Visual Narrative and Temporal Relevance: Segueing Instant Replay into Live Broadcast TV

Mark Perry
Brunel University

Mathias Broth
Linköping University

Arvid Engström
Liveling AB

Oskar Juhlin
Stockholm University

Professional production of live TV combines real-time and recorded video into a single broadcast stream. In "live" TV, non-live "instant replay" footage can help viewers to make sense of what has just happened. This article shows how multi-person TV production teams assemble timely and relevant instant replays that can be seamlessly combined with real-time footage during live broadcasts. Detailed interaction analysis demonstrates how this work is dependent on coordinated practices, and how team members achieve this by orienting to narrative concerns across multiple temporalities to produce topically useful instant replays, displaying clip relevance, and help segueing transitions between the ongoing action and replay. We conclude by examining the interrelationships between the sequential flow of visual content, the role of talk in mediating time-shifted visual alignments, and how members make their work visible and accountable to one another and to their intended audience.
Keywords: live television, social interaction, accountability, camera, media production, sports TV

INTRODUCTION

Live multicamera broadcasts show a rich picture of activity from different angles to create variations in tempo and emotional atmosphere that enliven the visual imagery and provide an enthralling televisual experience. Nevertheless, even well-produced multicamera productions cannot visually explain what is happening in the light of what has already happened, and instant replay—the selection and (re)use of just-recorded video material during the live broadcast—is often used to achieve this, drawing from camera angles that may not have been shown in the live broadcast. In this, we contribute to a small body of interactional work on television replays (Camus 2015; Engström et al. 2010) and, more generally, follow a recent theme of enquiry within interaction studies that explores the ways that people work with images as an integral and routine aspect of their daily work (e.g., Broth, Laurier, and Mondada 2014).

The workplace studies program pursued in this research is informed by ethnomethodology and conversation analysis, and it explores the constitutive practices of professional activities (Garfinkel 1986; Heath, Knoblauch, and Luff 2000; Luff, Hindmarsh, and Heath 2000). Its focus often lies on the ways in which tools and technologies feature in social action, unpacking situations in and through which participants themselves use and interact with artifacts in emergent activities. It differs from other forms of technology studies, which typically focus on meaning, representation, and the social construction of tools and artifacts, in its concern with the practical accomplishment of various professional activities. Analytically, it is concerned with the social and interactional organization of visual, vocal, and tactile aspects of human conduct (e.g., Goodwin 1994; vom Lehn, Heath, and Hindmarsh 2001). From this perspective, social order is ongoingly accomplished in and through witnessably ordered practices (e.g., Heritage 1984; Popova 2018). In this it takes an explicitly *interactionist* perspective to understanding social conduct, as it is only through a close examination of members' interactions that we can begin to understand their social organization and get insights into how behavior is both shaped by and maintains social structures (Heritage 1984). Our work here extends the body of work in workplace studies with its particular focus on temporovisual arrangements and the place of sense-making in image work.

Examining the threading together of visual image streams that are temporally separated, yet topically connected, under real-time conditions, offers us the opportunity to see how image work is understood, performed and communicated by members of a television crew. What makes image work—under these conditions—particularly interesting is that video images form the topic, means of communication between, and final output of the participants' work (Perry et al. 2009). With this in mind, the people working on the video content have two audiences: the audience of television viewers, but also their colleagues. In the same way that Mondada (2003) discusses a "double production" of views, as surgeons switch between endoscopic and external camera views when filming live video of internal surgery to produce relevant

sequential images of the ongoing action at the same time as creating intelligible images as records for the remote "witnessing audience," the TV production team also make their own image work on the video broadcast accountable to the viewers, while simultaneously making the actors in the setting accountable for their actions. Our article thus extends Mondada's work in that professional TV production teams are not only concerned with the production of sequential and intelligible visual records of the activity as it happens, but are able (and expected) to produce content that segues real time images with historic content to further enrich and accomplish a particular intelligibility of the real-time visual record.

This article is therefore an attempt to understand better how and why live broadcasts are made in the ways that they are, and more specifically, how live and non-live video footage is produced and sewn together through social interaction to create meaningful imagery within a coherent narrative structure. We will also see that (and *how*) images are produced in different ways and for different purposes (cf. Mondada 2003). To clarify the various ways that *time* is used in TV broadcasts, we distinguish between *real-time*, by which we refer to the ongoing sequential actions as they occur, and *live*, by which we refer to the broadcast, which may include both images that are broadcast as they occur in real-time as well as historically recorded images that are produced and broadcast alongside the real-time material.

Given the commonplace use of instant replay in television, we use the context of live sports as a "perspicuous setting" (Garfinkel 2002:118) for discussing the use of live and recorded imagery. Much of media studies' focus on live visual media within sports TV has been on the final product of the editing process — the video broadcast — semiotically read as a "text" to interpret the meanings that underpin it. This can be interesting and revealing about the normative orientations involved in making sequences of images and about the interpretations that may be made of the visual imagery and narrative created (cf. Jayyusi 1988), but this post-event analysis necessarily must ignore the contingent creative process, as well as the constraints on image production and the production practices that underlie the generation of these images. Gruneau (1989:216) makes this point forcefully about the production of sports on television, arguing that a focus on reading the broadcast program as a text will "downplay the political and economic limits that operate as a context for televised sport production, and it has all but ignored analysis of the actual technical and professional practices, the labour process, involved in producing sport for television." Our topic of concern falls squarely with the small but developing area in the cultural and media studies literature, known as production studies (e.g., Mayer, Banks, and Caldwell 2009) or critical media industry studies (e.g., Havens, Lotz, and Tinic 2009) of which television production forms a small component. This set of disciplines investigates the impact of micro-level production practices on media production, albeit addressing a largely different set of foci (the influence of economic, regulatory, and institutional forces on cultural output) to that covered by our research. Nevertheless, a small number of studies have examined the production practices of televised sport

in detail (e.g., Barnfield 2013; MacNeill 1996; Silk, Slack, and Amis 2000; Stoddart 2006; Williams 1977), although the mechanics of image production has barely been touched on (but see Camus 2015).

It is useful to begin by drawing a parallel between the experience of being present at a sports stadium and watching sport on television. Despite their common topic, these are strikingly different experiences, and it has been argued that this is largely due to the use of the mediating role of visual imagery in television production. While live television allows "broadcasting events exactly when and as they happen" (Lohr 1940:52), this perhaps oversimplifies the nature of televisation and it is not quite, as Dunlap Jr. (1948):8) states, that live viewers "see politics as practiced, sports as played, drama as enacted, news as it happens, history as it is made." In this respect, television does not provide a simple simulation of presence for its remote audience. Live events are mediated (through video technology), and in its mediation, the event is represented through selected shots that can amplify or conceal aspects of the unfolding action (whether deliberate or unintended, cf., e.g., Broth 2008; Camus 2015; Goldlust 1987; Heath and Luff 1992; Mondada 2009). Auslander (1999:153) describes this process as "mediatization," and cites Connor (1996) in showing how a number of factors can impact on the experience of *liveness* in mediatized events: "The intense 'reality' of the performance is not something that lies behind the particulars of the setting, the technology and the audience; its reality lies in all of that apparatus of representation." Thus, we do not experience the excitement of televisual liveness purely from the event itself, but through the sociotechnical machinery by which it is rendered.

