Video–based observation in impact evaluation

Simon Borg [*,1]

Department of Language, Literature, Mathematics and Interpretation, Western Norway University of Applied Sciences, Bergen, Norway

ARTICLE INFO	ABSTRACT
	This paper reviews the literature on the use of video–based observation (VBO) with particular attention to monitoring and evaluation (M&E) on development projects. While the use of video both as a research tool and as a strategy for supporting professional development is well–documented across several disciplines, the extent to which VBO has been utilized in M&E contexts is less clearly defined. In order to provide theoretically–grounded recommendations for the development and implementation of one organisation's innovative VBO impact evaluation scheme, this review examines recent evidence of VBO in M&E contexts and draws on VBO literature more generally to identify its advantages and challenges together with advice for enhancing its effectiveness. Based on this analysis, the paper highlights a number of practical issues that should be considered when VBO is being developed for M&E in development contexts, particularly where videos are being made by participants themselves. The value of VBO in responding to COVID–19 and reducing carbon emissions is also noted.

1. Introduction

While evidence of what people do in response to development interventions is a key facet of project impact evaluations, it can often be challenging (for reasons of access, budgets and, today, COVID–19) for evaluators to collect, in–person, a sufficient volume of good quality observational data. This paper discusses video–based observation (VBO) as a strategy that can facilitate analyses of behaviour during project impact evaluations.

The motivation for the review is first outlined along with the questions it is intended to address. The core of the paper consists of an analysis, with particular attention to monitoring and evaluation (M&E), of key themes in recent literature on the use of VBO in a range of disciplines. The paper concludes with concrete suggestions, informed by this review, for the design and implementation of a VBO approach to evaluating the impact of development interventions. While the evaluation of teaching is the primary area of concern here, the review, and the recommendations for practice emerging from it, are of relevance to the use of video–based impact evaluation more generally.

2. Background

The British Council is the UK's international organisation for cultural relations and educational opportunities. One of its strategic objectives is to 'develop a wider knowledge of the English language' and in relation to this it partners with organisations around the world – typically Ministries of Education – to deliver professional development for teachers of English in primary, secondary and tertiary settings. Interventions vary in scale, from smaller initiatives such as, for example, 'English for the Community' in Romania, which involves 150 teachers, to the National Teacher Training Programme in Egypt in which over 20,000 take part. The design of professional development interventions also varies, including both more conventional intensive workshop–based training programmes and, particularly more recently, an extended community of practice model (for a recent review of British Council projects adopting this model see Borg, Lightfoot & Gholkar (2020)). Project duration varies too, from six months to three years and sometimes even more. In terms of purpose, projects share the goal of improving the quality of English teaching and learning. The British Council's 'Continuing Professional Framework for Teachers'[2] (see Borg & Edmett, 2019 for a discussion) defines 12 professional practices (such as planning lessons, classroom management and assessing learning) that are seen to underpin effective teaching (and the specific areas of this framework that are addressed in individual projects will depend on local needs and the time available.

All British Council teacher development projects undergo an

---

* Corresponding author at: Klanec pri Kozini 21a, Kozina, 6240, Slovenia.
E-mail address: simon.borg@hvl.no.
[1] www.simon-borg.co.uk.
[2] See https://tinyurl.com/s4wrxit8.

evaluation. Evaluation is 'the assessment at one point in time of the impact of a project or programme and the extent to which stated objectives have been achieved' (Gosling & Edwards 2003, p.108). Impact concerns 'the long–term effects on identifiable populations or groups produced by a project or program' (Bamberger, Rugh, & Mabry, 2019, p.616) and 'the lasting or significant changes … in people's lives brought about by a given action or series of actions' (Roche 1999, p.21). Evaluation procedures will vary across projects but will typically draw on a range of qualitative and quantitative data from key stakeholders collected through, for example, surveys, focus group interviews, change stories written by teachers (Davies & Dart, 2005), lesson plans, student work and classroom observation.

While it is relatively straightforward to obtain evaluation data via surveys and other forms of self–reported data, the systematic collection and analysis of classroom observation data – a vital element in understanding project impact – presents several challenges. The number and geographical spread of teachers – for example, on the TEJAS project in Maharashtra, some 10,000 teachers are spread across an area of over 300,000 km^2 – typically mean (in combination with the modest size both of British Council project teams and the budgets available for evaluation work) that limited observations can be conducted in–person, and often only once at the start and end of an intervention. Lesson observation is also a sensitive and intrusive activity and one that is strongly (and negatively) associated with high–stakes performance evaluation in many of the educational systems where the British Council works. Gaining access to schools can, thus, be complicated. It is also the case that when project evaluators conduct fieldwork, the schools and teachers selected or who volunteer for observation are often considered to be excellent models, thus limiting the representativeness of results. Reactivity (Liang, 2015) is another challenge too, given that schools and teachers are often determined to demonstrate their best work to observers, leading to staged performances rather than naturalistic observations. Thus while classroom observation is recognised by the British Council as a significant source of insight into the impact of its professional development interventions, a range of factors related to workload (for project staff), budgets, access and educational cultures mean that the direct observation of teaching often contributes insufficiently to project evaluations. Thus, for example, the evaluation report for a recent teacher development project, where fewer than 0.2 % of the participating teachers were observed, noted in its recommendations that

> The goal of [the project] is to bring about change in the classroom and, in evaluating the programme, the observation of lessons in schools should therefore be a priority. While surveying teachers about what they do and which aspects of their work they need to develop has some value, direct assessment of their performance in the classroom will provide a more direct and informed understanding of such issues.[3]

In a review of the evaluations of six recent British Council teacher development projects adopting a community of practice model, Borg, Lightfoot & Gholkar, 2020) also noted that 'in terms of direct observations of … teacher performance in the classroom, the empirical evidence presented in evaluation reports was generally limited'.

