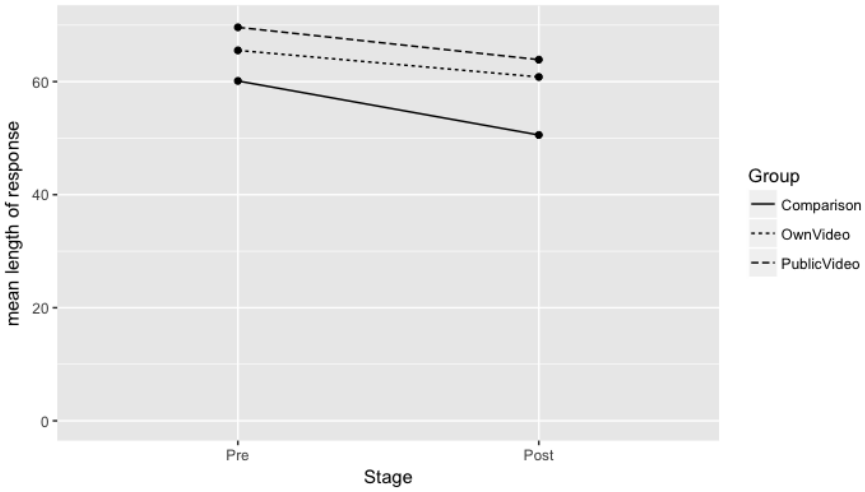

Figure 1: Mean response length of responses

Coding responses under each category results in compositional data: e.g. each unit of meaning coded for Stance can be descriptive, evaluatory, explanatory or theoretical; for each response, the proportion of units of meaning coded against each of these must total 100%. While other studies ignore this consideration, the proportions of response under each heading are not independent, so compositional analysis is important (Aitchison, 1982). Accounting for zeros using the Bayes-Laplace prior approach (Martin-Fernández, Palarea-Albaladejo, & Olea, 2011), data were transformed with an isometric log ratio before conducting a two-way MANOVA (multivariate ANOVA) with transformed proportions as dependent variables, group as a between-subject independent variable and stage as a within-subject independent variable.

For the Actor category, using Pillai's trace, there was no significant effect for group on the composition of the focus on different actors (V=0.01, F(2, 48)=0.16,

p=0.86) or stage (V=0.03, F(1,48)=1.36, p=0.24) (figure 2).