There is a long history of exploration into the live experience of sport in the literature on film and media studies and communication studies, and how this experience is delivered through different perspectives and media. In seminal work by Hastorf and Cantril (1954) examining different audience perceptions of an American football match, the authors concluded that there was no single "game" that existed as a commonly referenced reality which its audience merely "observed" (both when viewed live, and from recorded video). The claim was made that viewers deploy selective perception in complex circumstances to see a different game to that of the opposing team supporters. This itself sits within an even longer historical phenomenological perspective through Merleau-Ponty's (2012) clarifications on the relation between the objective world and the experienced world, in which the indeterminacies and ambiguities of the visible world are synthesized into new structures: perceived reality is effectively created through perceptual interpretation and affective judgment that is subjectively held to be true by each observer. Although Hastorf and Cantril present a very different study to our own approach here, their work resonates with our understanding of what TV production teams are trying to achieve: a narrative that shows and tells a common story of gameplay and highlights important features of the action, but which offers scope for sufficient breadth of interpretation for the viewer — to see it as a fan of either team, or of none. In this we ascribe to Altheide

and Snow's (1978) admonition that concerns of television and sports are some-what different, and that the logic and goal of television is to reframe the televised experience to map onto the viewers' interests and expectations, for commercial gain. While there may be different reasons for broadcasting and viewing sports, the game for fans is a matter of seeing the action, skill, and the eventual winning outcome. While this previous work does not show how such production work happens, the literature sets the scene for our own work on the mechanics of creating visual imagery in live sports that supports its viewers in making sense of the developing action.

This use of visual representations in mediating live events is not always intended to refocus, distort, or misrepresent events so as to provide a managed and "spun" dramatization of events, as commonly seen through the analytic lens of media or television studies and critical media theory. The application of technology in live television can also be used to support viewers in making sense of the reasons for previous and currently occurring events, by providing multiple viewing angles on events. For it to be relevant to the viewer, instant replays need to be reflexively tied to events that are unfolding live. In live televised sport, replaying different angles of events can lead to very different impressions of game play to that initially broadcast. Alongside the different live camera angles and verbal commentary, instant replay sequences can contribute to the viewer's understandings of the developing game itself, provide reasons (good or bad) for umpiring decisions, or display the skills and emotional involvement of individual players. Nevertheless, how this mediatization of events — that is, "the process of recording live events to be replayed at different points in time or space" (Morris 2008:59) — occurs in a live broadcast is not trivial. As shown by Camus (2015) and ourselves (Engström et al. 2010), the production processes that enable this interweaving of real-time and recorded content in real time is carefully managed in order to create an interesting, meaningful, and visu-ally compelling broadcast. This involves the interpretation of game play, location of appropriate and relevant video content, and detailed and often intricate technical work to create seamless and visually appealing video transitions. All of this must take place within an often rapidly changing game context. Given its live broadcast status, the production and use of replay footage is extremely time-dependent, with no possibility of taking time out, or re-doing the work. To understand how work can be achieved under these considerable demands and constraints requires us to take a close look at the production process, and the mechanics through which it operates.

Unpacking live TV demands that we take a careful look at how footage is aligned and combined to produce what is normally referred to in its totality as a "live" broad-cast, even though it may include some visual material that is *not* temporally concur-rent with the game (cf. Camus 2015; Marriott 2007). In order to do this, we focus here on the production processes that underlie the synchronization of real-time and recorded video in professionally produced broadcast television. We use empirical

data to show how the collaborative process of meshing relevant recorded footage into real-time footage is performed.

IMAGE PRODUCTION IN LIVE SPORTS TV

Replay has a relatively short history in live television, with a beginning in sport productions from the 1970s, when replay operators began to work with analogue tape machines (Verna 1987). Despite technical advances, it is still a relatively expensive and challenging technique, but it is increasingly used in TV-production both in sports and in other live genres. In the following sections, we briefly review the existing literature on how video feeds are selected for broadcast under live conditions and how live broadcasts are produced, focusing on sports programming. Although little is known about the use of instant replay, it fits into a wider picture of visual broadcast media production, and it shares a common technical platform, televisual conventions, and social practices with these media.

There is a small literature on the organization of camera selection for broadcast. The interactional work of multicamera work and vision mixing, as it is done in studio productions and live sports television, has most noticeably been examined in the ethnomethodology and conversation analysis literature. For example, Mondada (2009) presents a detailed analysis of camera selections and the use of a particular image configuration — the "split screen" — in the editing process of a broadcast TV-debate, and shows how this configuration may be related to moments of conflict in the debate. In terms of collaborative production practices, using recordings produced in the control room during live edited studio interviews, Broth (2008, 2009, 2014) has provided detailed analyses of the unfolding interaction between camera operators in the studio and the director in the control room, who continuously chooses between the shots that the operators propose for broadcast. Most pertinently, Broth (2008) shows how opposing participants are oriented to in TV productions and made relevant in the particular narrative presented to viewers in and through the sequencing of broadcast shots.

Unlike live camera selection, both Engström et al. (2010) and Perry, Juhlin, and Engström (2014) explore replay production in live sports television in which they examine activities around the introduction of recent, but *historical* (or post-production), footage into the broadcast video stream. Both papers discuss typical situations where work involves replay operators who, in response to significant game events, produce replay sequences of shots that need to be scrutinized and assembled very quickly. These shots then require additional work to negotiate and coordinate the insertion of the replay footage into the live broadcasts. Engström et al. (2010) present an analysis of the integration of replay footage into broadcast footage, exploring how the production team and camera operators collaboratively conduct search activities and synchronize replay production with game time. While narrative is considered as a resource in creating content for the live broadcast, their analysis focuses on a penalty in the game that the production team had not

noticed. In this context, the replay operator searched for relevant footage that would help explain the sequence of actions that led to the referee making a penalty call and to find an appropriate edit point for creating the replay. To achieve this, the whole production team pulled on a variety of visual and social resources to solve this problem together. As the replay operator scrolled back through the video, a format shift in the recorded footage was revealed that exposed an explanation for the penalty. This visible change in the recorded material enabled the replay operator to understand what had just happened in the game and consequently to search back in time for relevant footage that would be of value to viewers. Perry, Juhlin, and Engström (2014), on the other hand, explore how six EVS operators and a replay subeditor engage in sensemaking activities around meaningful aspects of gameplay and collaboratively "do looking together" in making relevant replay clip selections. The analytic focus is not concerned with interactions between the vision mixer (VM), commentators, or audience reactions as resources for producing time-suitable replay content selections, but lies in examining the alignment of visual content and talk in organizing team members' interaction to show how they make themselves accountable for their selection of video clip proposals for broadcast. The authors are concerned with narrative features of the game in activities or patterns that delimit the sequential relevance of replay items. However, their paper's focus on using narrative as an organizing feature is only discussed in passing; likewise, in the paper, talk was used by replay operators to make their replay selections observable rather than to interactively sequence live and recorded content together.

Similar to our own interests in collaborative editing, Laurier and Brown (2011) describe the moment-by-moment interactive work involved in post-production of a video documentary. At a superficial level, such editing has a common feature with the use of instant replay in that the production teams can return to events in the existing material to provide better explanations or visual perspectives. Yet unlike in live sports TV, such postproduction videos are "shaped in the edit" (Rabiger 1998), allowing a considered set of audio-visual materials to be constructed around a carefully orchestrated account of the narrative (cf. Barnfield 2013) that is not temporally dependent on the order of emerging events. However, there are some similar features in the form of this work compared to live editing. Laurier and Brown point to the ways that distributions of *knowledge* occur in multi-participant documentary video production, in addition to distributions of *work* between an editor and director. In this distinction, they explore the members' knowledge of the footage they have, where it is and what the footage is like, showing how production teams team face an on-going challenge of marrying together their differing sets of knowledge (e.g., of technique, footage available, or aesthetic or commercial demands) in fusing imagery into a meaningful and visually compelling set of video sequences. As with live TV, such activity is shown to be highly dependent on language use, gesture and temporal cues in interruptions and "noticings" of content, and making visible these concerns between members.