In response to such concerns, the British Council is developing a video–based approach to lesson observations which can be used across projects to obtain more extensive evaluation data. The model of VBO in M&E envisaged by the British Council is outlined in Fig. 1 and this provides some reference points for the review of literature in this paper. Two key features of this model are (a) the role of participants in making and uploading recordings of their lessons and (b) the analysis of
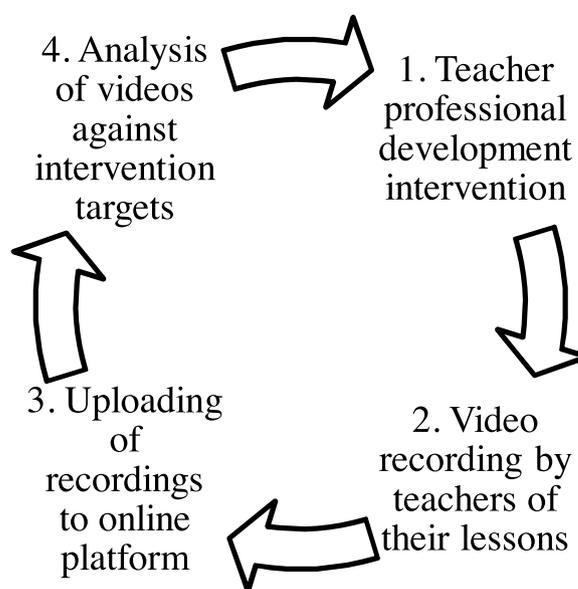

Fig. 1. VBO in M&E of British Council CPD work.

recordings through an online video tagging platform[4] . According to this model, teachers will take part in a professional development intervention that seeks to improve their teaching skills, and then, as part of the evaluation of that intervention, a sample of them will be asked (ideally at various points during a project) to make recordings of their own lessons (using their own devices) and to upload them to a web–based platform specifically designed for the analysis of videos. Raters will then access the videos and assess them using tagging tools built into the platform. The process will ideally also involve baseline (i.e. pre–intervention) lesson recordings too, but access to participating teachers is often not possible until the project has actually started.

Work is ongoing to incorporate into the tagging platform features requested by the British Council to support their VBO scheme, but the technicalities of the platform are beyond the scope of this paper, which is to provide theoretical and practical insights which can inform more specific decisions about the implementation of this envisaged model of VBO for project evaluation.

Five specific research questions will be addressed here:

1 RQ1: To what extent has VBO been utilised in M&E contexts?
2 RQ2: To what extent has VBO been utilised specifically for M&E on educational programmes?
3 RQ3: What advantages of VBO have been identified in the literature?
4 RQ4: What challenges associated with VBO have been identified in the literature?
5 RQ5: Based on the evidence available, what recommendations can be made for the effective use of VBO for M&E?

3. Method

As discussed by Erickson (2011), the use of video for social research has a long history; much has also been written about the use of video observation to support professional learning[5] (for example, Beisiegel, Mitchell, & Hill, 2018; Hockly, 2018; Quinn, Kane, Greenberg, & Thal, 2018; Mann, Crichton, & Edmett, 2020). Neither of these uses of video,

---

[3] The source of this quote is an internal report and further details cannot be disclosed to protect the identity of the project.

[4] The British Council will be using a platform called VEO – see www.veo.co.uk.

[5] This includes, for example, professionals recording themselves and reflecting on their videos or sharing the recordings with others who then provide developmental feedback.

though, was initially the focus of this literature review. Rather, the primary concern was VBO for the evaluation of interventions designed to bring about changes in behaviour. Several texts dedicated to M&E were first reviewed (Bamberger et al., 2019; Bell & Aggleton, 2016; Görgens-Albino & Kusek, 2009; Gosling & Edwards, 2003; Markiewicz & Patrick, 2016) in order to identify discussions in them of VBO. These searches were not productive; although observation was discussed as a data collection strategy, most texts did not even include 'video' or related terms in their index.

The second, more substantial, stage of the literature search employed a number of electronic databases, including Google Scholar, Scopus, Web of Science, Microsoft Academic and Semantic Scholar. The following search terms were used in various combinations across these databases: video, observation, teacher/teaching, monitoring and evaluation, project/programme evaluation, impact. Searches were limited to work in English and, after an initial review of search results, it was decided to include work published since 2015. While no significant material prior to this date was identified, and this cut-off point increased the currency of the review, it must be acknowledged that any earlier material not published in English was omitted.

It became clear quite early in the process that the volume of material addressing video–based teacher observation in M&E was very limited. For example, a Scopus search for items including 'video observation', 'teaching' or 'teacher' and 'evaluation' in the title, abstract or key words yielded four papers published since 2015, none of which involved the evaluation of programme impacts. A search using 'monitoring and evaluation', 'video' and 'teacher' also yielded no relevant results. It was necessary, therefore, to search using multiple combinations of key words across databases and to include all disciplines (the latter is a feature of this review which adds to its broader relevance to M&E across fields). Given that the anticipated volume of material was not going to be extensive, the criteria for inclusion in this review were not too stringent. Any sources that used VBO in the evaluation of a behavioural intervention were included, along with papers where video was used for research or professional learning and which had implications for the use of VBO in M&E contexts.

A final list of 25 sources were chosen for inclusion in this review. Most were research papers, along with some literature reviews and broader discussions of video–based inquiry. Most (10) came from health–related disciplines, with seven from education, four from organisational research and the remainder from childcare (2), psychotherapy and engineering (one each). The empirical work was international in scope, covering 13 different countries. Varying kinds of video recording equipment were also used in this body of work – iPads, smartphones, eye–glasses with built–in cameras, HERO action cameras, CCTV, video cameras on tripods and webcams.

