UNIVERSITY OF BIRMINGHAM

Using video clubs to develop teachers' thinking and practice in oral feedback and dialogic teaching

Perry, Tom; Davies, Peter; Brady, Josie

Link to publication on Research at Birmingham portal

Using video clubs to develop teachers' thinking and practice in oral feedback and dialogic teaching

Thomas Perry, Peter Davies and Josephine Brady

School of Education, University of Birmingham, Birmingham, UK

Abstract

We report the outcomes of an evaluation of a 'video club' intervention to improve the feedback and dialogic teaching practice of 91 teachers from 11 primary schools in England. Participating teachers worked collaboratively in a sequence of six video clubs over a six-month period. To understand teacher engagement we examine videos of video club meetings; online platform use metrics; surveys; selected videos of classroom practice; focus groups and interviews. We evaluate change in teachers' thinking and practice using survey results for participants compared to a comparison group of non-participating teachers at the intervention schools. The survey includes a new instrument for gathering evidence of teachers' thinking and practice in feedback. The results suggested changes in thinking and practice for teachers who self-reported as engaging highly with the intervention. We conclude by discussing the potential of video technology within professional development and the challenges of researching changes in thinking and practice.

Key Words: video clubs; professional development; formative assessment; feedback; reflection; classroom dialogue

1.   Introduction

This paper reports findings from an independent evaluation of an intervention using video technology to develop teachers' thinking and practice. Primary school teachers collaborated to review lessons with a focus on improving classroom dialogue and teachers' verbal feedback. Collaboration was organised through six 'video-clubs' over a six-month period in which participating teachers reviewed lesson videos from other schools and from their own teaching. They were supported by reading material and activities on an online platform. There has been considerable international interest in the scope for improving the quality of teaching through reviewing lessons with video technology. This paper contributes to the field by providing (i) a model of how video technology can be used to change teachers' thinking and practice in relation to feedback; (ii) quantitative evidence of impact from a newly-developed instrument which captures teacher thinking about classroom feedback; (iii) evidence that can inform the detailed design of video club interventions.

We start the paper by reviewing two areas of literature which informed the programme and its evaluation. Section 2 reviews the use of video within continuing professional development (CPD) programmes and the evidence for its effectiveness. Section 3 reviews the

literature on effective teacher feedback and classroom dialogue. We then provide details of the CPD programme and our evaluation aims and methods. The results section presents evidence of the extent of teachers' engagement with the intervention and evidence of impact on teachers' thinking and practice. The results suggested changes in thinking and practice for teachers who self-reported as engaging highly with the intervention. We close the paper by discussing the potential of video technology within CPD programmes, the use of measurement instruments to form and evaluate teacher development, and the difficulties of 'getting inside the black box' of professional development programmes to understand in detail the changes in thinking and practice required to bring about improvement in pupil learning.

2.   Video technology in collaborative professional development

As school effects on student achievement are largely attributable to the quality of individual teachers and teaching (Slater *et al.*, 2012; Kane *et al.*, 2013), teacher professional development is arguably the most direct and effective school improvement approach. Video technology has been proposed as a means of enabling and enhancing professional development activities. This review concentrates on using video technology for collaborative professional development (e.g Borko *et al.*, 2008), ignoring alternatives such as coaching which may be provided 'in-ear' during teaching (Quinn *et al.*, 2018; Kane *et al.*, 2015).

2.1   Arguments for video technology

Video technology shows promise for enabling and enhancing professional development for a number of reasons. First, video is able to vividly capture classroom interactions, allowing them to be carefully reviewed outside of the classroom (Tripp and Rich, 2012; Pehmer *et al.*, 2015; Brouwer *et al.*, 2017, p.61). Second, this ability to capture and revisit classroom practice lends itself to collaborative approaches to professional development: it becomes feasible for many colleagues and/or external experts to offer different perspectives on videos and jointly learn (Brouwer *et al.*, 2017). Group discussions of video evidence may help teachers to see things they would otherwise miss, creating space for conceptions to change (Gaudin and Chaliès, 2015). This idea has featured prominently in the literatures on 'Lesson Study' (e.g. Lewis *et al.*, 2006) and 'Learning Study' (Holmqvist, 2010; Davies and Dunnill, 2008). Third, online video platforms provide opportunities for teachers to prepare for discussions and for colleagues to observe the lessons of others at their own convenience and outside of teaching hours, avoiding supply cover costs (Quinn *et al.*, 2018). Repeat viewing of videos allows for greater efficiencies still, as particularly instructive videos can be viewed by a potentially unlimited number of teachers far removed from the original lesson. Finally, use of classroom videos through video platforms typically allows for additional functionality through the use of video tools, such as editing or annotation tools, which support teacher reflection, discussion and analysis (Rich and Hannafin, 2008). Support for these arguments is provided by reviews of using video technology for teacher development (e.g. Brouwer, 2011; Gaudin and Chaliès, 2015; Major and Watson, 2018).

2.2   Video club design (1): learning communities

Video clubs aim to exploit these possibilities by developing sustained learning communities in which teachers trust each other sufficiently to share and discuss their own visible practice and its impact on learners (Cordingley *et al.*, 2015; Darling-Hammond *et al.*, 2017; Major and Watson, 2018). Researchers using video technology to support teacher learning communities (e.g. Borko *et al.*, 2008; Sherin and van Es, 2008; van Es, 2012; Gröschner *et al.*, 2014; Vangrieken *et al.*, 2017; Alles *et al.*, 2019) have suggested several features which will make it more likely that teachers' thinking and practice will be developed: having shared discourse rules and trust; focused discussion of, and reference to, practices observed in the shared videos; sufficient contextual information for the video excerpts; explicit use of variation (alternative interpretations of pupils' responses and teaching strategies); active involvement and contribution by all group members; and supportive leadership. However, a culture of trust and mutual respect can require a significant amount of time to develop (van Es, 2012; Beisiegel *et al.*, 2018; Alles *et al.*, 2019). Video clubs aim to address this through a cyclical approach which develops teachers' confidence in the usefulness of the technology as well as their confidence in open dialogue with their peers.

Researchers have also pointed to the importance of guidance for groups of teachers when reviewing videos. This may be provided through an expert facilitator who promotes focused and productive discussion (Sherin and Han, 2004; van Es, 2012; Gaudin and Chaliès, 2015), However, the evidence here is not clear-cut and external facilitators increase the cost of interventions (Beisiegel *et al.*, 2018). Alternatively, teachers may be guided through written scripts which direct the focus and the process of discussions (Zhang *et al.*, 2011). Brouwer *et al.* (2017) reported a positive impact (compared to a control group) on targeted teaching behaviours when teachers used structured viewing guides in their review of videos.

These arguments suggest some questions for the evaluation of video club interventions: First, how does variation in the culture in participating schools predispose teachers towards effective use of the technology for open discussion? Second, does the length of the intervention give teachers' sufficient time to develop confidence in exploiting the opportunities created by the video technology? Third, how is the teachers' review of videos guided?

2.3   Video club design (2): choice of videos

In a video club initiative, teachers could review videos of their own teaching, videos of other teachers in their school or videos of teachers they do not know. Teachers may develop a new awareness of their teaching when they see themselves through the eye of a camera. This may be seen as having the benefit of authenticity, but it is not self-evident that teachers will see what they do in a new light simply as a consequence of viewing it through this medium. In particular, studies of dialogic teaching have found that teachers are often *unaware* of the teacher-student talk patterns and functions in their classrooms and their functions (Mercer *et al.*, 2009). Van Es (2012) and Borko *et al.* (2008) have argued that teachers will be more open to new ideas when viewing the work of others. Kleinknecht and Schneider (2013) reported greater teacher engagement in analysing 'problematic events' when reviewing videos of other teachers.

Nonetheless, two systematic comparisons of different combinations of teachers reviewing videos of their own teaching and the teaching of others (Seidel *et al.*, 2011; Beisiegel *et al.*, 2018) have reported greater benefits for a sequence of first watching selected videos of others and then reviewing their own teaching. Moreover, according to Sherin and van Es (2008), the benefits of using videos arrive when teachers develop their 'professional vision' through which they identify 'significant features' of teaching and learning. This has to be learned, rather than assumed (Brouwer *et al.*, 2017, p.61; Gaudin and Chaliès, 2015).