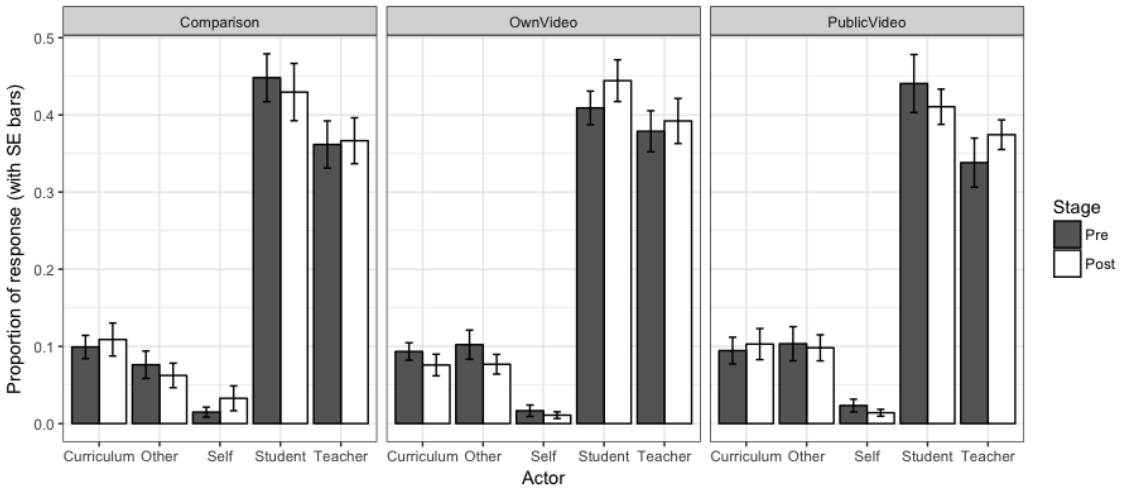

Figure 2: Distribution of Actor in responses

A similar approach was taken for classifying responses for Topic. There was no significant effect for group on the composition of the focus on different topics (V=0.09, F(2, 48)=2.25, p=0.12) or stage (V=0.02, F(1,48)=0.75, p=0.39) (figure 3).

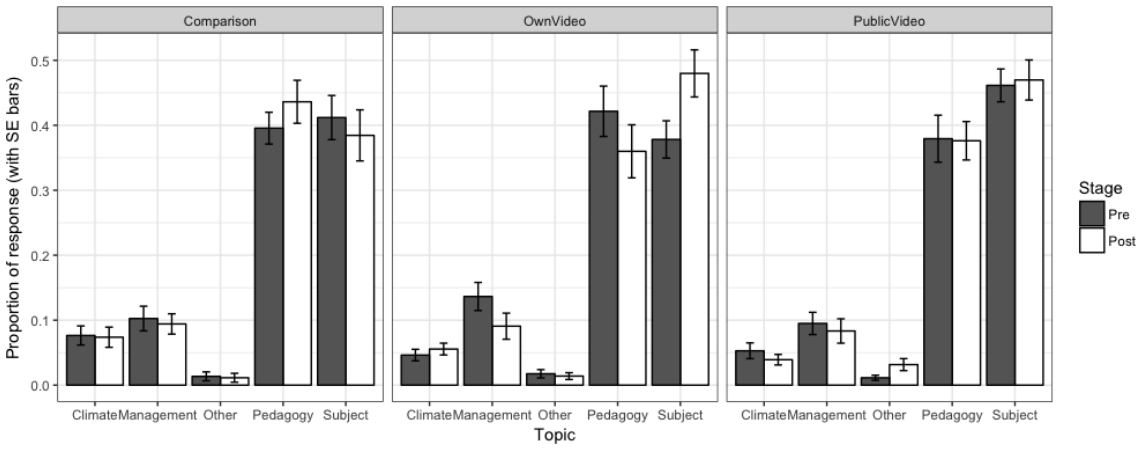

Figure 3: Distribution of Topic of responses

With Stance, there was no significant effect for group on the composition of Stance (V=0.03, F(2, 48)=0.86, p=0.42). However, there was an effect for stage (V=0.13, F(1,48)=7.39, p=0.009). Step-down t-tests for the pre-post intervention

patterns of response found only one significant difference, for the Theoretize category (t(100)=2.14, p=0.04) (figure 4).