Each article was read and key extracts transferred to a grid which contained the following headings: source, discipline, rationale for VBO, objectives, technology used, participants, methods, key results, challenges and 'other points'. The material for each of these themes was then subjected to closer thematic analysis and this led, for example, to the summaries of the benefits and challenges of VBO presented later in this paper.

3.1. VBO in M&E

Only three studies were identified where VBO was used for M&E. In several other papers (for example, Evans & Redmond, 2019; Diefenbacher, Sassenrath, Tatzel, & Keller, 2020), VBO was used to evaluate levels of compliance with regulations (for example, with hygiene protocols in factories or hospitals) but these were not considered to be impact evaluations because compliance was not being assessed in response to a prior training intervention.

Dyer et al. (2018) reported on the use of VBO in the evaluation of a training programme designed to improve the use of evidence–based practices by mentee nurses (specialising in maternal neonatal care) in India. This paper focuses on the process of developing and implementing a video–based M&E system rather than on the results of the M&E itself. Nurse mentors were first trained to run simulations and lead debriefing sessions with mentees. The mentors were then video recorded as they conducted these activities. The recording equipment was described simply as 'video cameras provided by the program' (p. 21). Video footage was uploaded to a computer, saved to USB drives, and couriered to the project headquarters where the data were uploaded to a secure server. A coding system was developed for the analysis of the videos and almost 500 hours of recordings were coded by two analysts. The paper concludes that VBO, as used in this study, has the potential to be applied at scale in low–resource settings. The use of ongoing video analyses to formatively inform decisions about programme design and content is also discussed. In terms of challenges, it was noted that 'the process of data collection, coding, and transferring large video files was labor intensive' (p. 25).

A second study (Lau, Chee, Ab Hamid, Leong, & Lau, 2019) where video was used for evaluation purposes also took place in a medical context, this time in Singapore. Video was used to evaluate the impact of a two–day cardiac life support training programme. Almost 450 medical and nursing students attended lectures and simulations. Teamwork performance during resuscitation simulations was evaluated quantitatively using video recordings (one before and one after the training). Two scales (the Clinical Teamwork Scale and the Communication and Teamwork Skills Assessment) were used as the basis of the analysis of the videos and the results indicated that on both scales trainees' performance improved significantly after the training. The simulations were recorded using video systems already present in the simulation laboratories.

In a study in Germany, Pletz and Zinn (2020) also used video recordings to formatively evaluate the impact of a new virtual training intervention for machine operators working in the field of mechanical and plant engineering. Participants (n = 13) first completed a virtual training programme; then they were asked to apply what they had learned to the use of the real machine. Both the training programme and the subsequent application of knowledge were video recorded (though details of how exactly the recording took place are not provided). Around 800 min of video were analysed and the results indicate that participants were able to effectively transfer learning from the virtual environment to the real world.

None of these examples are particularly close, in a disciplinary sense and in terms of the process, to the model of VBO for impact evaluation being proposed by the British Council. Dyer et al. (2018) shares some parallels in that recordings were made remotely (that is, the evaluators were not present) and mentors had some responsibility for ensuring recordings were completed and transferred to USB drives. In the other two cases, participants had no role in the making of the recordings. This initial analysis, then, suggested that the approach to VBO for impact evaluation being developed by the British Council, both in terms of participant roles and in the use of a web–based platform for analysis, is innovative.

3.2. Key themes in VBO

In this section I draw on a wider range of sources including, where they provide insights that are relevant to impact evaluation, accounts of VBO where it has been used for research and professional development purposes.

3.2.1. Uses of video

In the context of research, Whiting, Symon, Roby and Chamakiotis (2018) distinguish between four kinds of VBO:

- participatory video research — where participants are given control of the camera

- videography — where researchers film participants in the course of their work practices
- use of existing videos — analysis of existing videos not produced by researchers or participants
- video elicitation — the recording of video as stimulus for subsequent interviews.

These options are also relevant to impact evaluation contexts and that which is closest to the model envisaged by the British Council is the participatory approach. There is in fact a strand of M&E referred to as participatory M&E. This is defined as

a process through which stakeholders at various levels engage in monitoring or evaluating a particular project, program or policy, share control over the content, the process and the results of the monitoring and evaluation (M&E) activity and engage in taking or identifying corrective actions.[6]

A common approach to participatory M&E is that based on community video, where 'participatory videos are opportunities for communities to come together to tell their own stories, control their own narratives, and identify what is working for them and why'[7] (see also Lunch, 2006; Goodsmith & Acosta, 2011).

One interesting question here, then, regards the extent to which the approach envisaged by the British Council can be called participatory VBO. On the one hand, videos will be made and shared by teacher participants, using equipment available to them (according to Whiting et al., 2018 this approach is called 'distributed video study'). Teachers will also decide which lessons they want to record and make available for analysis. On the other hand, though, teachers may not have much ownership of the process; they will not be involved in the analysis and interpretation of the videos and will not expect to receive any feedback. Their role is to supply data which can be used for the purposes of M&E. This lack of fuller dialogic involvement may limit teachers' enthusiasm to engage in the process and such tendencies might be countered through some form of incentive, as suggested in the recommendations later in the paper.

Markiewicz and Patrick (2016) note that participation in M&E can be defined along a continuum from no participation, to provision of information, consultation, partnership and control. They also point out the distinction between *participation in evaluation* and *participatory evaluation*. Overall, the former of these terms is probably a closer description of the role that teachers will play in VBO as the British Council plans to use it for M&E. It may, of course, be possible for the British Council to consider ways in which this role can be made more participatory without losing sight of the M&E purposes VBO needs to fulfil.