Live sport TV itself is of course somewhat different to other genres of live TV and of video production in general. In live sports TV, camera operators "propose" shots (for selection by the VM) through the way they frame and focus on topic matter. One of the reasons for this is the often fast-moving nature of gameplay that is best shown from sequences of multiple angles to produce a coherent visual experience (cf. Jayyusi 1988). Perry et al. (2009), for example, describe the use of "lean" coordination mechanisms to help organize production, such as indexical gestures by camera operators in live sport to "point" to unfolding action to demonstrate their availability for broadcast selection after brief intermittent searches with the camera. This "pointing" is in contrast to the poorly framed or jittery camerawork seen while they "chase" the action, or search for footage in a field of play, in which multiple events occur simultaneously. In the absence of audio feedback to the broadcast studio, providing stable, game-relevant imagery is a visible indication they are ready to be broadcast, and will follow any events as they play out. To enable these production practices, production teams must rely on a mutual orientation to, and understanding of, the interactional and visual logic of what they are covering for coordinating their work (Broth 2008; Macbeth 1999). In a similar way, Camus (2015) described what he calls "la course aux ralentis" (approximately "slow motion competition"), which takes place between multiple replay operators in major football television productions; in order to have a chance of selection by the director, operators announce as early as possible the availability of "their" replayable segment.

The premise of live TV is that it is, as near as feasibly possible, simultaneously recorded, edited, and transmitted. In televised sports, this is achieved by operating multiple cameras and mixing their real-time content with recorded video sequences ("instant replay"), audio, and supporting graphics. The main direction and visual production of a live TV show is conducted in the production control room. This room contains a "gallery" of video monitors displaying all camera sources centered around a main broadcast monitor and a preview monitor. An intercom system enables communication between this room, the camera operators and adjacent production units.

A television production team's task in live sport is to provide relevant and visually interesting broadcast footage that will help viewers understand what is going on in the game *and* entertain them. In the words of one director discussing soccer,

> [Y]ou take the shot where the action is, that's most important, getting fancy is always the gravy, it's like you've got the bread and butter shot is the guy with the ball, that's your responsibility to the audience, to show them. Anything on top of that is making it fancy, it's the gravy which is a close up. (Silk, Slack, and Amis 2000:9)

However, cutting to real-time close-up footage is also extremely demanding when editing visual content, at least partly because of the visual editor's uncertainty about how long a camera operator will be able to, or attempt to, hold a

close up shot (see Perry et al. 2009). This makes the role of the instant replay operator particularly important, as it allows such close-up footage to be obtained retrospectively.

The ability to create instant replay material in the production of contemporary television relies on the use of nonlinear (tapeless) media, which allows immediate access to anywhere in the stored video footage. Video and audio materials are captured to a storage device, which allows recorded footage to be searched, segmented, resequenced, and played back (see Broth, Laurier, and Mondada 2014). In live sport that involves the use of multi-camera recordings, these systems allow program editors to cut into the live broadcast to show recorded footage from cameras that were not initially selected for broadcast, allowing the use of multiple angles on actions taking place during the game and at different playback speeds.

For very practical reasons, due to the speed of play, rules of the game, audience expectations, and the size of the rink, it appears that TV productions of live ice hockey are less of a planned exercise than other sports, such as athletics or golf. In the words of one producer,

> Hockey production is hockey production. It's one of the easiest sports to do because there's nothing to prepare for and the individual producer can only inject small differences in taste by the pace and flow of cutting cameras.... Every producer does it the same. (Shannon 1988, cited in MacNeill 1996:111)

Balanced against this self-depreciatory portrayal, hockey production teams are highly skilled and encultured within a set of professional and organizational practices. In what follows, we will closely examine these practices as they unfold in a series of activities.

EMPIRICAL WORK: METHOD AND SETTING

Data collection took place in the outside broadcast (OB) studio located in a custom-fitted bus outside an ice hockey arena in Sweden. Inside the arena, three manned and two fixed cameras were available. Two cameras positioned side by side covered the main action from an elevated perspective high up on the terrace; one camera was used for an overview (C1), another for in-detail, providing close-up shots from the same perspective (C2) (see Figures 1 and 2).

A third (C3) was mobile and positioned on the floor by the center line, just behind the wall of the rink and at eye-level with the players and referee. C3 was used for close-up shots at the rinkside and goal areas, its ground-level perspective between the player benches conveying closeness to the action, but its close proximity to the ice also meant a greater risk of the shot being obscured (as compared to the elevated C2). Its range covered the entire ice except for the near corners. The unmanned cameras (C4 and C5) were fixed on opposite sides of the rink behind each goal (see Figure 2 for typical framings from each camera). The OB studio just outside the
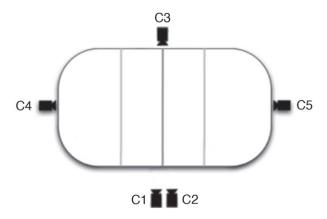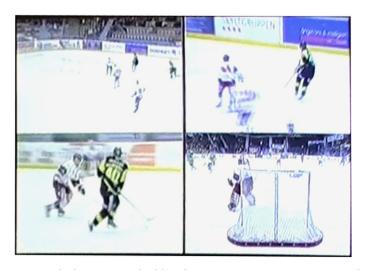
FIGURE 1. Rink and Camera Positions

FIGURE 2. Typical Shots Provided by the Main Cameras Covering the Same Action; Overview by C1 (Top Left), Close-Up from the Same Angle by C2 (Top Right), Low Close-Up (C3) and Fixed View behind Goal (C4). Captured from the EVS Workstation

arena had workstations for the VM, producer, production assistant, and graphics operator, all facing a video gallery (Figure 3). This video gallery displayed all of the camera feeds on wall-mounted monitors, the replay operator's work and graphics overlays. Directly behind the VM was the replay (EVS) operator's workstation. Because of their physical proximity, the VM and the replay operator can talk to each other directly. Everyone in the OB studio could hear the commentators over the loudspeakers inside the studio. The VM could speak directly to the commentators via an intercom headset, while the replay operator could speak back to the

FIGURE 3. Vision Mixer and Production Gallery

FIGURE 4. Replay Work Station Showing Monitors and the XT[2] Control Unit

commentators over a microphone by pressing a button to activate the intercom system.

We recorded two ice hockey matches; each match lasted for approximately 2.5 hours, and in total, the study generated over 15 hours of video recordings (including pre- and post-match events) from three cameras deployed by the researchers. One camera was directed at the monitors in the control room, another framed the replay operator's hands from the side, and a third overlooked the operator's face and the screens in front of him. Recordings were repeatedly viewed in team analysis sessions, and core events transcribed and categorized. The material has been translated from the original Swedish into English for an international readership

FIGURE 5. Replay Work Station Showing the XT[2] Control Unit

and this may obscure some of the subtleties of the original text. However, where appropriate, we have attempted to clarify these translational issues when we present them. Sequences have been numbered in the text below from our event transcripts; numbering is restarted for separate events (some lines have been removed for clarity).