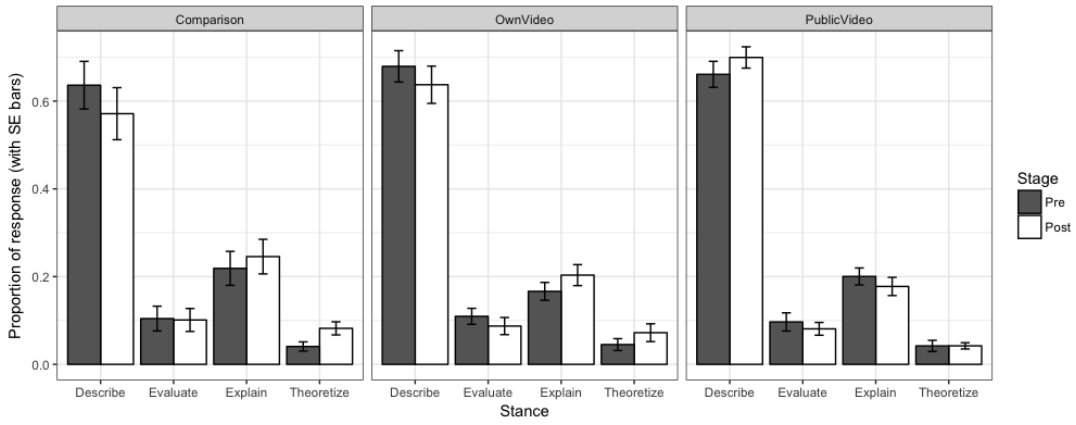

Figure 4: Distribution of Stance of responses

Since the Specificity category has only two values, there is no need to treat data as compositional. We conducted a two-way ANOVA on the proportion of responses classified as General. There was no effect for group (F(2,48)=2.45, p=0.10) or stage (F(1,48)=0.07, p=0.80) (figure 5).

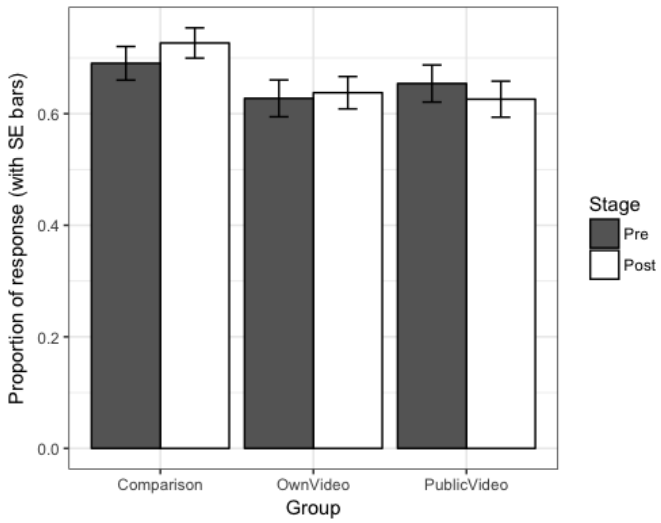

Figure 5: Distribution of Specificity of responses (General)

These results stand in marked contrast to existing literature. Previous studies with the van Es and Sherin framework have consistently showed decreasing focus on the self, increasing focus on students and subject matter with more specific responses. There is some difference in the literature with respect to Stance, most work showing decreasing evaluation with some showing a concomitant increase in description (e.g. Mitchell & Marin, 2015) and others showing an increase in interpretation (e.g. Sherin & van Es, 2009).

This study showed none of this. With the exception of a small increase (from a small base) in theoretizing after the intervention which does not seem to be related to the form of intervention, there were no statistically significant differences, nor in visual inspection, much evidence of even non-significance tendencies in this direction. However, the study was moderately well powered, having a 90% chance of detecting an effect size f=0.4 for the repeated measures, between factors MANOVAs (Faul et al., 2007). Indeed, the study had a relatively large sample compared to much key literature (e.g. Sherin & van Es, 2005; Mitchell & Marin, 2015) and treated data at the individual rather than group level.

4.2 Research question 2

Having noticed this contrast with existing literature, the second research question is moot, but is worth reporting for completeness.

Most literature reports studies with mathematics teachers (or trainees), while our study involved a range of subjects (mathematics, biology, art, social sciences). This led us to question whether there may be differences in responses from participants with different specialisms.

We undertook an analysis by subject. Given there were no large effects for pre-/post- condition and intervention type, we combined responses and conducted a one-way MANOVA by subject (following the procedures above).

For the Actor categories, there was a significant effect of subject on Actor ($V=0.66$, $F(3,12)=3.27$, $p=0.0004$). Step down ANOVAs showed statistically significant differences for Student and Other Actor categories, with Tukey HSD tests showing that biology participants focused less on students than those of other subjects, mathematics participants focused less on Other actors than art or biology participants and social science participants focused less on Others than biology participants (figure 6).