The literature suggests several features of video clubs that may affect the development of 'professional vision'. First, there is some evidence (Bates *et al.*, 2016) that teachers are less likely to develop this capacity for searching and critical scrutiny when they are simply given free access to a bank of videos from which to select what they will view. Second, video clips which focus attention on a particular issue seem to be more useful than longer videos which may raise a host of issues (Sherin *et al.*, 2009; Bates *et al.*, 2016). Third, there is a choice between focusing on implementation of a teaching strategy, focusing on what students learn or focusing on the relationship between a teaching strategy and what students learn. Videos are particularly adept at exemplifying teaching strategies and this has prompted calls to use video clips to demonstrate 'effective practice' (Brouwer *et al.*, 2017). However, it has also been argued that the essence of 'professional vision' lies in understanding (rather than taking for granted), relationships between teachers' strategies and *what* students learn. This approach is adopted in uses of video clips in professional development based on variation theory (e.g. Holmqvist, 2010). The idea here is that teachers need to be able to observe differences between what students learn that are associated with differences between the conduct of teaching.

3.   Teachers' thinking and practice in dialogic teaching and verbal feedback

Although the focus of this intervention was on teachers' verbal feedback, this has to be placed in the context of classroom dialogue. The possibilities for teachers' verbal feedback (e.g. in scaffolding pupils' ideas) depend on the patterns of discourse and participation rights that teachers manage in their classrooms (Walshaw and Anthony, 2008; Kiemer *et al.*, 2015). A contrast has frequently been drawn between an 'initiation-response-feedback' (IRF) pattern of discourse and 'dialogic' teaching (Howe and Abedin, 2013; Mercer and Dawes, 2014). An example of an IRF interaction would be a teacher initiating an exchange with a question, a student response, and then teacher feedback on the student response. IRF discourse is typically controlled by the teacher, encouraging teacher feedback to take the form of a judgement on the pupil's response. Nevertheless, the 'F-move' in this format can be characterised by different forms of feedback: a simple judgement of correctness; identification of a part of an answer that would be good to build on; or a suggestion about how to build a more complete understanding, for examples (Smith and Higgins, 2006). Feedback can be used to both close or sustain classroom dialogue and can be understood as a dialogic 'move' in terms of the elements from within Hennessy et al.'s (2016) *Scheme for Educational Dialogue Analysis (SEDA)*. Providing informative feedback (G4), referring back to prior contributions (C1), and inviting opinions, beliefs or ideas (E1) (Hennessy et al., 2016, Table 1) are dialogic moves which are or can replace feedback after a student contribution to classroom discussion. Dialogic teaching cedes

more control of the discourse pattern to pupils and encourages teachers to prompt pupils to extend and deepen their thinking, allowing pupils to build on others' ideas as the teacher models reflection and interest in alternatives (Alexander, 2008; Wiliam, 2011; Edwards-Groves *et al.*, 2014; Gillies, 2016; Khong *et al.*, 2019). Jay *et al.* (2017) reported results from an RCT of an intensive dialogic teaching intervention funded by the Education Endowment Fund (EEF) in England. The intervention, described and discussed at greater length in Alexander (2018), trained teachers in repertoires for dialogic teaching which promote pupils' ability to articulate and elaborate on their thoughts, examine and explain ideas, and participate within classroom discussion. The repertoires are situated within a framework justifying, outlining principles, and providing indicators for dialogic teaching. Using standardised assessment in English, Science and mathematics, Jay *et al.* (2017) found that 9-10 year-old children made an additional 2 months of progress in English and Science and an additional 1 month progress in mathematics. The intervention included videoing and reviewing lessons, but without an online platform.

In contrast, the EEF (2018) claims that interventions targeting teachers' feedback practices raise pupil attainment by an average of 8 months. This assertion rests heavily on Hattie and Timperley's (2007) extensive review of relationships between types of feedback and pupil attainment. They suggested four types of feedback in order from most to least effective in improving attainment: self-regulation; processing of a task; the task; the self. The huge scope of this study meant that it included evidence from a great range of settings including those far removed from school classrooms. Moreover, the protocol of many of the randomised-controlled tests (RCTs) included in the review relied on the control condition offering no scope for the type of feedback targeted in the intervention (Davies, 2018). Nonetheless, Hattie & Timperley's review has provided an important benchmark for subsequent research (e.g. Brown *et al.*, 2012).

The typology poses several questions for teachers' professional development. What is the relationship between teachers' conceptions of feedback and the hierarchy of impact on attainment suggested by the typology? How are teachers' conceptions of feedback related to their practice? How is teachers' feedback practice related to their management of classroom discourse? More specifically, for this study, can the use of video technology enhance professional development to change teachers' conception of feedback and/or subsequent practice? A small group of studies have started the task of building a body of evidence that addresses these final two questions. Kiemer *et al.* (2015) found that, compared with a traditional form of professional development, a video-based intervention improved students' sense of autonomy and motivation to learn. This outcome is very encouraging from the perspective of Hattie and Timperley's typology. Several other small-scale projects have also reported positive impacts of video clubs on teachers' thinking and practice about feedback (Van den Bergh *et al.*, 2014; Gröschner *et al.*, 2015; Schindler *et al.*, 2016).

4. Method

4.1   Research questions and research design
This study examines two questions in the light of previous research in this field:

(1) How do teachers participating in a video club intervention change their thinking and practice in oral feedback?
(2) Does evidence of the process of the intervention (how it worked in practice) support attributing change to participation in a video club?
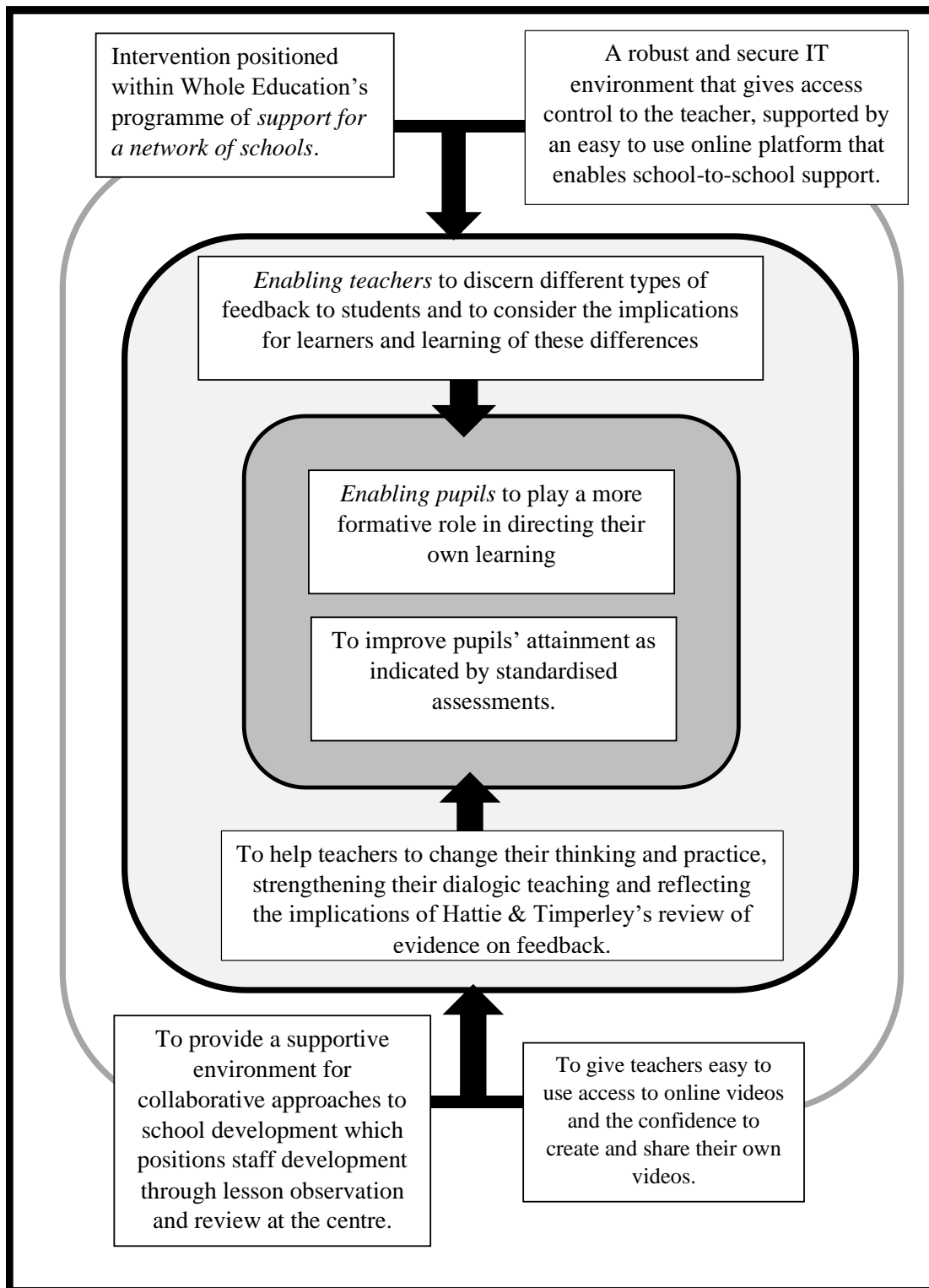
The study provides evidence of change in primary school teachers' thinking through a 'difference-in-difference' analysis comparing participating teachers with other teachers in the same schools who did not participate. During the course of any intervention that lasts for several months there are other factors in play. Schools may be involved in other initiatives and short-term changes (such as preparation for a school inspection) which affect teachers' focus and practice. Comparing teachers participating in an innovation with those who do not provides a way of controlling for these effects. Pre- and post- surveys provided the data for this comparison. This design adds to previous evidence (Van den Bergh *et al.*, 2014; Gröschner *et al.*, 2015; Schindler *et al.*, 2016) that did not measure the difference in outcomes between an intervention group and a comparison group. This study provides evidence of change in teachers' practice through analysis of video taped lessons. Evidence of teachers' engagement is provided through data gathered through the online platform that was used for the intervention and self-reports from teachers. Teachers were able to access videos through an online platform which recorded the number of times that videos were viewed and uploaded. Ethical approval was granted by the university ethics committee (ERN_15-0987A). In the following sub-sections we describe and justify: the design of the intervention; the sample of schools and teachers; and methods of data collection and analysis.