The replay operator in the data examined below used an EVS Multicam LSM (Live Slow Motion) system (see Figure 4), which allowed recording, and manipulation of nonlinear media recordings. This EVS unit was coupled with an XT[2] production server (Figure 5), allowing rapid interaction with the visual content. This combination drove a split screen monitor showing four real-time camera feeds (the larger, top screen) and two additional monitors for viewing the replay operation directly below it (foreground of Figure 4). This setup recorded multiple real-time camera feeds to the server continuously throughout the game, and enabled the operator to go back in time to any of the camera feeds, search within the video, and edit short sequences to be replayed. The controls to the XT[2] unit allowed camera selection from all of the 5 real-time camera feeds. It also had a video jog wheel for searching within the video stored on the server and a playback control lever (accelerating and reversing playback at up to 4× normal speed). The small screen that can be seen in Figure 5 provided access to a video bank for storing clips for later access; these clips could either be played individually or grouped together as playlists.

The EVS operator's work involved the continuous identification of potentially interesting situations in the game. When such a situation took place, he typically went through the cameras to examine which held a suitable framing of the situation by rewinding the video that had just been stored on the server. He would then select one (or more) video streams that showed this
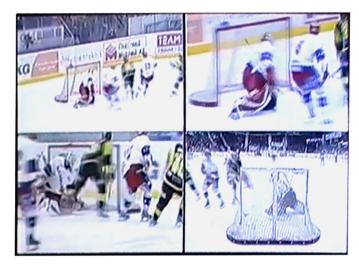
FIGURE 6. EVS Operator's Screen Displaying a Split of Four Selected Cameras

situation (see Figure 6, showing the view from four cameras, immediately following a goal).

On locating this, he would set an "in-point" to this feed and then wait for directions from the VM. If the VM, who relies on the EVS operator to have done just this, calls "EVS … ," the operator prepares to roll the sequence upon the command " … now." If no such call is made, the sequence is stored in the video bank for later access. In practice, the EVS operator's visual feed to the gallery (visual, because it contains no audio content) is very similar to that of a camera feed. The VM cannot directly edit this EVS stream; he can simply choose to broadcast it or switch to another visual feed. The EVS operator therefore acts like a subeditor for the replay footage, and he is able to make cuts, in real-time, between the pre-recorded camera footage as and when he chooses to do so, independently of the VM.

Next, we discuss critical components in the analysis that allow us to rise apart the topic of synchronizing real-time and recorded content: the narrative production practices and use of the recorded video material to produce meaningful footage of the game that can be cut into the ongoing live broadcast. This particular topic was selected for analysis because it illustrates important production practices through which media are selected from the different real-time and recorded feeds. It also shows how these feeds are cut together to produce meaningful footage that the production team judge to be relevant and aesthetically suitable for broadcast, and which allows historic visuals to be inserted into ongoing game play in such a way that it does not impinge on audience enjoyment of live events. These are of course subjective judgements, and in the absence of audience feedback, the production

team have to rely on their own interpretation of the viewers' perspective of the game.

NARRATIVE PRODUCTION PRACTICES

So far, we have referred to the production of narrative in the search for footage without closely exploring what we mean by narrative, how narrative is understood and exploited by the production team, or how it is produced under live broadcast conditions. The production and deployment of instant replay footage by the EVS operator plays a major role in this process, by presenting a perspective on the game through clips that focus the viewers' attention on particular developments in the game — necessarily to the exclusion of other parts. The real-time and recorded media used in the broadcast are also used to produce a sense of liveness and presence to the remote audience. This mediated content directs how the audience understands the unfolding action and experiences the excitement of game play. For instance, the overview camera (C1) is broadcast most of the time; this camera is usually centered on the action around the puck, restricting the viewer's access to action that is immediately peripheral to this. Of course, this action is often neither peripheral to game play or to the enjoyment of the game as a spectacle (cf. Gruneau 1989), but it is this constructed experience through the media that provides a look and feel for the edited TV broadcast that distinguishes it from the "real."

Replays perform an important role in enhancing this mediated live experience, by selecting items of interest and focus that can pull out aspects of the game that received less attention when broadcast live. That all cameras are operated to provide images of the central action at all times means that a substantial part of the "live" experience at the location is left out, or can only be broadcast in breaks between game play. The EVS operator's role in this is to fill pieces in to enhance the narrative structure of the game underway to help produce an edited experience of the match. Yet this is not a simple operation: the structure of the production is oriented towards the real-time aspect of the broadcast, and this sets practical constraints on what can and cannot be achieved. The way cameras are operated to frame the most significant action (MacNeill 1996) limits their usefulness in creating replayable footage; other than for a few well-scripted instances (e.g., goals, fouls, breaks, and end of play), none of the cameras are explicitly instructed to follow up on these actions by players *after* events that occur. The work of the EVS operator is therefore less like that of the VM who can expect relatively clean footage of action around the puck, and more like bricolage, as he assembles nonoptimal visual material that has to be cut together quickly.

It is this selection of recorded visual material for insertion into the real-time broadcast that we now turn to: how the production team themselves make sense of the developing game and re-present this to viewers, and how the process of managing transitions that fit replay sequences into the real-time broadcast is achieved.

Sensemaking and the Production of Narrative

Making sense of the ongoing game and the production of the game narrative is an ongoing and dynamic activity, managed and negotiated over time by distributed team members: VM, producer, EVS operator, and commentators. However, their resources for communication are relatively limited and asymmetrical in providing support for this. The bulk of this coordination is achieved within the real-time broadcast video channel and supplemented in the other direction by the broadcast audio, since the commentators and rinkside sounds are always live.

Our analysis shows some of the ways in which distributed collaboration and sense-making is involved in the contingent production of narrative: team members jointly negotiate and form a consensus of the meaning of events as the game unfolds. Over time, this consensus becomes a focal point for visual image and graphics selection for broadcast, as well as for the verbal discussion presented by the commentators. That the game narrative is collaboratively produced over time by physically separated parties within the broadcast team, and the potential problems this entails, is made clear on several occasions. Extract (1), which immediately follows a goal, illustrates this.

As we enter the action, a goal has just been scored in a game by the Västerås team against Huddinge's team. Immediately after this happens, the commentators conclude that there has been a goal, and this is where our extract begins. Simultaneously to this commentator announcement, in the OB studio, the Producer is celebrating this goal, while the VM continues to select the next camera for the live broadcast. Close by, the EVS operator is already preparing a sequence of replays for broadcast that can be brought into the live broadcast to help revisit and explain what had just happened. What we see as this activity unfolds is that the production team themselves use this material as a resource to make sense of the action so that they can introduce even more informative content into the rolling broadcast. In what follows, we present how the event plays out, drawing together the transcript of the verbal record and our observations of team members' actions, camera, and EVS footage from our video recordings inside the OB studio. To help present these sequentially, we have separated the extended excerpt into shorter segments, followed by a line-by-line analysis of how the participants interact and make sense of critical events. In the excerpt, the VM's reference to numbers refers to the camera numbers (see Figure 1) that he intends to select for broadcast. In the transcripts below, C(n) refers to camera numbers, EVS C(n) to camera used in EVS feed, and * to camera switch moments. Further, curly brackets,{}, are placed around talk that is simultaneously produced by commentators and in the control room, whereas square brackets, [], are placed around overlapping talk among co-present participants; pauses are timed in tenths of seconds; "=" denotes no pause and no overlap, ":" lengthening of sounds, "()" uncertain hearing, and "(())" transcriber's comment.