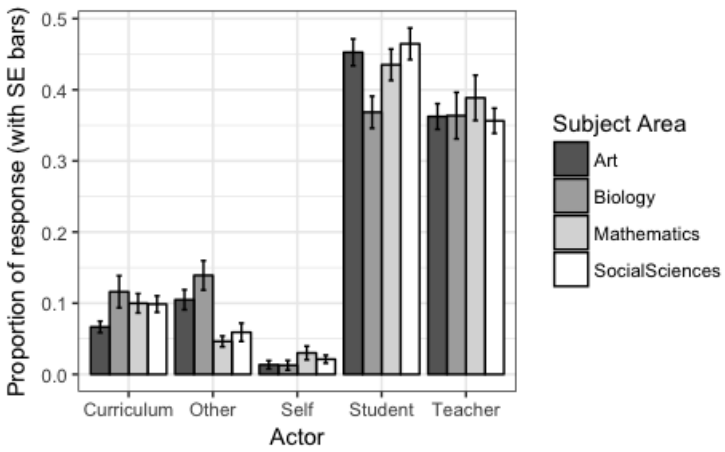

Figure 6: Distribution of Actor responses by subject

For Topic, there was a significant effect of subject on distribution of topics ($V=1.38$, $F(3,12)=9.77$, $p<0.0001$). Step down ANOVAs showed significant differences for Management ($F(3,47)=29.84$, $p<0.0001$), Pedagogy ($F(3,47)=29.84$, $p<0.0001$), Subject ($F(3,47)=9.28$, $p<0.0001$) and Other ($F(3,47)=9.04$, $p<0.0001$), with Tukey HSD tests showing that art participants discussed management and subject topics more (and pedagogy less) than others and biology participants commented on other topics more than others (figure 7).

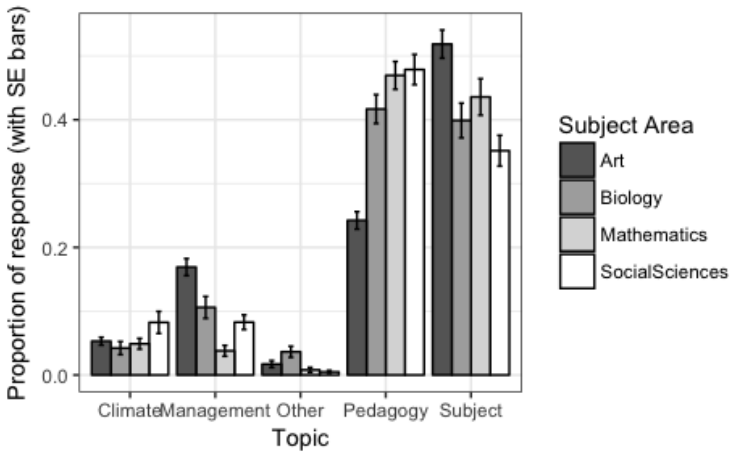

Figure 7: Distribution of Topic responses by subject

There was also a statistically significant effect of subject on Stance (V=0.66, F(3,9)=4.41, p<0.0001) with differences for evaluation (F(3,47)=6.10, p=0.001) and explanation (F(3,47)=3.49, p=0.022) with biology participants being more evaluative than those from art and mathematics and explaining less than biology participants (figure 8).