4.2 The design of the intervention

The 'video club' intervention was delivered through a partnership between IRIS Connect and Whole Education. IRIS Connect provides video equipment, an online platform and associated professional support for schools that buy a three-year subscription to their system. At the time of the evaluation, the proportion of state primary schools in England using IRIS Connect (=930, c. 5%) was considerably smaller than the proportion of state secondary schools (=723, c. 20%) using the system[i]. Whole Education is a not-for-profit organisation that supports a large network of schools to improve through professional development. Together, both organizations aim to support school improvement through collaborative professional development using video technology. The providers' strategy was influenced by their reading of Joyce and Showers (2002); the theory of change underpinning the intervention is presented in Figure 1. Participating schools received a publication outlining the IRIS Connect approach to professional development (IRIS Connect, 2014). The video clubs were designed to develop teachers' thinking and practice in relation to classroom talk and higher order feedback by 'making learning visible'. Each participating school appointed an intervention co-ordinator known as the school 'champion' who received training provided by the central team. Each school champion convened and led the programme of video clubs in their school. Videos of classroom interactions provided the stimulus for each meeting and the central team provided champions with guidance for each meeting (see Appendix 1 for a copy of this guidance). Each participating teacher had access to the IRIS-Connect online platform giving them access to

videos, discussion and guidance on videoing lessons. The detailed implementation of the intervention was placed in the hands of the schools. We used a wide range of data to check whether this range fell within acceptable limits to allow us to view the intervention as a whole, adhering to the specification in Appendix 1. This is the focus of the second research question.

Figure 1 - Theory of Change

| Intervention positioned within Whole Education's programme of *support for a network of schools*. | A robust and secure IT environment that gives access control to the teacher, supported by an easy to use online platform that enables school-to-school support. |

Enabling teachers to discern different types of feedback to students and to consider the implications for learners and learning of these differences

Enabling pupils to play a more formative role in directing their own learning

To improve pupils' attainment as indicated by standardised assessments.

To help teachers to change their thinking and practice, strengthening their dialogic teaching and reflecting the implications of Hattie & Timperley's review of evidence on feedback.

To provide a supportive environment for collaborative approaches to school development which positions staff development through lesson observation and review at the centre.

To give teachers easy to use access to online videos and the confidence to create and share their own videos.

The design of the intervention and data collection methods were refined as a result of experience in trialling the intervention with 12 primary schools during a "development phase". The main intervention reported in this paper lasted six months from January to July 2016. During this period, participating primary schools were asked to organise six video clubs on 'teacher feedback and dialogic teaching'. Video clubs one to three, between January and March, focused on classroom talk and teaching (video club one), questioning and group talk (video club two), and feedback (video club three). The support provided through the online platform included Hattie and Timperley's categorisation of different types of feedback. Teachers were asked to prepare for video clubs four and five by selecting and editing video clips of their own lessons which they believed would be useful for discussing the ways in which they were using dialogue and feedback in their teaching. The online guidance provided guidance to teachers on using the technology to do this.

In the final video club, teachers were asked to evaluate their experience and to discuss what they thought they had learned and how their practice had been influenced. Guidance provided to school champions asked for these reflections to be summarised using the online system. Video clubs typically lasted about an hour and took place at the end of the teaching day in time reserved for meetings, preparation and CPD. Further details on the guidance and structure provided for six video clubs are included in Appendix 1.

The intervention relied on designated 'school champions' rather than external experts to co-ordinate and lead the sessions. The school champion was responsible for leading and co-ordinating the use of the video technology in schools and for liaising with senior management to align this practice with school improvement objectives. They were also the key point of contact between the school, IRIS Connect, and Whole Education. Two project meetings were held with the school champions to clarify expectations and address their questions about the implementation.

4.3 Sample of schools and teachers

The target was to recruit 12 primary schools to participate in the intervention. Six schools which had participated in the development phase agreed to take part in the pilot. Six relatively new schools to IRIS Connect were also recruited, bringing the initial total to 12. The sampling aimed to recruit one-form as well as two-form entry schools to check the feasibility of the intervention in small as well as medium-sized schools. School details are presented in Table 1. School 12 dropped out when the school champion left the school, leaving 11 schools in the sample. School 7 received notification of an OfSTED school inspection shortly after the start of the pilot. They were subsequently graded 'Inadequate' but chose to remain in the pilot as part of their response to inspectors' recommendations.

Schools were given freedom regarding the number of teachers and the process which led to these teachers participating in the project. Three schools (4, 8 and 10) required all teachers to participate. In four schools, teachers were asked to volunteer. In four schools the head teacher and school champion identified participants. Schools were encouraged to concentrate on Year 5 (age 9 to 10) classes where possible. As primary schools in England usually expect teachers to take responsibility for a range of subjects with their class we did not collect data on subject specialisms.

Table 1 Characteristics of participating schools

School	First use[1]	In Dev. Phase[2]	Age range	School roll	School type[3]	% FSM [4]	% EAL [5]	% KS2 Level 4[6]	No. of teachers Total	No. of teachers Partici-pating
1	Sep11	Yes	7-11	389	Com	5	4	85	20	4
2	Feb14	Yes	5-11	610	Com	39	95	71	33	4
3	Apr14	Yes	3-11	456	Com	31	43	88	25	7
4	Jun14	No	2-11	439	Com	47	6	79	24	19
5	Jan15	No	4-11	303	F	8	9	77	13	4
6	Jul15	No	4-11	503	Com	20	12	82	21	5
7	Jul15	No	7-11	355	Com	17	2	82	18	9
8	Sep15	No	4-11	391	Com	26	5	75	21	17
9	Oct15	No	3-11	212	Ac	39	5	86	18	3
10	Jan16	Yes	5-11	422	Ac	6	1	88	18	16
11	Jan16	Yes	4-11	241	Com	18	2	59	12	4
12	n/a	Yes	3-7	252	Com	15	3	15	13	none

[1] Date school first used the IRIS Connect technology
[2] School was involved in the development phase prior to the intervention
[3] School governance type (Ac=Academy, Com = State maintained community, F = state mainated foundation)
[4] Percentage of children eligible for free school meals in last six years
[5] Percentage of children achieving at least level 4 (expected standard) in reading, writing and maths age 11
[6] Percentage of children with English not as first language

4.4 Data collection and Analysis

Our analyses drew on several data sources, described in this section (4.4.1 to 4.4.4).

4.4.1 *Pre- and post-intervention survey evidence of teachers' thinking*

All teachers in each school were invited to complete a questionnaire administered in January and again in July. Completed surveys on both occasions were received from 68% (n= 63) of teachers reported as participating in the video clubs and from 25% (n=32) of teachers reported as not participating. Almost 100% of participating teachers and 61% of non-participating teachers completed at least one of the surveys, respectively. We infer that missing data from participating teachers did not reflect lack of engagement with the intervention. See below for further discussion of missing data and engagement.