Some of the sources reviewed here did involve participants in making their own videos. In one study (Allen & Hadjistassou, 2018), Swedish student teachers who were completing a practicum in Tanzania and Kenya recorded lessons on iPads and uploaded these so that their supervisors in Sweden were able to provide feedback. Also with professional development rather than M&E in mind, Perry and Boylan (2018) asked participants on a facilitator training initiative to record and share with other participants examples of their work in supporting teachers. In another example (Colliers, Coenen, Remmen, Philips, & Anthierens, 2019), medical practitioners in Belgium recorded their consultations with patients. This study is particularly relevant here as its purpose was to evaluate the use of VBO (in studies of doctor–patient interactions). It was highly participatory in the sense that practitioners were involved in design decisions throughout the process and provided input that shaped the

approach to VBO that was adopted. Various suggestions for improving the use of VBO emerged from this study and these are included in the section on 'Enhancing VBO' below.

3.2.2. Advantages of VBO

Several papers (see Box 1) highlight the advantages of VBO (particularly in relation to in–person observation – see James, Desborough, McInnes, & Halcomb, 2019 for a detailed comparison). The sources here include both remote and direct (i.e. with the observers present) uses of VBO. In some cases, such as Liang (2015) and Biery, Bond, Smith, Leclair, and Foster (2015), remote VBO was live and did not involve any recording.

The issues in Box 1 can be summarised in relation to three main issues: validity, reliability and flexibility.

1 Video is less intrusive and reduces the tendency of people to modify their behaviour (in the direction of what is believed to be desirable) when they are conscious of being observed. This is known as reactivity or the Hawthorne Effect (see Liang, 2015 for a discussion). The more authentic behaviour associated with VBO increases the validity of the data obtained. Also, a larger number of observations can be conducted and analysed, thus further contributing to the validity of conclusions.
2 Particularly when analysed quantitatively, video lends itself to review by multiple analysts, who can also compare their analyses and resolve discrepancies by reviewing the recorded material. The analysis of video by multiple raters allows for increased levels of reliability (i.e. consistency of analysis).
3 Video offers analysts more flexibility than in–person observation. No (often time–consuming and costly) travel is required and recordings can be analysed at any time. Video generates permanent records that can be reviewed in a focused manner (for example, less relevant extracts can be fast–forwarded) and studied more than once if necessary.

It is important to stress that simply employing VBO does not guarantee these benefits. Facilitating conditions must exist, such as good quality footage, effective coding systems and trained analysts, as discussed later in this paper. One advantage that VBO does guarantee, though, and which is, however, not highlighted in the literature, is that it can reduce the need for evaluators to travel (often by plane) and is thus good for the environment.

3.2.3. Challenges in VBO

The literature does, also, highlight various challenges associated with the use of VBO. These are summarised in Box 2 .

This list highlights a range of technical and ethical challenges in VBO. Technical problems, such as ineffective operation or placement of the recording equipment, will limit the quality of audio and video that are generated. Ethical issues, while central to all research and impact evaluation work, assume an even higher profile when video is involved given the potentially personal, sensitive and revealing nature of video recorded data. Thus, even when participants agree to be recorded, they may place limits on how the material can be used. For example, Colliers et al. (2019) found that GPs participating in their video–based study were happy for the recordings to be used for research purposes but less willing for the videos to be seen by peers and used for educational purposes. This example suggests that attitudinal barriers to VBO can be significant. In many contexts, participants will not be accustomed to having their routine or professional behaviours observed and/or recorded; this novelty, combined with concerns about how the recordings will be used, means that one key initial challenge in VBO can be securing participants.

Two additional issues noted in Box 2 both come from the field of education and are relevant to the model of VBO being developed by the British Council. The first relates to cases when videos are made and

---

Box 1
Advantages of VBO.

- Participants are less nervous than with in–person observations (Ault et al., 2019)
- Behaviours can be observed remotely (Ault et al., 2019; Biery, Bond, Smith, Leclair, & Foster, 2015)
- Less demands on observers' time in terms of travel and being able to analyse recordings more rapidly and at their convenience (Ault et al., 2019; Dagnaes-Hansen et al., 2018; Diefenbacher et al., 2020; Yanes et al., 2016)
- More valid results due to reduced reactivity among participants (Brotfain et al., 2017; Bui et al., 2018; Lebaron et al., 2018; Liang, 2015; Mac Mahon, Ó Grádaigh, & Ní Ghuidhir, 2019; McKay, Shaban, & Ferguson, 2020; Simons, Beltramo, Blalock, & Levine, 2017)
- Recordings can be reviewed by multiple analysts, improving reliability (Christianson, 2018; Dagnaes-Hansen et al., 2018; Lebaron et al., 2018; Mesman, 2020)
- Video provides a permanent record that can be returned to (Christianson, 2018; Lebaron et al., 2018)
- Video provides a detailed record of behaviour (Lebaron et al., 2018) including of classroom behaviour (Allen & Hadjistassou, 2018)
- Video allows the duration of behaviours and events to be analysed (Christianson, 2018)

Box 2
Challenges in VBO.

- For participant–made videos, recordings may not be produced and shared to the agreed timetable (Allen & Hadjistassou, 2018)
- Key decision–makers, such as managers, may object to the use of video cameras to record the work of their employees (Ault et al., 2019)
- The quality of video and audio may limit what can be seen and accurately heard (Ault et al., 2019; Mac Mahon et al., 2019)
- Individuals operating the recording equipment may lack the technical expertise required (Biery et al., 2015)
- The positioning of the video equipment may limit what is actually recorded (Brotfain et al., 2017; Gold & Windscheid, 2020; Mac Mahon et al., 2019; Yanes et al., 2016)
- Participants may have concerns about data privacy (Biery et al., 2015; Lebaron et al., 2018) and about how the recordings will be used (Colliers et al., 2019)
- In classrooms, video footage alone does not provide sufficient context for interpreting behaviour (Molbæk & Kristensen, 2019)

---

shared by participants. Allen and Hadjistassou (2018) note that the student teachers in their study did not upload videos according to the agreed timetable and concluded that effective remote observation of teaching was 'heavily contingent on student teacher discipline and organization in submitting videos' (p. 362). This will be relevant to the British Council's VBO scheme, which will depend on the timely uploading of videos by participating teachers. The second issue, regarding the lack of a broader context for video–recorded behaviours (Christianson, 2018), is also important and suggests that the analysis of video should, unless supported by additional explanatory data, focus on observable behaviour and not the cognitions (i.e. beliefs, thinking and intentions – see Borg, 2006) that underpin what teachers do.