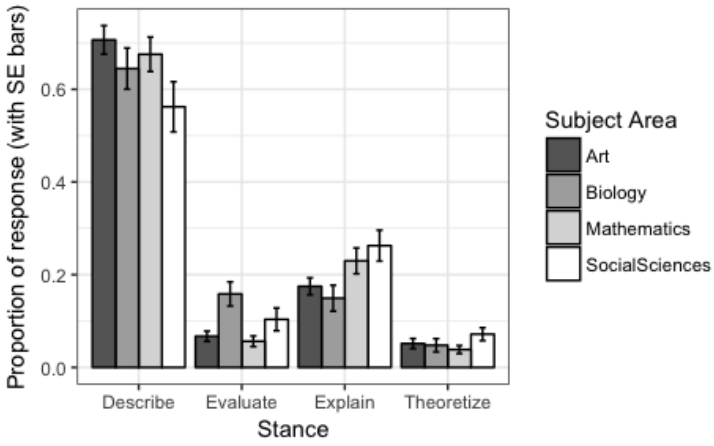

Figure 8: Distribution of stance by Subject

We conducted a one-way ANOVA on the proportion of general statements against subject. There was a statistically significant difference (F(3,47)=7.54, p=0.0003)

with Tukey HSD showing that art participants were less general than those from biology and mathematics (figure 9).

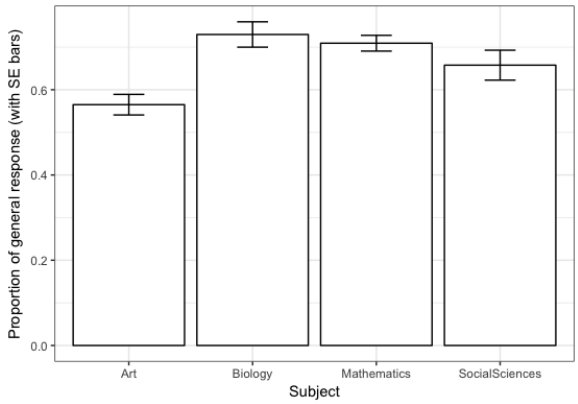

Figure 9: Proportion of general responses by subject

5 Discussion and conclusions

The subsidiary research question asked whether professional vision varied by subject. This picture is mixed: there were apparently some key differences for the elementary pre-service teachers, specializing in art: they tended to be more specific, saying more about the subject and classroom management, less about pedagogy. There were less stark differences with secondary biology participants who were slightly more evaluative and wrote less about students.

These results should be treated with caution. One cannot disaggregate participants' subject from the content of the videos they watched: understandably, art participants watched art lessons, mathematics participants watched mathematics lessons and so on. We cannot rule out that the art videos gave less opportunity to talk about pedagogy than the mathematics videos, rather than that the art participants' professional vision is different from the mathematics participants' vision.

However, RQ2 was a subsidiary question and rendered somewhat moot by the result of our main question concerning whether different sources of the videos ('own' or 'public') was associated with different changes in professional vision.

There were no effects for different intervention types nor, more importantly, for pre- and post-intervention (except the small change in 'theorizing' which applied equally to the comparison group). Thus, before discussing whether the participants' development of professional vision varied by video type, we have to ask whether there was any noticeable development at all. That does not appear to be the case.

This starkly contrasts with much literature motivating this study, which showed consistent movement away from participants writing about themselves and from evaluating; towards writing specifically about subject and students. One might question whether there was a similar tendency here, but one which did not reach statistical significance. This is not supported by the data: not only was the study relatively well powered but examination of figures 2-5 does not suggest a tendency in these directions.

We could thus think of the study as a *divergent replication*, at odds with existing research. It is important to note that this study was not *designed* as a replication; the design assumed that, as previous research showed, the intervention would lead to clear change. The research question was focussed on how that change might be moderated by the type of video. Given the study closely followed the methods of Simpson et al. (2018) (itself grounded on the methods in the literature following the work of Sherin and van Es), explanations for the divergence of these results are unlikely to lie in differences in process or task design. We should consider the consequences of the study as if it were a traditional replication and consider explanations for divergence by considering other reasons for a failure to replicate results.

Replication studies are rare in most fields: Makel and Plucker (2014) suggest only around 1 in 750 education articles is a replication. In studies of the development of professional vision, there appears to be only one: Star et al.'s (2011) replication of Star and Strickland's (2008) study partially diverged from previous findings. While they found similar development of professional vision in relation to classroom environment and communication, in contrast to the original study, they saw no noticeable development in attention to tasks or mathematical content.