Our instrument for gathering evidence of teachers' thinking had 7 items, each of which was designed to capture a difference in the power of feedback according to Hattie and Timperley's (2007) review. The format of the instrument presented pairs of contrasting statements. Each statement in the pair was designed to offer a plausible justification from a particular perspective on teaching. For example, the first pair was 'provides the child with a simple next step they can easily achieve' and 'presents the child with a serious challenge'. Teachers were invited to indicate on an 11 point scale whether their feedback was more like the first or the second statement. The 'more powerful feedback' statement was sometimes

placed first in the pair and sometimes second. The full set of statements is provided in Table 5 in the results section. The table indicates in bold which of each pair of statements articulates a more powerful type of oral feedback according to Hattie and Timperley. Our categorisation of each statement in terms of Hattie and Timperley's typology is shown in italics in brackets after each statement. Our analysis of these items focused on whether there was a difference between the pre- to post- survey change in responses by teachers participating in the intervention compared with the change in responses by teachers not participating in the intervention. The post-intervention survey also included 9 statements to elicit teachers' beliefs about the efficacy of the intervention (see Table 3). Analysis of these items informed our judgement about teachers' engagement with the intervention.

4.4.2 *Evidence of teachers' practice from shared lesson videos*

We gathered some indicative evidence of teachers' practice through video clips of lessons that they chose to share. We interpreted teachers' selections as indicating what they considered to be, at least, acceptable practice. The IRIS Connect online platform enabled teachers to share videos with the research team. Teachers in 8 of the 11 schools chose to share a total of 45 video clips and these included any comments that teachers had added and any analysis they had conducted using the 'forms' facility on the online system. Shared video clips lasted from about 3 to 30 minutes with an average length of 11 minutes. Teachers typically chose to share either a single activity or a teacher-led discussion from a lesson.

Lesson clips were scored using the same scale of 7 items on teacher feedback used in the survey of teachers. Although no formal reliability testing of the instrument as an observation tool was undertaken, a selection of the clips were viewed independently by two of the researchers when testing the use of the scale for observation. The researchers then met to review and discuss the videos and whether the scale and items were suitable for recording the feedback practice in the clips. All clips were then coded by a single researcher. These data need to be treated with some caution: First, we were not able to compare practice in these shared videos with teachers' practice before the intervention. Second, the terms on which schools participated in the project (and a fundamental principle of the IRIS-Connect system) is that teachers choose whether, and with whom, to share a video. It is likely that their selection was somewhat affected by 'social desirability bias' (Krumpal, 2013) as they became familiar with the intentions of the intervention. Moreover, giving teachers control over which videos they chose to share resulted in videos of differing lengths (see above), subjects and lesson contexts (e.g. small group work versus whole-class sessions). Finally, the videos were shared during the course of the project and, therefore, cannot be treated as evidence of teachers' practice at the end of project. These limitations aside, the lesson videos were a valuable source of evidence to assess, through triangulating evidence, whether teacher surveys, interviews, practice and the aims of the intervention were in alignment.

4.4.3   Online platform use

The online platform automatically recorded usage data. Teachers were required to log into the platform to view or add content. The platform logged every time a teacher visited a page, viewed a video, uploaded a video or used one of the video analysis forms. These data allowed us to compare use of the platform by school and over time. However, these data do not reveal what teachers were doing on each occasion.

4.4.4   Additional engagement checks

We conducted short, roughly ten-minute, telephone interviews with 8 teachers (who were not project leaders) during final stage of the intervention. We used these interviews to check our interpretation of the other data we had regarding the school's engagement with the intervention. Interviews were recorded and transcribed to allow consideration and discussion by the research team. We also conducted a short survey of school champions to gather information on variation in the implementation of the intervention and their views about teacher engagement and the efficacy of the intervention.

5.   Results

We first provide evidence of engagement with the intervention and details of its implementation. This process evidence, as described above, comes from data on teachers' use of the online platform, review of videos of video club meetings, analysis of lesson videos, teachers' self-reporting on their engagement and school champions' reports on the progress of the intervention.

We follow this with a difference-in-difference analysis examining changes in teachers' thinking and self-reported practice, comparing participating teachers with other teachers in participating schools and then by level of engagement with the intervention.

5.1   Intervention engagement and processes

5.1.1 School and teacher participation in, and response to, the intervention

The first question is how strongly the schools and teachers participated in the intervention. The attrition rate was low. One school (12) dropped out very early when the school champion left the school. The engagement of School 11 was disrupted by an external school inspection, but even after receiving a disappointing grade from OfSTED the school chose to use its participation in the project as a means of addressing priorities for development. We are not aware of any individual teacher dropping out of the project within their school although there were, of course, some absences from particular events. The online platform logged each instance when a teacher viewed a video (see Table 2). On average, participants used the online platform once every three days.

Table 2 - Number of separate occasions on which the platform was accessed (per participant or teacher) over the whole pilot project period

School ref no.	School new to IRIS Connect	School adopted whole-school participation	Number of online platform hits per project participant	Number of Online platform hits per teacher in the school
1	0	0	121	28
2	0	0	64	7
3	0	0	23	10
4	0	1	81	81
5	0	0	45	19
6	1	0	199	47
7	1	0	34	20
8	1	1	47	47
9	1	0	27	11
10	1	1	37	37
11	1	0	44	15
Group Averages	Schools new to IRIS Connect		65	29
	Schools with pre-project experience		67	29
	Schools with whole-school approach		55	55
	Schools with small group approach		70	20

Teachers' use of the online system was strongly associated with the timing of video club meetings. This was clear from the spikes in metrics data and confirmed by reports from school champions and the end-of-project survey. School champions agreed (n=5) or strongly agreed (n=6), that the 'Teachers at my school positively engaged with the video clubs'. They also agreed (n=4) or strongly agreed (n=7) that the "The online content on the IRIS-Connect platform to support video clubs was stimulating".

Teachers' self-reported responses to using the technology were generally positive (Table 3). Nearly 90% of participants reported that the video technology was easy to use and roughly 80% of participants reported that the focus on formative feedback was relevant to their professional development needs. Roughly three-quarters of participants reported that the intervention had made them more likely to reflect on their practice. In the end of project survey over 95% of participants asserted they had changed their practice as a consequence of the intervention. Change in approach to using questions was the most frequently cited area of practice but they also referred to changes in feedback more generally. Participants interviewed at the end of project expressed an equally positive view of the effect of the intervention on their practice, referring chiefly to feedback, dialogue and pupil ownership of classroom talk.

Table 3 - Teachers' beliefs about value and practicality of the intervention

	Statement	Strongly Disagree	Disagree	Unsure	Agree	Strongly Agree	Total
1	It is easy to use IRIS-Connect cameras to record videos	2 2.2%	15 16.7%	13 14.4%	52 57.8%	8 8.9%	90
2	It is easy to use the IRIS-Connect platform to view my recorded videos	1 1.1%	7 7.8%	4 4.4%	63 70.0%	15 16.7%	90
3	The amount of preparation needed for video clubs has been too much to manage	3 3.5%	44 51.2%	26 30.2%	12 14.0%	1 1.2%	86
4	The time I have spent using IRIS-Connect has been worth it	0 0.0%	5 5.6%	20 22.5%	50 56.2%	14 15.7%	89
5	Using IRIS-Connect has become an important part of my professional development practice	0 0.0%	20 22.2%	20 22.2%	43 47.8%	7 7.8%	90
6	The focus on formative feedback has fit within my own professional development needs	0 0.0%	5 5.7%	11 12.5%	61 69.3%	11 12.5%	88
7	Attending video clubs has made me more likely to reflect on my own practice	0 0.0%	7 8.2%	18 21.2%	50 58.8%	10 11.8%	85
8	Video clubs have led to increased teacher collaboration in my school	0 0.0%	9 10.7%	25 29.8%	36 42.9%	14 16.7%	84
9	Video clubs have been an effective approach to professional learning	0 0.0%	2 2.4%	21 24.7%	50 58.8%	12 14.1%	85

5.1.2 Adherence of school co-ordinators (champions) to guidance on leading video clubs

The design of the intervention relied on leadership by school champions rather than external experts. This raises the question of whether the implementation of the intervention adhered to the guidance that was provided by the intervention team. This was judged by observing videos of video club meetings and by school champions' self-reports of those meetings. These showed that the video clubs did focus on the quality of feedback and dialogic teaching: e.g. the need for teachers to refrain from giving the correct answer 'too soon' to encourage dialogue; the relationship between the task and the ensuing dialogue; pupil independence, ownership and engagement; and the role of praise in promoting dialogue.