Extract 1a. Participants are COM (commentator), EXP (expert commentator), PRO (producer), VM (vision mixer), and EVS (EVS operator).

```
01 COM: a {de e mål}
        yea it's a goal
02 PRO:   {Ö:::::::}:::::
            ((celebrating exclamation))
        (0.4)
03 VM:  tvåan (där)
        two   (there)
04 COM: (.) ä:: å de {e de e Sebastian Meyer som (sticker) fram}
            eh  an it's it's Sebastian Meyer who (goes) forward
05 VM:              {[trean   ]     =binga trean   (.) trean   }
                     three color correct three        three
06 PRO:             {[(de e) i] bild=                         }
                     (it's) on the screen
07 VM:  (1.2) där (0.5) ettan
             there       one
08 COM: (0.2) å de e ett tidit mål i matchen efter t{vå:} noll nie
             an it's an early goal in this match after two oh nine
09 VM:                                            {där}
                                                  there
10 COM: (.) {ett noll Västerås.     }
             one nil Västerås
11 VM:      {tvåan (.) (ta re lugnt)}
             two        take it easy
12 VM:  (0.5) tvåan (1.6) där {(push:)    }
             two        there (push)
13 PRO:                  {va re en mål}vaktstavla, eller
                          was it a slip by the goalie or
14 VM:  (.) fyran (1.1) där (0.8) tvåan (.) final (0.9) nu
           four       there      two       final     now
15 VM:  (1.9) e ve ess:
             EVS
16 COM: (  {     )} ä en bidragande orsak eh:
        (        ) is a contributing factor eh
17 EVS:    {a:   }
            yes
18 COM: (.) {((till))} att de går bra för*Västerås just nu
            that things are going well for V. right now
19 VM:     { nu    }
            now
                                  *VM switches to EVS feed, C1
20 COM: (0.5) tavla eller inte,
            slip or not
```

Here, we see an example of strain between commentators on the one hand, and the broadcast production crew on the other, and how they negotiate to arrive at a common understanding of gameplay leading up to a goal. This exchange is instigated by the Producer with the question "was it a slip by the goalie, or" spoken out aloud in the control room in response to what had happened on the rink (line 13). The VM then cuts to the EVS (lines 15 and 19), allowing the commentators as well as the audience to make judgment of this event based on replayed footage. The two-part syntax of the VM's command to the EVS operator, "EVS … now," can be seen to be tailored to meet the demands of the task of transitioning from the live feed to the EVS feed, which requires a high degree of synchronization between the VM and

EVS operator. The first part ("EVS" line 15) serves to establish the readiness of the EVS operator (here confirmed by "yes," line 17), which is a pre-condition for the close coordination of the second verbal part of the command ("now," line 19) and the actual execution of the start of the sequence of replays. This exchange continues in Extract 1b.

Extract 1b

```
20 COM: (0.5) tavla eller inte,
              slip or not
        (2.6)
21 VM:  jävla passning asså
        what a bloody pass
22 COM: (0.4) ja: va säje*ru
              yes what do you say
                    EVS C1*EVS C2
23 COM: (0.5) ska han ta [ren där.  ]
              should he take that one
24 EXP:                  [näj vi ser] att han--
                          no we see that he
25 EXP: han ä väl inte riktit me på var puck{en ska gå  } eh
        he is not really prepared for where the puck is going
26 EVS:                                    {(den) sista}
                                            (the) last one
27 EXP: {(.)} Sjögren i Huddingemålet å {(.) stöter fram me klub}ban
             Sjögren in Huddinge's goal     pushes his stick
28 PRO: {bra}                            {dåli passning av backen}=
        good                              bad pass by defender
29 VM:  =tvåan=
         two
30 EXP: {(.) å} mi*ssar pucken {(0.7)} (.) {eh::
             an misses the puck          eh
                    *VM switches to live feed
31 VM:  ={där  }                {ettan}     {släpp match, ettan}
         there                   one         release game, one
32 EXP: (0.4) men de e bra gjort av Sebastian Meyer å {de e vi}ktigt för
              but it's well done by Sebastian Meyer an it's important for
33 VM:                                                {tack    }
                                                       thanks
34 EXP: (.) Väs{terås (0.8)    } (0.3) att direkt ta kommandot här=
            Västerås                    to take control immediately here
35 VM:         {trean (0.5) där}
                three       there
36 EXP: =(efter den här matchen)
         (after this game)
```

To help understand the ways that image production connects to this verbal exchange, material from the EVS screen corresponding to this extract can be seen in Figure 7 showing the visual imagery and the point at which cuts are made. Following review of the replay that is broadcast, the commentator (who can hear this over the intercom) appears to respond to the producer's query about the keeper's slip in his next voiceover comment (line 20). The topic of a possible slip is thus brought up verbally in the commentary as the first replay segment begins to be shown (line 22), which leads us to link this to the producer's question as he refers to the same topic, using the same term ("tavla," informal Swedish for "mistake"). Viewing this first replay segment, the producer comments aloud on a bad pass by a defender
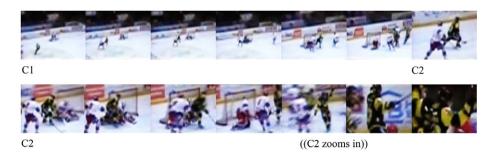
FIGURE 7. Screenshots Showing EVS Sequence from Lines 19 to 30 in Extracts 1a-1b (Timeline Flowing from Left to Right, Top to Bottom)

prior to the goal, implying that this is the reason for such poor game play (line 21). Having finished viewing the first replay of the goal (segment finishes in line 22), the commentators address the issue of the "save-ability" of this goal in their commentary (line 22 and onwards). After the EVS operator's announcement that the replay footage is due to finish (line 26), the VM announces and carries out a cut to camera 2 for broadcast (lines 29–31), as the commentators continue to complement the attacking player's skillfulness (line 32). Although it might be argued that the skilled reading of the game by the producer, VM and commentators could have led them to focus on exactly the same question (and use the same term) on the reason for this goal, a more plausible explanation is that they use the instant replay and audio link as resources for them to mutually orient towards the same topics. What we see then is that the EVS operator, VM, and commentators reflexively provide resources (questions, relevant EVS footage, and interpretive commentary) to the others on this topic as it develops, so that they can themselves in turn provide additional resources to interrogate this explanation.

Making sense of what happened in this instance is very much a collaborative exercise through micro-adjustments towards the needs of the other participants in developing a coherent narrative of the game to help provide an explanation of what had happened and why it had occurred. In this case, the production team in the control room was free to voice their opinions and share comments among themselves in order to make sense of the live action as the game unfolded, insofar as these comments did not interfere with their operational communication. This communication and the multiple camera angles available in the video gallery and EVS logs provided additional resources that they used to make sense of action in the rink. The commentators, in turn, are listening to the key members of the production team (producer and VM) over the intercom. However, as their voices are constantly being broadcast, their commentary is primarily directed towards their TV audience. This means that any comments on the resolution of ambiguity within game play by commentators during the live broadcast need to be formulated so that they make sense to the audience, although they may also be implicitly directed at the production team in requesting additional footage.

This analysis shows how distributed sense-making is undertaken by the whole production team, utilizing their own skilled readings of the game and technical resources to shape the broadcast into a cohering narrative of explainable gameplay — all of which occurs within a tight set of technical and broadcast constraints that limit both conversational synchronicity and opportunities for traditional turn-taking symmetries. It clearly illustrates how team members work to jointly negotiate and form a consensus over what they understand has happened. This sensemaking activity is woven into and through the broadcast audio and footage to collaboratively arrive at a coherent audio-visual account that explains to the viewers what has been happening leading up to the current state of play, who was at fault, and what is likely to happen next.