Maker and Plucker (2014) caution against dismissing diverging replications; to do so is 'a serious misunderstanding of both science and creativity' (p. 312). Instead, one must consider what the new study's failure to reproduce the expected results might mean, given the context of previous studies.

Schmidt (2009) lists five functions of replication: controlling fraud, controlling sampling error, controlling artefacts, generalising to different populations and assessing hypotheses in existing literature. Given this study sits as an outlier among different studies which consistently showed change after video interventions, one cannot see our divergent replication as an indication of fraud, sampling error or artefacts in those previous studies. However, while we can be sure our own results were not the product of fraud (honestly!), we need to consider whether our study might be the outlier because of sampling error or artefacts, whether it indicates a difference related to generalisation or whether it allows us to raise new hypotheses.

Given that our sample was similar to that of Simpson et al. (2018), it seems unlikely the outcomes reflect sampling error. It seems unlikely they are artefacts of design (such as the task or processes), since the methods followed closely Simpson et al.'s (2018) design (and thus that of van Es & Sherin, 2010, and subsequent literature). The sample came from a population which was somewhat different from the wider

literature, in being a range of different subject participants, in a particular country but given the impact of video interventions has been relatively consistent across countries and across some subjects, it is hard to argue that the divergent results should be attributed to the generalisation to other populations.

This leaves only that this divergent replication indicates an opportunity to generate new hypotheses.

The literature reviewed above suggests a mechanism by which video interventions impact on professional vision, though permitting reflection (with minimal distraction) on careful and controlled viewing of scenes, leading to differentiating important from unimportant aspects of those scenes. However, such a mechanism is unlikely to be universal, but will be affected by the contexts in which it operates. The divergence in this study inevitably leads us to question what about the context might differ from the contexts of previous studies which might plausibly impact on that mechanism.

Star et al. (2011) argue that the divergence of their replication of Star and Strickland (2008) could have two causes. The first is that the original and replication studies used somewhat different assessment methods for judging participants' performance. That is, one cause may be an artefact of different processes involved in generating data. The second, however, is that participants appeared to start from weaker positions (scoring notably lower on attention to tasks and communication than the original participants). That is, the context in which they are trying to develop professional vision is different, so the underlying mechanisms may not work. Star et al. (2011) argue that improvement in some areas of professional development may need a stronger starting position.

Our study was a close operational replication (in the sense of Lykken, 1968) of Simpson et al. (2018), with very similar assessment methods. So, we can reasonably argue that the difference in outcome did not result from a difference in assessment, nor other artefacts of process. Star et al.'s alternative, that there was a difference in the initial state of the participants, is more convincing; albeit for the opposite reason. Rather than having a lower level of professional vision, participants in this study have a consistently *higher* relative level of focus on student thinking, on pedagogy and subject matter and of specific responses, prior to the intervention. Comparing the proportion of statements coded under each heading for this study with Simpson et al. (2018), the participants in this study *start* from the position that those in the Simpson et al. study *finish*. The group from which our sample was drawn may already have relatively well developed professional vision and the mechanism by which this particular video intervention acts to alter professional vision may not be cumulative: it may not be enough to give pre-service teachers more video interventions of the same type, once they have an already established level of professional vision. This fits Stürmer, Seidel, and Holzberer's (2016) question about how high performing students can be supported in their learning trajectories still further.

Cartwright and Hardie (2012) contrast causal factors with support factors: a spark may cause a fire, but only with the presence of combustible material and oxygen. Understanding the mechanism by which video may lead to the development of professional vision also requires us to identify support factors, which means distinguishing factors present when interventions are successful from those present when they are not.