5.1.3 Analysis of participants' teaching

Although the online platform metrics showed when a teacher had created a lesson video we were only able to observe lessons when teachers granted permission. It seems appropriate to interpret teachers' selection of videos as tending to reflect what they considered to be good teaching in the context of the stated purpose of the intervention. Therefore, our question in analysing the lesson videos was 'To what extent does this teaching align with the view of powerful feedback promoted by the intervention?' Participating teachers shared 43 lesson videos and Table 4 presents an analysis of these videos using the same categories and 11-point scale as the survey of teachers. Between 50% and 75% of the lesson videos were rated as

displaying an approach to feedback in line with aims of the intervention (positive mean scores for items 1, 3, 4 and 6 in Table 4 and negative figures for items 2, 5 and 7). The most positive ratings were for 'poses serious challenges' and 'leaves the judgement to the pupil'.

Table 4 - Characterisation of teachers' feedback visible in shared lesson videos using the same format as the teacher survey (frequencies of lesson videos at each point in the scale)

	Strength of tendency towards <u>left</u> hand statement					Strength of tendency towards <u>right</u> hand statement							mean
	-5	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**	**5**		
1 Gives easy/simple next steps	0	2	6	2	0	0	2	10	11	9	1	*Poses serious challenges*	1.53
2 Highlights differences between alternative responses/method	0	5	9	5	2	7	1	6	5	1	2	Gives correct response/method	-0.40
3 Gives correct series of steps or points	0	1	5	5	1	3	5	4	14	4	1	*Reviews pupil's steps or points*	1.07
4 Makes pupil feel good	0	1	1	2	2	7	0	0	5	18	7	*Leaves the judgement to the pupil*	2.53
5 Reflects the teacher's judgement about pupil's current understanding	0	5	10	9	2	2	4	2	6	3	0	Tells the pupil how close they were to teacher's answer	-0.74
6 Tells a pupil the problem with their method/thinking	0	2	3	4	2	5	3	5	9	9	1	*Helps the pupil to work out what the problem is with their thinking/method*	1.26
7 Helps pupils understand different ways of seeing a problem	0	1	8	10	2	3	2	7	4	4	2	Helps pupils know what they are expected to do/say	0.09

5.1.4 Variation in implementation and engagement by school and teacher

We were able to examine several ways in which engagement and implementation varied between schools. First, given the emphasis on 'learning communities' in the literature on video clubs, it is reasonable to suppose that schools which had been using the technology before participating in the intervention would have developed a more favourable culture for successful implementation. Teachers' fears over 'surveillance' of their lessons, would have had more opportunity to be assuaged. On this reading, longer use of IRIS Connect technology would have been associated with stronger commitment by senior management to school improvement through collaborative professional development. However, the metrics collected through the online system revealed no difference in the number of online platform hits per teacher between experienced schools and schools new to IRIS Connect.

Second, engagement per participant was about 25% more in schools adopting a small group format, although engagement per teacher when considering all teacher in the school was naturally higher (nearly 3 times higher) in schools that adopted a whole school approach. In each term, teachers created an average of 20 lesson videos per school, but the level and timing of activity varied considerably between schools. On average, each participating teacher created just under five lesson videos, though this average was inflated by very high rates of use in schools 1 and 6.

Third, school champions believed that willingness to discuss practice was helped by the sequence of video clubs focusing on others' lessons followed by video clubs in which teachers shared their own lessons. They also believed that the focus on feedback and classroom dialogue was consistent with schools' improvement plans. However, teachers in three schools (Schools 2, 5 and 9) created no lesson videos in the final 3 months of the intervention. The headteacher of School 5 attributed this tail-off to staff turnover, but no explanation was offered by the other two schools.

Fourth, in some schools, champions led video clubs in a manner that fostered open discussion during which teachers considered implications for their own practice. Discussion between teachers focused on possible developments in their practice with the talk being conducted in terms of 'what we want to develop' and 'reflecting on how [we] can best facilitate this.' In other schools, the video club discussions kept very close to the guidance provided through the online platform and, in these schools, teachers made relatively few comparisons with their own practice. Relatively few opinions were aired, followed by a small amount of clarification/agreement before the next question was 'read out'. Nonetheless, even in schools which started with a rigid format, later video clubs became more relaxed. We observed no relationship between the number of teachers attending the video club and the style in which it was led.

We compared the lesson videos (24) from schools in which video clubs had been led in an open style with those (9) from schools in which video clubs were led in a closed style. On average, the lesson reviewers judged that feedback in schools in which video clubs had been led in an open style gave more emphasis to each of the qualities of feedback judged by the intervention as more powerful (using the data in Table 4). This is a small sample which must be treated with caution, but a simple t-test suggested that each of the differences was statistically significant at the 0.05% level and five of the seven were statistically significant at the 1% level. Whether this is a causal link and its direction is unclear.

Finally, we also gathered some data on how teachers believed their own feedback practice varied according to context. Teachers were asked in the end of intervention survey to report how their feedback practice varied. They claimed to provide more scaffolding and focused feedback earlier in the year, whilst giving greater emphasis to challenge, extension and making links later in the school year. They also claimed that they adjusted their type of feedback according to differences between children. They justified this difference on the basis that type of feedback had to be adjusted to take account of levels of pupil understanding and ability. If this turned out to be a general pattern, then we would observe differences between schools related to their intakes of pupils.

Table 5 - Teacher Self-Characterisation of Feedback (reduced scale)

Statement 1 "My feedback typically…"	Survey Group		% Responses (0dp)			Statement 2 "My feedback typically…"	Effect size †
			Left	Neutral	Right		
1. provides the child with a simple next step they can easily achieve *(2)*$^\Omega$	Baseline	Participant	45	15	40	**presents the child with a serious challenge *(4)***	-0.01
		Non-Participant	52	12	36		
	Endpoint	Participant	27	16	57		
		Non-Participant	38	11	51		
2. **carefully highlights the differences between alternative ways of solving a problem** *(3)*	Baseline	Participant	40	21	39	gives pupils a correct way of solving a problem *(2)*	.07‡
		Non-Participant	37	15	48		
	Endpoint	Participant	52	20	28		
		Non-Participant	54	13	33		
3. makes plain the sequence of steps a child should follow in solving a problem *(2)*	Baseline	Participant	25	13	62	**concentrates on helping a child to review the steps they have taken to solve a problem *(4)***	.26
		Non-Participant	17	24	59		
	Endpoint	Participant	31	22	46		
		Non-Participant	40	20	40		
4. leaves the child feeling good about what they have done *(1)*	Baseline	Participant	56	14	31	**leaves the child to judge whether they have done a good job *(4)***	.24
		Non-Participant	47	13	40		
	Endpoint	Participant	54	18	28		
		Non-Participant	60	12	29		
5. **reflects my judgement about why a child has given a particular answer** *(3)*	Baseline	Participant	39	39	22	tells the child how close their answer was to my *answer* *(2)*	-.08‡
		Non-Participant	49	28	23		
	Endpoint	Participant	58	22	19		
		Non-Participant	59	23	18		
6. tells a child what the problem is with their thinking *(3)*	Baseline	Participant	19	6	75	**helps a child to work out what the problem is with their thinking *(4)***	.26
		Non-Participant	12	7	81		
	Endpoint	Participant	15	12	72		
		Non-Participant	23	8	69		
7. **concentrates on helping children to understand different ways of seeing a problem** *(4)*	Baseline	Participant	48	14	38	concentrates on making sure that children know what they are expected to do *(2)*	-.07‡
		Non-Participant	49	12	39		
	Endpoint	Participant	59	20	21		
		Non-Participant	50	17	33		

† Change for participants relative to non-participants. Calculated using Cohen's d

‡ The intervention aimed to increase use of feedback described by the left hand side of this row, so for this row a negative effect size is desirable.

$^\Omega$ Power of feedback categorised following Hattie and Timperley (2007): (1) feedback on self; (2) feedback on task (3) feedback on processing of task; (4) feedback on self-regulation

5.2   Change in teachers' self-reported feedback practice

5.2.1 Difference-in-difference analysis by participation

The comparison between baseline and end-of-project surveys is used to examine change in participating teachers' thinking about feedback. Table 5, above, presents survey results for the questions asking participants to indicate their typical approach to each of seven aspects of feedback. As noted in Section 4.4.1, there were differences in the number of participants versus non-participants responding to both survey points. To counteract and investigate the impact of these missing data within the difference-in-difference analysis, we ran a multiple imputation[ii]

to estimate missing data and thereby retain the information from participants with a single survey response or other missing data within the analyses. There were some small but appreciable differences between the imputed and complete case results. Therefore, the imputed results are reported in this section; complete case results for Table 5 are provided in Appendix 2, Table A2.