Playing with, and for, Time in Creating Explanations

Narrative concerns require content to be produced that support interpretation, but that is also made available at an appropriate moment. The VM faces a problem in managing these transitions between real-time and instant replay footage during a game in support of this narrative. Such segueing of replay sequences into the live broadcast is complicated by the contingent character of events unfolding in the rink: real-time broadcast footage of game-relevant developments should not be obscured by potentially less-relevant and timely replay footage.

Similarly, where the game is "out of play," for whatever reason, there is arguably an expectation that pre-recorded material should be brought into the broadcast to either help explain the current situation, or to fill in a relatively boring break in the action. The EVS operator does have a means of artificially manipulating the length of his clips to mesh with the real-time footage by dynamically controlling the playback speed to accentuate and enhance actions by stretching them out temporally. Nevertheless, this cannot be used continuously, nor does it solve problems of cutting the broadcast replay neatly back into the real-time footage for when gameplay begins again. This temporal synchronization between feed transitions is therefore most commonly handled through close coordination between the VM and EVS to ensure that broadcasts of the recorded clips dovetail neatly into the real-time images.

This coordination is illustrated in the following situation (Extract 2) after a goal had just occurred, with the replays again being used to show the action leading up to it. Extract 2 begins around nine and a half minutes after the end of extract 1. Within the extract, the instant replay is deployed by the VM, allowing the EVS operator to directly control the replay sequence within the live broadcast. The EVS then uses a lightly pre-edited set of five replays (taken from four cameras) that run in sequence to aid the commentators in their ongoing explanation of how the goal had happened, with the final clip of the replay intended to provide a display of emotion by the scoring player. However, the decision to air the last camera shot turns out to be a mistake, when the promised final clip does not stay on the victorious player and instead closes the sequence on a confusing blurred shot. The example shows how

the task of transitioning between real-time video and replay is managed collaboratively between VM and EVS operator, how they establish the joint focus needed for the playback of the replay, how the operator makes the VM aware that this particular replay will be extended in time and makes use of multiple cameras, and how the operator communicates his intended ending of the sequence to the VM who can then resume full control of the broadcast. The breakdown in the intended outcome is especially interesting as it reveals how an unexpected error, arising out of the EVS operator's misinterpretation of a piece of camerawork, is handled and concluded. As before, the extract is broken down into shorter segments for the purposes of explanation and analysis (Extract 2a):

Extract 2a

```
01 VM:     tvåan
           two
02 COM:    (0.7) men två {noll} e re
                 but two  nil it is
03 VM:                    {här }
                           here
04 COM:    (1.0) ah de stämmer
                 yeah that's right
05         ((3.5 s untranscribed))
06 COM:    pucken lite turligt når fram dit s-
           the puck with a bit of luck arrives there s-
07 EXP:    (0.3) ah den e luri, vi kommer se här, den e: eh,
             yes that is a tricky one, we'll see here, it's eh,
```

As only real-time footage is still available in the broadcast at this point, the commentators describe from memory the way the goal came to be scored. In their description, they categorize what happened as "with a bit a luck" (line 6) and "that is a tricky one" (line 7). Immediately linked to the second categorization there is a comment that anticipates the upcoming replay ("we'll see here," second part of line 7). The extract continues:

Extract 2b

```
08 EVS: (0.4) jass{[å::]            }
              oh yeah
09 VM:             {[eve] äss:::     }
                    EVS
10 EXP:            {de e meningen av} eh:
                    it's on purpose by eh
11 VM:  (.) nu
            now
12 EXP: (0.5) A*lbrant att den skulle bli {svår (så han drar) sig ur}=
             Albrant that it would be difficult (so he withdraws)
              *VM switches to EVS feed, C1
13 VM:                                    {ta : eh : : : : :        }
                                           take eh
```

The comment in line 7 (Extract 2a) from the commentators—which is audible by not only the audience of viewers but also by the people in the control room—is responded to by both the VM and the EVS operator. Whereas the EVS operator, a bit

mockingly, questions that anticipation "oh yeah" (line 8), the VM cuts in and initiates the first part of the command to the EVS operator to get ready to start providing the replay (line 9). Following this, the EVS operator halts his current editing task. Half a second later, the go ahead is given (line 11) and the EVS operator's new replay is put on the air (line 12). As soon as the EVS feed is aired, the commentators can comment on the sequence of replayed versions of the events leading up to the goal, and do so in Extract 2c as follows.

Extract 2c

```
12 EXP:   (0.5) A*lbrant att den skulle bli {svår (så han drar) sig ur}=
                Albrant that it would be difficult (so he withdraws)
                   *VM switches to EVS feed, C1
13 VM:                                      {ta : eh : : : : : :        }
                                              take eh
14 EXP:   =så lägger han en liten {(                              )} studsa
           then he makes a little  (                            ) bumps
15 EVS:                        {ja tömmer hela skiten Janne}
                                I'll empty the whole damn thing Janne
16 EXP:   (1.2) (precis {innan x)         de e  (.)}
                (just before x,)           it is
17 EVS:              {ja tömmer he*la skiten Janne}
                      I'll empty the whole damn thing Janne
                          EVS C1*EVS C2
18 EXP:   jättesvårt för Huddingeförsvararen där å ta ren
          very difficult for the H. defender there to take it
19 EXP:   (0.7) han går ner å (ren) kryper under
                he goes down and it crawl under
20 EVS:   (1.1) * (1.6) så ä re gjort=
                        so that it's done
          EVS C2*EVS C3
21 EXP:   =under händerna på nummer tjutvå ska re va (istället)
          under the hands of number twenty-two it should be (instead)
22 EVS:   (1.5) * (1.7) kommer glädje också
                        joy is coming too
          EVS C3*EVS C4
```

As we show in the data, it turns out that the EVS operator has prepared footage of four different versions of the goal, which will be presented to the viewers in the order of cameras 1, 2, 3 and then the unmanned camera behind the goal. This is what he refers to when he announces, twice, "I'll empty the whole damn thing" (lines 15 and 17). As he glances over his monitor (see Figure 4) and hears that the VM is involved in another side conversation on the first occasion that he says this (line 15), the EVS operator repeats his utterance a second time (line 17). During his second announcement, he also makes his first visual transition, from the replay of camera 1 to the replay of camera 2. As the replay of camera 2 is being broadcast, the EVS operator scrolls backward to the beginning of the segment to be replayed from camera 3, and as that segment is broadcast, he scrolls back to the beginning of the goal footage from the unmanned camera behind the goal. As the segment from this camera is broadcast, he scrolls just a little bit forward, to find the beginning of the shot of camera 2 in his right hand monitor, and stops at an image where a player raises his arms in the air, expressing the joy of having scored the goal.

This last image turns out to be the starting point of a fifth and final component of the sequence of replay items. However, as there are only four cameras available (C5 being at the other end of the rink), a fifth component might not therefore be expected by the VM (having heard that the EVS operator proposes to "empty the whole damn thing"), with whom the EVS operator needs to coordinate the return to real-time footage at the end of the replay. So, countering the risk of the VM switching back into real-time broadcast at the end of the fourth component (i.e., showing footage of the camera just behind the goal), the EVS operator says that there will be a replay of the expression of joy as well (line 22) as the fourth replay segment is broadcast. The first frames of the full set of replay feeds from the various cameras broadcast can be found in Figure 8.