3.2.4. Enhancing VBO

It is clear from the discussion so far that VBO offers several advantages for impact evaluations but that there are also various challenges that the use of video observations creates. This section reviews recommendations made in the literature for enhancing the quality of VBO. Box 3 summarises key points and these are discussed below (see also Colliers et al. 2019, p.5 for detailed guidance on setting up a video–based study).

Eliciting the views regarding VBO of participants and other stakeholders can provide insights that facilitate the process of using video in research and M&E contexts. Colliers et al. (2019) argue that 'it is essential to understand which conditions must be met for the participants to enhance participation and successful implementation of a video observation study' (p.2). They also illustrate the value of engaging with target VBO participants from the outset and using participant input (for example, concerns about privacy and demands on their time) to inform decisions about how VBO will be used. An understanding of stakeholder views can also limit unexpected problems; Ault et al. (2019), in their study of remote video observations of student teachers, found that one school principal did not give permission for the recording equipment to be used in the classroom and conclude that obtaining permission in

advance is desirable in work of this kind. Organisations and researchers wishing to implement VBO, therefore, need to anticipate and find out about potential concerns, and seek ways of addressing them. Another example of this is provided by Lebaron et al. (2018) in the field of organisational research, where it is reported that organisations often oppose video–based studies because they are concerned about privacy and about how data will be stored, accessed and used. Addressing such concerns by, for example, explaining how data will be stored securely and only used for agreed purposes, is thus important.

Many of the issues just discussed, of course, relate more broadly to ethical issues and, as noted earlier, it is important that VBO meets the requirements of any ethical (and legal) regulations that apply (in their five steps for conducting a video study, James et al., 2019 dedicate step 3 to 'legal and ethical issues'). In the context of education, for example, stringent conditions often apply to video recording children in classrooms and it is important that such conditions are met.

The preparation of participants for VBO is also particularly relevant here given that, as part of the British Council's scheme, teachers will be invited to record themselves. Allen and Hadjistassou (2018), for example, explain that before the Swedish student teachers in their study travelled to Africa, they were 'briefly trained on video recording for educational research purposes, such as positioning the recording device in the classroom' (p.356). Whiting et al. (2018) discuss the preparation of the participants in their VBO study in more detail:

Videocam, instructions for taking part in video study, and consent form are posted to participant in preparation for briefing … We used a mixture of one–to–one and group briefings dependent on participant availability. A few were face–to–face (F2F), but most were conducted via Skype using the Share Screen option to present the PowerPoint briefing and answer participants' questions. We outlined ethical guidance of "what not to film'' in the briefing presentation. This included: anything confidential, sensitive, or highly personal;

Box 3
Enhancing VBO.

- Understand participants' and other stakeholders' concerns about VBO and take steps to address these (Colliers et al., 2019; Diefenbacher et al., 2020; Lebaron et al., 2018)
- Comply with legal and ethical requirements that regulate research generally and the use of VBO specifically (especially data privacy, storage, access and use) (Colliers et al., 2019; Waller & Kaplan, 2018)
- Prepare participants for VBO, including, where necessary, through training for making and uploading recordings (Allen & Hadjistassou, 2018)
- Observe several times to increase coverage and reduce reactivity (Molbæk & Kristensen, 2019)
- Develop a systematic coding scheme for the analysis of videos and provide coder training to maximise inter–rater reliability (Ault et al., 2019; Dyer et al., 2018; Lau et al., 2019; Waller & Kaplan, 2018)

---

children (unless participant's own and both parents agreed); other people unless in a public place where they might reasonably expect to be observed or were people they knew who consented to being filmed; in shopping centers or areas with high security status; and while driving or cycling. (p. 325)

It is clear that VBO will be enhanced when participants are prepared, both technically and in terms of advice regarding the kinds of video–recording behaviours that are appropriate and inappropriate.

Another way of enhancing VBO, as noted by Molbæk and Kristensen (2019), is to observe teachers more than once, and this seems particularly feasible where participants are providing the videos themselves. Multiple observations also provide an increased volume of data and a reduction in reactivity (as teachers become accustomed to being recorded). It must be acknowledged, though, that multiple observations also increase the demands being placed on the observees and may make them less inclined to participate.

The final point in Box 3, regarding coding and coder training, is repeatedly stressed in quantitative VBO studies. Multiple raters are often involved in such work and papers typically provide substantial detail both of the coding framework and of the steps that were taken to prepare raters to use it in a consistent manner. Waller & Kaplan (2018, p.502) propose four key elements in VBO studies that focus on behaviour and two of these relate to 'coder selection and training' and 'coding scheme and interval[8]' and in their discussion of remote observation of teaching, Ault et al. (2019) discuss in detail how observer training took place:

The university supervisors contributed to the development of the observation form, so they were highly familiar with each item. They developed the operational definitions and discussed how each item would be scored during the observations. They then conducted three practice observations in local classrooms using student teachers not associated with the study. The observers independently scored the form and debriefed following each observation to compare scores. If disagreements occurred, the observers refined the operational definitions. The third observation resulted in acceptable interobserver agreement reliability of 80 % or better on each section of the form, after which the observers began observations for the study. In addition, following each observation conducted for the study, the observers discussed any disagreements and refined definitions of items prior to the next observation. (p. 127)

In common with many studies where multiple raters of video were involved (for example, Dagnaes-Hansen et al., 2018; Dyer et al., 2018; Lau et al., 2019), this paper also cites statistics for inter–rater reliability. Evidence that the analysis of videos has been conducted robustly is an essential element in enhancing the perceived quality of VBO. This applies to M&E as well as to research and it is thus necessary to be explicit,

when M&E results based on VBO are reported, about preparations for and the conduct and reliability of the analysis. It is also vital that the coding framework being applied (and which is typically embodied in a structured observation tool) is theoretically sound. Much advice on such matters is available in the research methods literature (for example, Newby, 2010; Cohen, Manion, & Morrison, 2011; Bryman, 2016).