The table truncates the 11 points in the survey to 3 categories (left, neutral, right) for clarity of presentation. The scores for participating and non-participating teachers were largely similar at the outset. While there were some differences, a comparison of baseline scores using chi-squared tests on all of the items in Table 5 returned no statistically significant test statistics. We compared the change in reported practice (participants vs non-participants) through an effect size (final column of Table 5). The effect size (Cohen's d) was calculated from the original 11-point scale:

$$Effect\ size = \frac{\frac{\sum_{1}^{n_1}(f_{p2} - f_{p1})}{n_1} - \frac{\sum_{1}^{n_2}(f_{o2} - f_{o1})}{n_2}}{\sqrt{\frac{\sigma_{\Delta p}^2 + \sigma_{\Delta o}^2}{2}}}$$

Where:

f_{p1} is the score (on 11 point scale) on that item for participant p in the pre-survey

f_{o2} is the score (on 11 point scale) on that item for non-participant o in the post-survey

n_1 is the number of participants and n_2 is the number of non-participants

$\sigma_{\Delta p}^2$ is the variance of the difference between pre- and post- survey scores for participants

Of the 7 effect sizes calculated, three (3, 4 and 6) had an absolute magnitude of greater than 0.2 and can therefore be considered to be 'small' effect sizes. The direction of the effect size in each case suggests that the intervention encouraged teachers to adopt more powerful types of feedback, or at least to believe that they were doing so. However, in all three cases, the effect size has been driven by non-participants *moving away* from the feedback approach targeted. This tendency is also apparent in the complete case results (see Table A2, Appendix 2). We cannot be sure what has driven this result and whether it stems from a feature of the research design, such as 'contamination' or self-selection effects for participants and non-participants based at the same schools; an issue with the self-report instrument and how it was interpreted by teachers; natural variation in teacher self-reported feedback practice as the school year developed which was arrested by the intervention; or some other factor.

5.2.2 Change in self-reported feedback by level of engagement

Given the inconclusive results pertaining to changes in teacher self-reported practice in the previous section and the difficulties with the difference-in-difference analysis discussed, in this final analysis we use participant self-reported engagement with the intervention to identify and

compare the results by intervention engagement level. For this we identified two survey items relating to engagement: 'Using IRIS-Connect has become an important part of my professional development practice' and 'Film clubs have been an effective approach to professional learning'. For each of these items, a score of 2 was given for a 'strongly agree' response and 1 for 'agree', otherwise 0. The sum of these two scores was calculated and the participants grouped of 'none' for scores of 0 (28% of 172 teachers), 'low' for a score of 1 (21%), 'medium' for 2 (36%) and 'high' for scores of 3 and 4 (14%).

Using the original 11-point scale, the mean change in response was calculated for all levels of engagement (see Table 6, below)

Table 6 – Change[†] in feedback by self-reported engagement

	Engagement Level				
	None	*Low*	*Med*	*High*	
1. provides the child with a simple next step they can easily achieve	1.17	0.92	1.26	1.03	**presents the child with a serious challenge**
2. **carefully highlights the differences between alternative ways of solving a problem**	-0.69	-1.14	-0.91	-1.60	gives pupils a correct way of solving a problem
3. makes plain the sequence of steps a child should follow in solving a problem	-0.22	-0.94	-1.20	-1.04	**concentrates on helping a child to review the steps they have taken to solve a problem**
4. leaves the child feeling good about what they have done	-0.73	-0.86	-0.89	0.37	**leaves the child to judge whether they have done a good job**
5. **reflects my judgement about why a child has given a particular answer**	-0.61	-0.79	-0.86	-0.80	tells the child how close their answer was to my *answer*
6. tells a child what the problem is with their thinking	-0.56	0.07	-0.50	-0.38	**helps a child to work out what the problem is with their thinking**
7. **concentrates on helping children to understand different ways of seeing a problem**	0.10	-0.53	-0.79	-1.00	concentrates on making sure that children know what they are expected to do

[†] Negative scores indicate a shift to the left-hand statement on the 11-point scale.

As with the previous results, the change in feedback by engagement presented in Table 6 suggests that the instrument is identifying changes in thinking, but the results by engagement level do not exhibit a clear and consistent pattern. There is no meaningful relationship between engagement and feedback change for items 1, 5 and 6. Other items suggest differences between non-engagers and the rest (item 3), high engagers and the rest (item 4) and a relationship between engagement and change (items 2 and 7). Three of the four of these (items 2, 4 and 7) are in the expected direction but, overall, the results are inconclusive.

6.   Discussion

We close the paper by, first, summarising the main findings and discussing what this study suggests about the potential of video technology within CPD programmes; and, second, commenting on the methodological difficulties encountered by this study and what future research looking to evidence changes in teacher thinking and practice can do to address them.

6.1   The potential impact of and barriers to using video technology within CPD programmes

In general, the process and implementation evidence supported the feasibility of the programme to participants, the majority of whom (71.9%) felt the time they spend using IRIS-Connect had been worthwhile. The engagement data revealed that all schools with the exception of one, completed the core requirements of running video clubs and wider data suggests that, for many, there was good engagement with the programme, particularly evident in the platform metrics around the video clubs, which structured and motivated sustained engagement.

The notable differences in implementation we have reported related to the openness and richness of the discussion observed in video clubs, the level of engagement with the video platform and the length of engagement (with activity dropping in the Summer term for three schools). Our evidence suggested that these factors mattered. For the teachers most engaged with the programme, who shared videos and agreed to be interviewed, the videos shared demonstrated feedback practice in line with the intervention aims. Another factor of apparent importance was the 'openness' of the discussion we observed in the video club meetings. More rigid earlier sessions in some schools suggested lower ownership, engagement and depth in discussion and enquiry. It is likely that the discussion we observed reflected numerous factors, including the wider, and pre-existing professional learning culture of the school and the confidence of the teachers, but nonetheless appears to be a supporting factor for effectiveness.

Taken as a whole, the results suggest that depth of engagement was a key factor in success and that *whether* the intervention works is a less valuable question than asking what conditions it works under. Interviews, survey data and videos suggest that professional learning culture, quality of facilitation and time provided for teachers to engage with the intervention all mattered for success. There were also suggestions that the buy-in of teachers (i.e. seeing feedback as an area they wished to improve) was also supportive of impact. In sum, the evidence suggests that video clubs were supportive of, creates opportunities for and were a medium for professional learning, but that there were other factors which determined whether or not this was successful.

6.2   Capturing Changes in Teacher's Thinking

One aspect of this evaluation study was go beyond basic implementation and engagement evidence to assess changes in teachers' thinking and indicative evidence of relationships between thinking and practice. For the latter, we developed and used a new instrument to assess changes in teachers' principles for feedback, comparing participants and non-participants to assess the impact of the intervention. The findings in this respect were inconclusive: many items did not register clear differences and the ones that did seemed to suggest that it was the

non-participants who were changing. This is likely to have stemmed from a combination of, first, limitations in the validity and reliability of the instrument and, second, difficulties creating a comparison group from non-participants at the same school. The instrument, while a reasonable interpretation of the feedback literature, also left considerable scope for interpretation, and, combined with inadequacies in the comparison group, produced an inconclusive picture. In sum, it is likely that a combination of not having a hard cut off between participants and non-participants, coupled with the difficulties of capturing changes in teacher thinking and self-reported practice led to a combination of comparison groups, intervention, impact and instrument which were too loosely coupled to both identify and explain a clear result.

These issues notwithstanding, we argue that there is great value in process evaluation going beyond basic compliance and attendance indicators to better understand why and in what conditions programmes work. For this, there is value in developing instruments which can capture major aspects of teaching practice, such as classroom feedback, for use as a professional development as well as a research tool. Participants in this intervention faced the challenge of translating principles about effective feedback in the literature into, first, professional development and, second, classroom practice. Researchers faced the parallel problem of capturing evidence of this process. We hold that bridging this gap is an important aspect of realising evidence-informed practice. Recent developments in our ability to measure and record dialogic teaching, as discussed earlier in this paper, are good examples of assessment tools which can record realistic classroom practice.

7.   References

Alexander, R. (2008). Culture, dialogue and learning: Notes on an emerging pedagogy. In N. Mercer & S. Hodgkinson (Eds.), *Exploring Talk in School: Inspired by the Work of Douglas Barnes* (pp. 91-114). London: Sage.

Alexander, R. (2018). Developing dialogic teaching: Genesis, process, trial. *Research Papers in Education*, *33*(5), 561-598.

Alles, M., Seidel, T., and Gröschner, A. (2019). Establishing a positive learning atmosphere and conversation culture in the context of a video-based teacher learning community. *Professional Development in Education*, *45*(2), 250-263.