Close to the end of the fourth image segment, the EVS operator switches to the footage of camera 2 (the final image in the sequence), the beginning of which he had prepared just a few moments earlier. After only a few moments of broadcasting the fifth segment, however, something happens that is treated as an accident by the EVS operator. Contrary to what he had apparently anticipated, camera 2 does not stay on the shot of the joyful player for very long. Instead, after only a short moment, it makes a swift pan away from the rink to show footage of the audience (line 25). As the conversation in the studio progresses, the camera pan is clearly treated as something that was not suitable for broadcast.

Extract 2d (Commentator talk not transcribed)

```
22 EVS:   (1.5) * (1.7) kommer glädje också
                         joy is coming too
          EVS C3*EVS C4
23 PRO: =ska (vi) stanna rå:
          do (we) hold/stop then
24 VM:   (0.7) håll match*=
                hold game
                    EVS C4*EVS C2
25 PRO: =a:: (0.7)+(0.3) låt den kör[a:]
          yes                let it run
                    +C2 pans quickly right, shot of audience
26 EVS:                      [a-] häj de gjorde de inte
                                 yea- no it didn't
27 EVS   (0.3) tack (0.4)*
                thanks
                         *VM switches to live feed
28 EVS: (1.9) my bad (1.1) sorry gurkan
          my bad          sorry gurkan ((to camera operator??))
29       ((7.2 s untranscribed))
30 EVS: åh: (va) dålit
          oh  how bad
31 VM:   (0.8) (en) jävla tur att inte du (är me mer)
                we're damn lucky you don't (participate more)
32 EVS: (0.6) m:
                m
33 VM:   (0.3) he he he h: ((VM laughter))
34 EVS: (0.5) va fan man tar en bild
                what the fuck you take a shot
35 EVS: (.) (head to toe) då tror man att den ska va kva:r där
            (head to toe) they you'd think that it would stay there
```

FIGURE 8. Left to Right: Cameras 1, 2, 3, Unmanned Camera 4 and "Joy" from Camera 2

After viewing the pan, the EVS operator shows that his earlier anticipation did not prove to be true "yea — no it didn't" (line 26) and announces that he is done by "thanks" (line 27), after which the VM switches back to live (line 27). Shortly thereafter, the EVS operator states that it was indeed his fault, and also apologizes, most likely to the camera operator (line 28). After some time, the EVS operator assesses, as if talking to himself, this broadcast content as poor work (30), and provides, in response to a tease by the VM (lines 31–33), a hearable and accountable explanation that when you see a player taking a shot in that way, then you would also expect the camera to "stay there" (lines 34–35). In doing so, he appears to be offloading some of the blame for this failure onto the camera operator for not doing his job as expected: in his comment, he asserts (whether true or not) that the camera operator had *not* oriented to the production team's need for narratively relevant imagery.

In this second analysis, we show how the entire team work around solving the problem of creating a visual account of what had just happened in making sense of how the goal had been scored under highly time-restricted circumstances. To do this, the VM and EVS operator have to support the commentators (first) and viewers (later) in making interpretations about the lead-up to this sequence of events, and to tie this in at an appropriate time for the explanation to be relevant to the viewers and so that it slots neatly into the ongoing real-time action on the ice so that it does not occlude further narratively relevant developments in the game. The incident illustrates how the task of transitioning between real-time video and replay was collaboratively managed, allowing control of the broadcast footage to be passed back and forth between operators, and for the replay action to be inserted for an extended period of time and showing multiple camera perspectives. In this respect, the breakdown of expectations evident in the example here illustrates that narrative is doing double-duty in production work: it is something that helps the viewers at home (and commentators) make sense of the progress of the game, but it is also a resource for organizing teamwork in image production. This secondary role of narrative — that footage be relevant to explaining sequential actions — allows the tightly coordinated visual sensemaking activities that are critical to the function of making time-dependent instant replays.

DISCUSSION

The insertion of instant replays in live sports TV-broadcasts manifests an expectation, by the audience (or at least an audience expectation that was anticipated by the TV-crew) to see what happened *again and in different ways to that originally broadcast*. Instant replay therefore identifies moments in the emerging real-time and audio-visual narrative where this additional resource for making out what just happened in the game not only becomes relevant and expected but also possible, considering the current state of the ongoing game. While temporality has been extensively discussed in previous work (Engström et al. 2010; Perry, Juhlin, and Engström 2014), narrative has been attended to only in passing, and this article explicitly examines the interrelationships between narrative and temporal organization in image work.

We have seen how the production and broadcasting of replay sequences involves many different temporalities: for instance, that of the game itself, of the online commentary about the broadcast visual feed, of the talk in the OB studio, and of the playback and sequencing of the recorded video material itself. These different temporalities are both skillfully managed and oriented to by the professional TV-crew as it works to produce a coherent and intelligible narrative (cf. Broth 2008; Camus 2015; Jayyusi 1988) in the form of an emerging sequence of real-time or just recorded shots for their audience of viewers. Members of the crew manage their accountable participation (Goodwin and Goodwin 2004) in the collaborative production work using their resources at hand, which are not only used for crafting the broadcast product into a coherent narrative, but also, in a "double-duty" way, for communicating within the crew. For instance, we have shown how commentator talk and broadcast replayed visual footage may be tightly fitted to, and reflexively configuring, each other. Thus, on the one hand, the commentators may implicitly request particular visual content through their commentary; on the other hand, showing particular visual material clearly makes specific ways of commenting the shown action relevant and expected. These serve to both enrich the broadcast game's intense reality experience of "liveness," as well as providing materials for making sense of gameplay (cf. Camus 2015).

Whatever their level of expertise in working with multiple image streams, the members of the production team are limited in what they are able to attend to; as we have seen, even skilled professionals struggle to attend to multiple visual feeds, let alone fuse them together into a visually and narratively smooth sequence of images. Creating a broadcast that allows temporal shifting between historical and current feeds is highly complex in that recorded visual content needs to be both relevant and produced to fit into moments that do not disrupt ongoing activities. The collaborative production of instant replay material depends crucially on the VM and the EVS operator's "professional vision" (Goodwin 1994). This involves, for example, what aspects in the rich visual environment are relevant for their current work tasks and where — here especially in what screen(s) — this can be seen or expected to appear. It also depends on their "professional hearing" within the complex phenomenal

environment in the OB studio, which allows them to know which auditive aspects to attend to and listen for, as well as what these aspects project as relevant next actions in the game and in the team's production work. The way they see and hear what goes on around them is contingent on their respective tasks within the activity. Whereas the VM is constantly on the lookout for the next shot to broadcast (cf. Broth 2008), one of which may be to insert a sequence of instant replay, the EVS operator looks and listens for the next salient event from which to prepare a replayable shot or sequence of shots. This highly socially distributed activity concerning the alignment of temporally fragmented content is very different from other settings where visual materials are likewise crucial for the activity, such as post-production film editing (Laurier and Brown 2011), CCTV surveillance (Heath, Luff, and Svensson 2002) or laparoscopic surgery (Koschmann et al. 2007; Mondada 2003). The empirical data presented here on the selection of visual content shows that replay material is selected in a piecemeal fashion that is contingent on its ready availability and the social resources that can be brought to bear on its search, interpretation, selection, and broadcast.

Our analysis shows that there are three critical, and related features of visual production in segueing replays with real-time content.