4. Lessons learned

RQ1 and RQ2 defined at the start of this review asked about the extent to which VBO has been used in M&E both generally and in education. It is clear that, while observation is widely discussed in the M&E literature as a data collection tool, little evidence seems to exist (beyond the community–oriented use of participatory video) that VBO has been applied in M&E settings. This is true not only in education but across disciplines more broadly. Even in medical settings, where video–based observation is well–established, very few examples were found where M&E (rather than compliance with health protocols) was the main purpose of such observation. This was a surprising finding here and this review is thus timely given the increased interest in remote VBO for M&E that can be expected in global development contexts as a result of the limitations COVID–19 has placed on travel and the ability of evaluators to conduct observational fieldwork in person. Research–oriented discussions of VBO are much more widely available, including special issues of journals dedicated to video–based studies (Lebaron et al., 2018; Mesman, 2020), while VBO for professional learning also features in the literature.

RQ3 and RQ4 focused respectively on the advantages and challenges associated with VBO and these were summarised in turn in Box 1 and Box 2 above. In relation to the British Council's intended use of VBO, the most relevant advantages are that observational data from teachers around the world will become available via a web–based platform through which multiple raters will be able to evaluate project impacts. In terms of challenges, all of those listed in Box 2 are relevant to the British Council scheme. It should be anticipated, for example, that attitudes to VBO among teachers and their employers will not always be positive, that technical expertise in producing good quality videos among the teachers will vary, and that concerns will arise about data privacy and access.

It is clear that the M&E scheme the British Council is developing is innovative, particularly in the use of the web–based VEO platform through which uploaded videos can be tagged and rated by multiple coders against pre–defined coding frameworks. However, the effectiveness of the scheme more generally will not depend solely on how well the web–based platform works. Several other factors will come into play and to conclude this review I will draw on the literature review to address RQ5 and to make recommendations for the development, trialling and implementation of VBO in M&E. These will be relevant not only to the British Council's work but to the use of VBO in evaluation contexts by a range of organisations and in different disciplines.

---

[8] 'Interval' here refers to decisions about how often observed behaviours will be coded – for example, continuously or at intervals.

1 A standardised package for target participants will need to be developed that includes, for example, an invitation to participate in VBO, a description of the process and what they would be required to do, guidance on how they should (and should not) do it, advice on making and uploading videos, a consent form, and links to the hosting organisation's 'VBO webpages' where all the relevant information is available. A similar package will be required for organisational leaders given the important gate-–keeping role they play.

2 Training will also be required for the hosting organisation's 'VBO project managers' who will be responsible, for example, for promoting the scheme locally, consulting with stakeholders and monitoring participant progress and following up cases who have not adhered to the agreed timetable.

3 Participants may be given the opportunity to record and upload a short video as part of their training to use the new VBO system. This would allow any potential problems, for example, with video or audio quality, to be addressed before the actual behaviours to be evaluated are recorded.

4 Decisions will also need to be made about the length of the re-cordings participants are asked to share. In the context of teaching, full lessons may be more authentic than shorter re-cordings that illustrate a particular aspect of teaching; shorter videos may present less of a challenge to teachers and will be easier to upload and analyse. Various options should be explored as part of the trialling of any remote VBO scheme.

5 Potential objections or obstacles to participating in VBO should be anticipated and as far as possible addressed in the 'participant VBO package'. Initial consultation with, for example, target teachers, educational authorities, local British Council offices (and possibly even local communities) can provide insight into potential concerns regarding VBO (for example, data costs asso-ciated with uploading videos or limitations on the participation of females) as well as practical issues such as the kinds of recording equipment participants have access to.

6 It is useful to think about ways of incentivising participation in VBO. Although the focus of this discussion was M&E, this does not rule out the possibility that participants might be offered some feedback, if this were something that might motivate them to take part. Some form of certification for teachers who contribute videos may also be an attractive option in certain contexts.

7 Given the global nature of the British Council's teacher devel-opment work, it will be necessary to obtain country–level advice on relevant legal and ethical issues that VBO will need to address and how best to accomplish this in different contexts.

8 The observation tool and the coding framework for analysing uploaded videos will be built into the web–based video tagging platform. An observational tool is currently being developed for this purpose and will be accompanied by supporting documents for users. These are all important elements of a good quality VBO scheme. It will be important to ensure that the VBO tool and the coding framework used with it are theoretically robust (i.e. that the principles underpinning them can be articulated with refer-ence to existing theory from the observational research methods literature and work on the observation of teachers more generally).

9 Evaluations of video-based lessons will be enhanced when raters have access to additional contextual information about the teaching. Teachers, for example, may thus be asked to provide, along with the video, a short description of where the lesson fits into the curriculum, a lesson plan, and a short profile of the learners.

10 Raters responsible for assessing the videos will need to be given appropriate training to familiarise them with the observation tool and the coding framework and to enable a team of raters to use these tools consistently. Uploaded videos should be evaluated by at least two raters and inter–rater reliability statistics should be cited along with M&E results based on VBO.