Bates, M. S., Phalen, L. and Moran, C. G. (2016). If you build it, will they reflect? Examining teachers' use of an online video-based learning website. *Teaching and Teacher Education, 58*, 17-27.

Beisiegel, M., Mitchell, R., and Hill, H. C. (2018). The design of video-based professional development: An exploratory experiment intended to identify effective features. *Journal of Teacher Education*, *69*(1), 69-89.

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

Brouwer, C. (2011). *Equipping teachers visually.* Eindhoven: Eindhoven School of Education.

Brouwer, N., Besselink, E. and Oosterheert, I. (2017). The power of video feedback with structured viewing guides. *Teaching and Teacher Education, 66*, 60-73.

Brown, G. T., Harris, L. R. and Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education, 28*(7), 968-978.

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., Crisp, B., Saunders, L. and Coe, R. (2015). *Developing great teaching: lessons from the international reviews into effective professional development*. London: Teacher Development Trust.

Darling-Hammond, L., Maria E. Hyler, Gardner, M. and Espinoza, D. (2017). *Effective Teacher Professional Development*. Palo Alto, CA: Learning Policy Institute.

Davies, P. (2018) *Paying for Education: Debating the Price of Progress.* Routledge.

Davies, P. and Dunnill, R. (2008) ''Learning Study'as a model of collaborative practice in initial teacher education', *Journal of Education for Teaching,* 34(1), pp. 3-16.

Department for Education (2016). *Schools, pupils and their characteristics. January 2016. (SFR 20/2016).* London: Department for Education.

Edwards-Groves, C., Anstey, M. and Bull, G. (2014). *Classroom talk: Understanding dialogue, pedagogy and practice.* Newtown, PETAA.

EEF (2018) *Teaching and Learning Toolkit: Feedback*: EEF. Retrieved from https://educationendowmentfoundation.org.uk/evidence-summaries/teaching-learning-toolkit/feedback/.

Gaudin, C. and Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review, 16*, 41-67.

Gillies, R. M. (2016). Dialogic interactions in the cooperative classroom', *International Journal of Educational Research. 76*, 178-189.

Gröschner, A., Seidel, T., Kiemer, K. and Pehmer, A.-K. (2015). Through the lens of teacher professional development components: the 'Dialogic Video Cycle' as an innovative program to foster classroom dialogue. *Professional Development in Education, 41*(4), 729-756.

Gröschner, A., Seidel, T., Pehmer, A.-K. and Kiemer, K. (2014). Facilitating collaborative teacher learning: the role of "mindfulness" in video-based teacher professional development programs. *Gruppendynamik und Organisationsberatung, 45*(3), 273-290.

Hattie, J. and Timperley, H. (2007). The Power of Feedback. *Review of Educational Research, 77*, 81-112.

Hennessy, S., Rojas-Drummond, S., Higham, R., Márquez, A. M., Maine, F., Ríos, R. M., García-Carrión, R., Torreblanca, O. and Barrera, M. J. (2016). Developing a coding scheme for analysing classroom dialogue across educational contexts. *Learning, Culture and Social Interaction, 9*, 16-44.

Holmqvist, M. (2010). Teachers' learning in a learning study. *Instructional Science, 39*(4), 497-511.

Howe, C. and Abedin, M. (2013). Classroom dialogue: a systematic review across four decades of research. *Cambridge Journal of Education, 43*(3), 325-356.

IRIS Connect (2014) *Teaching is learning programme: 6 steps to success*. Brighton: IRIS Connect.

Jay, T., Willis, B., Thomas, P., Taylor, R., Moore, N., Burnett, C., Merchant, G. and Stevens, A. (2017). *Dialogic Teaching: Evaluation report and executive summary*. London: EEF. Retreived from https://educationendowmentfoundation.org.uk/public/files/Projects/Evaluation_Reports/Dialogic_Teaching_Evaluation_Report.pdf.

Joyce, B. R. and Showers, B. (2002). *Student achievement through staff development*. Alexandria, VA: ASCD.

Kane, T., Gehlbach, H., Greenberg, M., Quinn, D. and Thal, D. (2015). *The best foot forward project: Substituting teacher-collected video for in-person classroom observations.* Center for Education Policy Research: Harvard Graduate School of Education. Retrieved from http://cepr.harvard.edu/files/cepr/files/l4a_best_foot_forward_research_brief1.pdf.

Kane, T. J., McCaffrey, D. F., Miller, T. and Staiger, D. O. (2013). *Have We Identified Effective Teachers? Validating Measures of Effective Teaching Using Random Assignment.* Seattle: Bill & Melinda Gates Foundation.

Khong, T. D. H., Saito, E., & Gillies, R. M. (2019). Key issues in productive classroom talk and interventions. *Educational Review*, *71*(3), 334-349.

Kiemer, K., Gröschner, A., Pehmer, A.-K. and Seidel, T. (2015). Effects of a classroom discourse intervention on teachers' practice and students' motivation to learn mathematics and science. *Learning and Instruction, 35*, 94-103.

Kleinknecht, M. and Schneider, J. (2013). What do teachers think and feel when analyzing videos of themselves and other teachers teaching?. *Teaching and Teacher Education, 33*, 13-23.

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: a literature review. *Quality & Quantity, 47*(4), 2025-2047.

Lewis, C., Perry, R. and Murata, A. (2006). How should research contribute to instructional improvement? The case of lesson study. *Educational researcher, 35*(3), 3-14.

Major, L., & Watson, S. (2018). Using video to support in-service teacher professional development: the state of the field, limitations and possibilities. *Technology, Pedagogy and Education*, *27*(1), 49-68.

Mercer, N. and Dawes, L. (2014). The study of talk between teachers and students, from the 1970s until the 2010s. *Oxford Review of Education, 40*(4), 430-445.

Mercer, N., Dawes, L. and Staarman, J. K. (2009). Dialogic teaching in the primary science classroom. *Language and Education, 23*(4), 353-369.

Pehmer, A.-K., Gröschner, A. and Seidel, T. (2015). Fostering and scaffolding student engagement in productive classroom discourse: Teachers' practice changes and

reflections in light of teacher professional development. *Learning, Culture and Social Interaction, 7*, 12-27.

Quinn, D. M., Kane, T. J., Greenberg, M., & Thal, D. (2018). Effects of a video-based teacher observation program on the de-privatization of instruction: evidence from a randomized experiment. *Educational Administration Quarterly*, *54*(4), 529-558.

Rich, P. J. and Hannafin, M. (2008). Video Annotation Tools. *Journal of Teacher Education, 60*(1), 52-67.

Schindler, A.-K., Gröschner, A. and Seidel, T. (2016). Teaching Science Effectively: A Case Study on Student Verbal Engagement in Classroom Dialogue. *Orbis Scholae, 2015*(2), 9-34.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M. and 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*, 259-267.

Sherin, M. G. and Han, S. Y. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education, 20*(2), 163-183.

Sherin, M. G., Linsenmeier, K. A. and van Es, E. A. (2009). Selecting Video Clips to Promote Mathematics Teachers' Discussion of Student Thinking. *Journal of Teacher Education, 60*(3), 213-230.

Sherin, M. G. and van Es, E. A. (2008). Effects of Video Club Participation on Teachers' Professional Vision. *Journal of Teacher Education, 60*(1), 20-37.

Slater, H., Davies, N. M. and Burgess, S. (2012). Do Teachers Matter? Measuring the Variation in Teacher Effectiveness in England. *Oxford Bulletin of Economics and Statistics, 74*(5), 629-645.

Smith, H. and Higgins, S. (2006). Opening classroom interaction: the importance of feedback. *Cambridge Journal of Education, 36*, 485-502.

Tripp, T. R. and Rich, P. J. (2012). The influence of video analysis on the process of teacher change. *Teaching and Teacher Education, 28*(5), 728-739.

Van den Bergh, L., Ros, A. and Beijaard, D. (2014). Improving Teacher Feedback During Active Learning: Effects of a Professional Development Program. *American Educational Research Journal, 51*(4), 772-809.

van Es, E. A. (2012). Examining the development of a teacher learning community: The case of a video club. *Teaching and Teacher Education, 28*(2), 182-192.

Vangrieken, K., Meredith, C., Packer, T. and Kyndt, E. (2017). Teacher communities as a context for professional development: A systematic review. *Teaching and Teacher Education, 61*, 47-59.

Walshaw, M. and Anthony, G. (2008). The teacher's role in classroom discourse: A review of recent research into mathematics classrooms. *Review of educational research, 78*(3), 516-551.