Put together with Star et al. (2011), our study leads to a valuable hypothesis for further study: there may be a 'sweet spot' in which video interventions support the

development of professional vision. Video interventions of this type may fail when participants are underdeveloped or overdeveloped. Star et al. argue that unless people have developed a baseline vision, they cannot build on it. Our study adds to this the suggestion that there may be a limit to how video (at least, as used here) may allow further development of professional vision and, once teachers have achieved it, this type of intervention may no longer be associated with improvement.

Acknowledgements: We are grateful to M. Novotná, L. Pavlasová, J. Robová, J. Stará, and K. Uličná without whose support in teaching the courses and coding the data, this paper could not have been written.

This research has been financially supported by the project PROGRES Q17, "Teacher Preparation and the Teaching Profession in the Context of Science and Research" funded by Charles University.

References

Aitchison, J. (1982). The statistical analysis of compositional data. *Journal of the Royal Statistical Society. Series B (Methodological)*, *44*(2), 139–177.

Blomberg, G., Sherin, M. G., Renkl, A., Glogger, I., & Seidel, T. (2014). Understanding video as a tool for teacher education: investigating instructional strategies to promote reflection. *Instructional Science*, *42*(3), 443–463.

Blomberg, G., Stürmer, K., & Seidel, T. (2011). How pre-service teachers observe teaching on video: Effects of viewers' teaching subjects and the subject of the video. *Teaching and Teacher Education*, *27*(7), 1131–1140.

Borko, H., Jacobs, J., Eiteljorg, E., & Pittman, M. E. (2008). Video as a tool for fostering productive discussions in mathematics professional development. *Teaching and Teacher Education*, *24*(2), 417–436.

Cartwright, N. & Hardie, J. (2012). *Evidence-based policy: A practical guide to doing it better*. Oxford, UK: Oxford University Press.

Faul, F., E. Erdfelder, A.G. Lang, & Buchner, A. (2007). G∗Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Goldman, R. (2007). Video representations and the perspectivity framework. In R. Goldman, R. Pea, B. Barron, & S. J. Derry (Eds.), *Video research in the learning sciences* (pp. 3–37). London, UK: Lawrence Erlbaum Associates.

Krammer, K., Hugener, I., Frommelt, M., Fürrer Auf der Maur, G., & Biaggi, S. (2015). Case-based learning in initial teacher education: Assessing the benefits and challenges of working with student videos and other teachers' videos. *Orbis Scholae*, *9*(2), 119–137.

Ling Wong, S., Wai Yung, B. H., Cheng, M. W., Lam, K. L., & Hodson, D. (2006). Setting the stage for developing pre-service teachers' conceptions of good science teaching: The role of classroom videos. *International Journal of Science Education, 28*(1), 1–24.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin, 70*(3), 151–159.

Makel, M. C. & Plucker, J. A. (2014). Facts are more important than novelty: Replication in the education sciences. *Educational Researcher*, *43*(6), 304–316.

Martin-Fernández, J. A., Palarea-Albaladejo, J., & Olea, R.A. (2011). Dealing with Zeroes. In V. Pawlowsky-Glahn & A. Buccianti (Eds.), *Compositional Data Analysis*: *Theory and Applications* (pp. 43–58). New York: John Wiley and Sons.

Mitchell, R. N. & Marin, K. A. (2015). Examining the use of a structured analysis framework to support prospective teacher noticing. *Journal of Mathematics Teacher Education*, *18*(6), 551–575.

Pavlasová, L. (2017). Profesní vidění studentů učitelství biologie zaměřené na obor a oborově didaktické jevy. [Professional vision of biology pre-service teachers with a focus on subject and subject didactic phenomena.] *Scientia in educatione, 8*(2), 84–99.

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., & Wickler, N. I. Z. (2011). Video based lesson analysis: Effective science PD for teacher and student learning. *Journal of Research in Science Teaching*, *48*(2), 117–148.

Roth McDuffie, A., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., Drake, C., & Land, T. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *Journal of Mathematics Teacher Education*, *17*(3), 245–270.