11 As presently conceived, the British Council's approach to VBO requires teachers to participate rather than be participatory. There will be value in considering what realistic steps might be taken to make the VBO process a more participatory one for participants (and their organisations).

12 Decisions will also need to be made about when and how many times an individual participant will take part in VBO. What is feasible will vary across projects and contexts; the goal, though, should be to provide a sound empirical basis for judgements about the impact of interventions on what participants do. Single exit observation measures (for example of teacher classroom performance) are weak in this respect.

13 Finally, it is important to remember that while VBO provides insight into what participants do and say, it does not allow for conclusions about the thinking and motivations that underpin behaviour. To address such issues, additional data collection tools such as, for example, interviews or written reflective commen-taries by participants, need to be used.

In conclusion, it is also relevant to note that as a result of COVID–19 many routine activities in educational and business organisations are being conducted online. Teaching is being delivered via Zoom, for example, while business meetings now routinely take place on Teams. In some ways these developments facilitate the digital capture of behav-iour, given that video conferencing apps come with the ability to record (though the limiting impact of online work on interpersonal interactions – for example, between teachers and students – must also be recognised). The benefits, challenges and recommendations for the use of VBO in M&E made in this paper, though, will apply generally irrespective of whether the behaviour being analysed is taking place online or in–person.

Author statement

The author was the sole contributor to all aspects of the paper.

Funding declaration

Declaration of Competing Interest

The author reports no declarations of interest.

References

Allen, C., & Hadjistassou, S. (2018). Remote tutoring of pre-service EFL teachers using ipads. *ELT Journal, 72*(4), 353–364. https://doi.org/10.1093/elt/ccy016

Ault, M. J., Spriggs, A. D., Bausch, M. E., & Courtade, G. R. (2019). Evaluation of remote versus face-to-face observation of teacher candidates in an alternative certification program. *Rural Special Education Quarterly, 38*(3), 124–136. https://doi.org/10.1177/8756870519861030

Bamberger, M., Rugh, J., & Mabry, L. (2019). *Realworld evaluation: Working under budget, time, data, and political constraints* (third ed.). Thousand Oaks, CA: Sage.

Beisiegel, M., Mitchell, R., & 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. https://doi.org/10.1177/0022487117705096

Bell, S., & Aggleton, P. (2016). *Monitoring and evaluation in health and social development: Interpretive and ethnographic perspectives.* London, England: Routledge.

Biery, N., Bond, W., Smith, A. B., Leclair, M., & Foster, E. (2015). Using telemedicine technology to assess physician outpatient teaching. *Family Medicine, 47*(10), 807–810.

Borg, S. (2006). *Teacher cognition and language education: Research and practice.* London, England: Continuum.

Borg, S., & Edmett, A. (2019). Developing a self-assessment tool for English language teachers. *Language Teaching Research, 23*(5), 655–679.

Borg, S., Lightfoot, A., & Gholkar, R. (2020). *Professional development through Teacher Activity Groups: A review of evidence from international projects*. London, England: British Council.

Brotfain, E., Livshiz-Riven, I., Gushansky, A., Erblat, A., Koyfman, L., Ziv, T., … Borer, A. (2017). Monitoring the hand hygiene compliance of health care workers in a general intensive care unit: Use of continuous closed circle television versus overt observation. *American Journal of Infection Control, 45*(8), 849–854. https://doi.org/10.1016/j.ajic.2017.03.015

Bryman, A. (2016). *Social research methods* (5th ed.). Oxford: Oxford University Press.

Bui, A. H., Guerrier, S., Feldman, D. L., Kischak, P., Mudiraj, S., Somerville, D., … Leitman, I. M. (2018). Is video observation as effective as live observation in improving teamwork in the operating room? *Surgery, 163*(6), 1191–1196. https://doi.org/10.1016/j.surg.2018.01.019

Christianson, M. K. (2018). Mapping the terrain:The use of video-based research in top-tier organizational journals. *Organizational Research Methods, 21*(2), 261–287. https://doi.org/10.1177/1094428116663636

Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education* (7th ed.). London, England: Routledge.

Colliers, A., Coenen, S., Remmen, R., Philips, H., & Anthierens, S. (2019). Looking inside the out-of-hours primary care consultation: General practitioners' and researchers' experiences of using video observations as a method. *International Journal of Qualitative Methods, 18*, Article 1609406919879341.

Dagnaes-Hansen, J., Mahmood, O., Bube, S., Bjerrum, F., Subhi, Y., Rohrsted, M., & Konge, L. (2018). Direct observation vs. Video-based assessment in flexible cystoscopy. *Journal of Surgical Education, 75*(3), 671–677.

Davies, R., & Dart, J. (2005). *The 'Most Significant Change' and MSC technique. A guide to its use*. Retrieved from http://www.mande.co.uk/docs/MSCGuide.pdf.

Diefenbacher, S., Sassenrath, C., Tatzel, J., & Keller, J. (2020). Evaluating healthcare workers' hand hygiene performance using first-person view video observation in a standardized patient-care scenario. *American Journal of Infection Control, 48*(5), 496–502. https://doi.org/10.1016/j.ajic.2019.11.032

Dyer, J., Spindler, H., Christmas, A., Shah, M. B., Morgan, M., Cohen, S. R., … Walker, D. (2018). Video monitoring a simulation-based quality improvement program in Bihar, India. *Clinical Simulation in Nursing, 17*, 19–27. https://doi.org/10.1016/j.ecns.2017.11.007

Erickson, F. (2011). Uses of video in social research: A brief history. *International Journal of Social Research Methodology, 14*(3), 179–189. https://doi.org/10.1080/13645579.2011.563615

Evans, E. W., & Redmond, E. C. (2019). Video observation of hand-hygiene compliance in a manufacturer of ready-to-eat pie and pastry products. *International Journal of Environmental Health Research, 29*(6), 593–606. https://doi.org/10.1080/09603123.2018.1558183

Gold, B., & Windscheid, J. (2020). Observing 360-degree classroom videos – Effects of video type on presence, emotions, workload, classroom observations, and ratings of teaching quality. *Computers & Education, 156*, Article 103960. https://doi.org/10.1016/j.compedu.2020.103960

Goodsmith, L., & Acosta, A. (2011). *Community video for social change: A toolkit*. Minneapolis, MN: American Refugee Committee International.