Wiliam, D. (2011). What is assessment for learning?. *Studies in educational evaluation*, *37*(1), 3-14.

Zhang, M., Lundeberg, M., Koehler, M. J. and Eberhardt, J. (2011). Understanding affordances and challenges of three types of video for teacher professional development. *Teaching and Teacher Education, 27*(2), 454-462.

Appendices

Appendix 1 - Guidance for film clubs

Film clubs 1-3: Teachers viewing and discussing video clips provided through the online platform
Guidance for each of film clubs 1-3 was provided online (through the 'group' function in online platform) under the following sub-headings:

(i) **Introduction**;
(ii) **Classroom talk & (Film Club 1) teaching; (Film club 2) questioning and group talk; (Film club 3) feedback**: one 'page of guidance for each module supplemented by 1-2 'talking heads' videos of 30 sec – 1 min 30 second duration.
(iii) **Things to consider**: more note-form guidance including a section 'what sort of questions should I be asking myself when watching lesson video clips?'
(iv) **Pre-film club reflection**: more guidance on the topic
(v) **Film club**: guidance (linked to two 1-2 minute video clips) on how to organise the film club meeting –
  1 Ensure everyone is aware of the lesson context
  2 Watch the clip once without pausing
  3 Consider the following questions:
     What do you think about the nature of the True or False task?
     What other ways could this maths question have been presented?
     How is the task introduced?
     How does the task affect the nature of the discussion?
  4 Watch the clip a second time with these questions in mind.
  5 Watch the clip a third time. Feel free to pause and discuss where you feel it appropriate. Use the time-stamped comments feature to make a note of anything that stands out or that you find interesting (you may notice comments from other teachers involved in the project; you may wish to consider these when having your own discussions).
(vi) **Report your film club findings**: Champions were encouraged to record their film club sessions and to share these online with the evaluation team. Alternatively, they were given a set of open response items through which they could summarise the film club. For example, after film club 2 the prompts asked about: attendees; the main points of discussion during 'your own reflections'; the main points of discussion after viewing the film clips; any changes in perspective during the course of the discussion; the intentions of the project team in future practice following the film club; judgement about the effectiveness of the film club.
(vii) **Reflections**: prompts on how to review own practice following the film club.

Film clubs 4 & 5: Teachers reviewing and discussing videos of their own lessons
Teachers were mencouraged to organise their fourth and fifth film clubs (in which they used their own lesson or lessons) in six steps:

1. **Collaborative planning:** in which they were asked to plan together a 'rich task or question'. Four short video clips were provided as exemplification. The planning guidance also included a 'reminder about dialogue' (with a list of 11 desirable characteristics of dialogue) and they were also reminded about feedback with a short clip featuring Dylan Wiliam and Table A2A.1 which was adapted from Hattie & Timperley's (2007) review of research on feedback.

Table A1.1 (which was untitled in the online guidance)

Feedback level	Example	Effectiveness
Feedback about the self	What a superstar! You did that really well.	Weak
Feedback about the task	So you're saying the two sums would not be equal; can you work out what the final equation would read that proves that?	Good when supported by strategies for learners to try
Feedback about strategies or processes needed in tasks	By suggesting something might go wrong you make the reader worry about your character, which creates suspense and makes them want to read on. Is there any way you could introduce this earlier to engage the reader right from the beginning?	Powerful in the short term
Feedback about self -regulation	So you adjusted the weight at the front of your model plane, after you found it kept nose-diving, by removing a paperclip. Is there any way you can make it glide for longer? Where can you find out how to improve this? Has anyone managed to make theirs go further? What can you learn from theirs?	Powerful in the long term

2. **Reflecting, editing, commenting and sharing:** in which they were again encouraged to focus on episodes in the lesson which they found interesting in terms of dialogic teaching and feedback. Instructions were provided on how to edit lessons to focus on the episodes and how these could be shared with other teachers in the group through the online platform.
3. **Peer feedback:** in which teachers were invited to share their lesson with one other colleague who would add comments to the video clip using the online platform. This invitation was accompanied by guidance that comments should "(i) respect and support the teachers and pupils in the video; (ii) suspend judgement to avoid faulty assumptions; (iii) focus on interactions and how they are linked; and (iv) look deeper and seek context".
4. **Sharing** with the project group: in which the video clip(s) and associated comments would be shared with the whole project group and comments invited from all group members.
5. **Pre film club reflections:** in which group participants were invited to view the clips and to reflect on aspects of the teaching and learning they wanted to discuss at the film club.

6. **Film club:** in which the group met to discuss the video clip(s) with the following focus: "(i) the effectiveness of the planned task/question; (ii) how learners are invited into dialogue; (iii) how dialogue develops; (iv) how language promotes a positive learning culture; (v) what pupil responses tell you about their learning; (vi) the source of feedback (teacher, self, task, peer) and how it is offered; (vii) how feedback relates to the perceived learning intention; (viii) how feedback promotes independent learning? (ix) how feedback creates thinking and furthers understanding; (x) how learning is led (pupil-led / teacher-led) and the impact this has".

Film Club 6 Evaluating the experience

In Film Club 6 teachers were asked to reflect on their experience using the following questions and champions were asked to summarise the thoughts expressed using the online platform. The questions were:

1. Please give an overview of your school approached the project, including who was involved and how this was organised.
2. Did you like this approach to professional learning?
3. Why?
4. What did you and the other participants learn?
5. What changes, if any, have been made to classroom practice?
6. What benefits for learners did you observe? Do you have any other evidence to back this up?
7. Were there any negative effects?
8. What challenges did you experience? How did you overcome these?
9. Where do you plan to go from here?

Appendix 2 – Complete Case Analysis

Table A2 (complete case) - Teacher Self-Characterisation of Feedback (reduced scale)

Statement 1 "My feedback typically…"	Survey Group		% Responses (0dp)			Statement 2 "My feedback typically…"	Effect size †
			Left	Neutral	Right		
1. provides the child with a simple next step they can easily achieve *(2)*^Ω	Baseline	Participant	42	15	42	**presents the child with a serious challenge** *(4)*	.16
		Non-Participant	52	19	29		
	Endpoint	Participant	23	19	58		
		Non-Participant	50	16	34		
2. **carefully highlights the differences between alternative ways of solving a problem** *(3)*	Baseline	Participant	39	20	41	gives pupils a correct way of solving a problem *(2)*	-.04‡
		Non-Participant	35	23	42		
	Endpoint	Participant	53	21	26		
		Non-Participant	58	9	33		
3. makes plain the sequence of steps a child should follow in solving a problem *(2)*	Baseline	Participant	25	12	63	**concentrates on helping a child to review the steps they have taken to solve a problem** *(4)*	.42
		Non-Participant	10	32	58		
	Endpoint	Participant	29	27	44		
		Non-Participant	34	34	31		
4. leaves the child feeling good about what they have done *(1)*	Baseline	Participant	59	16	26	**leaves the child to judge whether they have done a good job** *(4)*	.29
		Non-Participant	61	16	23		
	Endpoint	Participant	50	21	29		
		Non-Participant	75	9	16		
5. **reflects my judgement about why a child has given a particular answer** *(3)*	Baseline	Participant	40	40	21	tells the child how close their answer was to my *answer (2)*	-.49‡
		Non-Participant	66	21	14		
	Endpoint	Participant	60	24	16		
		Non-Participant	55	27	18		
6. tells a child what the problem is with their thinking *(3)*	Baseline	Participant	20	3	76	**helps a child to work out what the problem is with their thinking** *(4)*	.36
		Non-Participant	10	3	87		
	Endpoint	Participant	11	15	74		
		Non-Participant	18	12	70		
7. **concentrates on helping children to understand different ways of seeing a problem** *(4)*	Baseline	Participant	49	10	41	concentrates on making sure that children know what they are expected to do *(2)*	-.11‡
		Non-Participant	52	16	32		
	Endpoint	Participant	56	24	19		
		Non-Participant	55	21	24		

† Change for participants relative to non-participants. Calculated using Cohen's d
‡ The intervention aimed to increase use of feedback described by the left hand side of this row, so for this row a negative effect size is desirable.
^Ω Power of feedback categorised following Hattie and Timperley (2007): (1) feedback on self; (2) feedback on task (3) feedback on processing of task; (4) feedback on self-regulation

Notes

ⁱ Based on figures provided by IRIS Connect and data from the Department for Education: (Department for Education, 2016)

ⁱⁱ The multiple imputation used imputation by chained equations, ordered logit to reflect the ordinal scale. It produced 10 imputed datasets and was conducted in Stata 13