Santagata, R. & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM*, *43*(1), 133–145.

Santagata, R. & Yeh, C. (2014). Learning to teach mathematics and to analyze teaching effectiveness: evidence from a video- and practice-based approach. *Journal of Mathematics Teacher Education*, *17*(6), 491–514.

Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video based field experience. *Journal of Mathematics Teacher Education*, *10*(2), 123–140.

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90-100.

Schoenfeld, A. H. (2011). Noticing matters. A lot. Now what? In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics Teacher Noticing: Seeing Through Teachers' Eyes* (pp. 223–238). New York and London: Taylor & Francis.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, *27*(2), 259–267.

Sherin, M. G. & van Es, E. A. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of Technology and Teacher Education*, *13*(3), 475–491.

Sherin, M.G. & van Es, E.A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education*, *60*(1), 20–37.

Simpson, A., Vondrová, N., & Žalská, J. (2018). Sources of shifts in pre-service teachers' patterns of attention: the roles of teaching experience and of observational experience. *Journal of Mathematics Teacher Education*, *21*(6), 607-630. doi:10.1007/s10857-017-9370-6

Sonmez, D. & Hakverdi-Can, M. (2012). Videos as an instructional tool in pre-service science teacher education. *Egitim Arastirmalari-Eurasian Journal of Educational Research*, *46*, 141–158.

Star, J.R., Lynch, K., & Perova, N. (2011). Using video to improve preservice mathematics teachers' abilities to attend to classroom features: A replication study. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics*

Teacher Noticing: Seeing Through Teachers' Eyes (pp. 117–133). New York and London, Taylor & Francis.

Star, J. & Strickland, S. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*(2), 107–125.

Steffensky, M., Gold, B., Holdynski, M., & Möller, K. (2015). Professional vision of classroom management and learning support in science classrooms—Does professional vision differ across general and content-specific classroom interactions? *International Journal of Science and Mathematics Education, 13*(2), 351–368.

Stockero, S. L. (2008). Using a video-based curriculum to develop a reflective stance in prospective mathematics teachers. *Journal of Mathematics Teacher Education, 11*(5), 373–394.

Stockero, S. L., Rupnow, R. L., & Pascoe, A. E. (2017). Learning to notice important student mathematical thinking in complex classroom interactions. *Teaching and Teacher Education, 63*, 384–395.

Stürmer, K, Seidel, T., & Holzberger, D. (2016). Intra-individual differences in developing professional vision: preservice teachers' changes in the course of an innovative teacher education program. *Instructional Science, 44*(3), 293–309.

Talanquer, V., Tomanek, D., & Novodvorsky, I. (2013). Assessing students' understanding of inquiry: What do prospective science teachers notice? *Journal of Research in Science Teaching, 50*(2), 189–208.

Uličná, K. (2017). Professional vision of future English language teachers: subject-specific noticing and knowledge-based reasoning. *E-pedagogium*, II, 38–49.

Uličná, K., Stará, J., & Novotná, M. (2017). Teacher in the eyes of future primary school teachers. In M. Houška et al. (Eds.), *Efficiency and Resonsibility in Education, 14th International Conference* (pp. 490–497). Prague, CZ: Czech University of Life Sciences.

van Es, E. A. & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education, 13*(2), 155–176.

Vondrová, N. & Žalská, J. (2015). Ability to notice mathematics specific phenomena: what exactly do student teachers attend to? *Orbis scholae, 9*(2), 77–101.

Whitcomb, J. (2003). Learning and pedagogy in initial teacher preparation. In W. M. Reynolds & G. E. Miller (Eds.), & I. B. Weiner (Editor in chief), *Handbook of psychology, Volume 7: Educational psychology* (pp. 533–556). New York: John Wiley.

Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review, 19*(3), 327–342.

Zhang, M., Lundeberg, M., Koehler, M. J., & Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education, 27*(1), 454−462.