Görgens-Albino, M., & Kusek, J. Z. (2009). *Making monitoring and evaluation systems work: A capacity development toolkit*. Washington, DC: World Bank.

Gosling, L., & Edwards, M. (2003). *Toolkits: A practical guide to assessment, monitoring, review and evaluation*. London, England: Save the Children Fund.

Hockly, N. (2018). Video-based observation in language teacher education. *ELT Journal*. https://doi.org/10.1093/elt/ccy022. ccy022-ccy022.

James, S., Desborough, J., McInnes, S., & Halcomb, E. (2019). Strategies for using non-participatory video research methods in general practice. *Nurse Researcher, 27*(2), 32–37. https://doi.org/10.7748/nr.2019.e1667

Lau, Y., Chee, D. G. H., Ab Hamid, Z. B., Leong, B. S.-H., & Lau, S. T. (2019). Interprofessional simulation–based advanced cardiac life support training: Video-based observational study. *Clinical Simulation in Nursing, 30*, 16–24. https://doi.org/10.1016/j.ecns.2019.03.001

Lebaron, C., Jarzabkowski, P., Pratt, M. G., & Fetzer, G. (2018). An introduction to video methods in organizational research. *Organizational Research Methods, 21*(2), 239–260. https://doi.org/10.1177/1094428117745649

Liang, J. (2015). Live video classroom observation: An effective approach to reducing reactivity in collecting observational information for teacher professional development. *Journal of Education for Teaching, 41*(3), 235–253. https://doi.org/10.1080/02607476.2015.1045314

Lunch, C. (2006). *Participatory video for monitoring and evaluation*. Retrieved from https://insightshare.org/wp-content/uploads/2017/05/PV-for-ME-Experiences-with-the-MSC-approach-English.pdf.

Mac Mahon, B., Ó Grádaigh, S., & Ní Ghuidhir, S. (2019). Super vision: The role of remote observation in the professional learning of student teachers and novice placement tutors. *TechTrends, 63*(6), 703–710. https://doi.org/10.1007/s11528-019-00432-z

Mann, S., Crichton, R., & Edmett, A. (2020). Evaluating the role of video in supporting reflection beyond inset. *System, 90*. https://doi.org/10.1016/j.system.2019.102195

Markiewicz, A., & Patrick, I. (2016). *Developing monitoring and evaluation frameworks*. Thousand Oaks, CA: Sage.

McKay, K. J., Shaban, R. Z., & Ferguson, P. (2020). Hand hygiene compliance monitoring: Do video-based technologies offer opportunities for the future? *Infection Disease & Health, 25*(2), 92–100. https://doi.org/10.1016/j.idh.2019.12.002

Mesman, J. (2020). Video observations of sensitive caregiving "off the beaten track": Introduction to the special issue. *Attachment & Human Development*, 1–9. https://doi.org/10.1080/14616734.2020.1828511

Molbæk, M., & Kristensen, R. M. (2019). Triangulation with video observation when studying teachers' practice. *Qualitative Research Journal, 20*(1), 152–162. https://doi.org/10.1108/QRJ-07-2019-0053

Newby, P. (2010). *Research methods for education*. Harlow: Pearson Education Limited.

Perry, E., & Boylan, M. (2018). Developing the developers: Supporting and researching the learning of professional development facilitators. *Professional Development in Education, 44*(2), 254–271. https://doi.org/10.1080/19415257.2017.1287767

Pletz, C., & Zinn, B. (2020). Evaluation of an immersive virtual learning environment for operator training in mechanical and plant engineering using video analysis. *British Journal of Educational Technology, 51*(6), 2159–2179. https://doi.org/10.1111/bjet.13024

Quinn, D. M., Kane, T. J., Greenberg, M. L., & 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*(529), - 558.

Roche, C. (1999). *Impact assessment for development agencies: Learning to value change*. Oxford, UK: Oxfam.

Simons, A. M., Beltramo, T., Blalock, G., & Levine, D. I. (2017). Using unobtrusive sensors to measure and minimize hawthorne effects: Evidence from cookstoves. *Journal of Environmental Economics and Management, 86*, 68–80. https://doi.org/10.1016/j.jeem.2017.05.007

Waller, M. J., & Kaplan, S. A. (2018). Systematic behavioral observation for emergent team phenomena: Key considerations for quantitative video-based approaches. *Organizational Research Methods, 21*(2), 500–515. https://doi.org/10.1177/1094428116647785

Whiting, R., Symon, G., Roby, H., & Chamakiotis, P. (2018). Who's behind the lens?:A reflexive analysis of roles in participatory video research. *Organizational Research Methods, 21*(2), 316–340. https://doi.org/10.1177/1094428116669818

Yanes, A. F., McElroy, L. M., Abecassis, Z. A., Holl, J., Woods, D., & Ladner, D. P. (2016). Observation for assessment of clinician performance: A narrative review. *BMJ Quality & Safety, 25*(1), 46. https://doi.org/10.1136/bmjqs-2015-004171

Simon Borg has been involved in teaching and teacher education for over 30 years. His academic areas of expertise are teacher cognition and teacher education. His current work as an educational consultant focuses on the design, implementation and evaluation of teacher professional development programmes and policies in a range of international contexts. Details of his work are available at http://simon-borg.co.uk/.