First, rich conversational interaction is critical to the production of complex and temporally shifted visual sequences. As we have shown, the OB studio is a very visual environment but the audio channel is also available to core participants. Verbal and other audible cues from the director, the commentators, or the stadium audience are an important resource to the EVS operator since he is often working with the recorded video material and cannot be fully focused on the real-time feed. This is not the case for mixing purely real-time visual feeds in sports production (cf. Perry et al. 2009).

Second, a reflexive orientation by the team towards emerging narrative features in the broadcast images is an important feature of the production process because of its role in generating relevant and meaningful TV for viewers. The production team look to what has been shown and said in the real-time audio-visual broadcast in guiding their next actions and collaborative effort. They attempt to both complement and interrogate this developing footage by re-presenting visual material and referring to existing and anticipated/required visual content. This orienting role of narrative is therefore intimately related to temporality, through reference to historic, current, and projected game play.

Third, the team have to "make time" for the others to do their work. The temporal alignment of real-time and recorded content is a real and ongoing practical problem for the production team: events in the world unfold in ways that are not wholly pre-dictable, and both generating and finding visual sequences that can be brought into use at a relevant moment, and returning from recorded visual content to the ongoing, real-time game play without losing narratively relevant material from either tempo-ral state, are not easy. Members of the production team perform work to provide

opportunities for each other to do their work, and in this, the output of the professional production team is highly collaborative.

CONCLUSION

Viewers of live television have become used to seeing the salient features of sports events multiple times, in close temporal proximity to their occurrence, from various directions, and at various playback speeds. What we refer to as *live* sport spectating on television is highly temporally fragmented. Broadcast of this content is possible because of the technologies that the production team have at their disposal in storing, "scrubbing" backwards and forwards through video and interplaying real-time and recorded video feeds. However, we have also shown how the production of meaningful and compelling visual content is not simply dependent on the technical provision of sophisticated editing equipment for individuals to mix together visual data streams. Crucially, it involves social interaction across a skilled team that together makes the technology work to project forwards and backwards across temporally displaced content in the selection and mixing together of visual media.

Examining the situated video practices of the production team us to explore what this work is composed of, and how the participants interact and make sense of the resources that they have to hand. When we unpack how this merger of different temporalities is designed and put together, we see a very complex set of activities in which the production team's orientation to narrative plays a major part. The production of instant replay and its role in the production process is twofold. First, it supports the work of production, through its role in supporting the production team's understanding of, and problem solving around the filmed event, and second by its role in providing narrative accountability for the practical purposes of the viewer in making sense of the ongoing game.

As with Mondada's (2003:67) work on video in surgery, this video collaboration is a finely coordinated "collaborative matter," in which the participants "do not just *follow* but *anticipate*" the actions of others. In large part, this occurs through their shared expectations of what visual content is likely to be required to ensure a smooth and intelligible narrative structure, and in their collaborative work to both make sense of what is happening and has happened, and to make time for the other members to do their work. It is this work, and these practices, that allow temporally discontinuous visual content to be segued together under the real-time pressures of a live broadcast to provide an intelligible, relevant, and visually engaging record of events.

REFERENCES

Altheide, David L. and Robert P. 1978. "Sports Versus the Mass Media." *Urban Life* 7(2):189–204.
Auslander, Philip. 1999. *Liveness: Performance in a Mediatized Culture*. London: Routledge.
Barnfield, Andrew. 2013. "Soccer, Broadcasting, and Narrative: On Televising a Live Soccer Match." *Communication & Sport* 1(4):326–341.

Broth, Mathias. 2008. "The Studio Interaction as a Contextual Resource for TV-Production." *Pragmatics* 40:904–926.

Broth, Mathias. 2009. "Seeing through Screens, Hearing through Speakers: Managing Distant Studio Space in TV-Control Room Interaction." *Pragmatics* 41:1998–2016.

Broth, Mathias. 2014. "Pans, Tilts and Zooms. Conventional Camera Gestures in TV-Production." Pp. 63–96 in *Studies of Video Practices: Video at Work*, edited by M. Broth, E. Laurier, and L. Mondada. New York: Routledge.

Broth, Mathias, Eric Laurier, and Lorenza Mondada. 2014. *Studies of Video Practices: Video at Work*. New York: Routledge.

Camus, Laurent. 2015. "Réaliser en Direct: Une Vidéo-Ethnographie de la Production Interactionnelle du Match de Football Télévisé Depuis la Régie." Ph.D. dissertation, Sociologie, Télécom ParisTech, France. Retrieved from PASTEL Archive: NNT: 2015ENST0036, HAL Id: tel-01336803.

Connor, Steven. 1996. *Postmodernist Culture: An Introduction to Theories of the Contemporary*. 2nd ed. Oxford: Wiley-Blackwell.

Dunlap, Orrin E., Jr. 1948. *Understanding Television: What It Is and How It Works*. New York: Greenberg.

Engström, Arvid, Oskar Juhlin, Mark Perry, and Mathias Broth. 2010. "Temporal Hybridity: Mixing Live Video Footage with Instant Replay in Real Time." Pp. 1495–1504 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'10)*. New York: ACM Press.

Garfinkel, Harold. 1986. "Introduction." Pp. vii–viii in *Ethnomethodological Studies of Work*, edited by H. Garfinkel. London: Routledge and Kegan Paul.

Garfinkel, Harold. 2002. *Ethnomethodology's Program*. New York: Rowman and Littlefield.

Goldlust, John. 1987. *Playing for Keeps: Sport, the Media and Society*. Melbourne, Australia: Longman.

Goodwin, Charles. 1994. "Professional Vision." *American Anthropologist* 96(3):606–633.

Goodwin, Charles and Marjorie Harness Goodwin. 2004. "Participation." Pp. 222–244 in *A Companion to Linguistic Anthropology*, edited by A. Duranti. Oxford: Blackwell.

Gruneau, Richard. 1989. "Making Spectacle: A Case Study in Television Sports Production." Pp. 134–154 in *Media, Sports and Society*, edited by L. Wenner. Thousand Oaks, CA: Sage.

Hastorf, Albert H. and Hadley Cantril. 1954. "They Saw a Game: A Case Study." *The Journal of Abnormal and Social Psychology* 49(1):129–134.

Havens, Timothy, Amanda D. Lotz, and Serra Tinic. 2009. "Critical Media Industry Studies: A Research Approach." *Communication, Culture and Critique* 2:234–253.

Heath, Christian and Paul Luff. 1992. "Disembodied Conduct: Interactional Asymmetries in Video-Mediated Communication." Pp. 35–54 in *Technology in Working Order: Studies of Work, Interaction, and Technology*, edited by G. Button. London: Routledge.

Heath, Christian, Hubert Knoblauch, and Paul Luff. 2000. "Technology and Social Interaction: The Emergence of 'Workplace Studies'." *British Journal of Sociology* 51(2):299–320.

Heath, Christian, Paul Luff, and Marcus S. Svensson. 2002. "Overseeing Organizations: Configuring Action and its Environment." *British Journal of Sociology* 53(2):181–201.

Heritage, John. 1984. *Garfinkel and Ethnomethodology*. Cambridge: Polity Press.

Jayyusi, Lena. 1988. "Toward a Socio-Logic of the Film Text." *Semiotica* 68:271–296.

Koschmann, Timothy, Curtis LeBaron, Charles Goodwin, Alan Zemel, and Gary Dunnington. 2007. "Formulating the Triangle of Doom." *Gesture* 7(1):97–118.

Laurier, Eric and Barry Brown. 2011. "The Reservations of the Editor: The Routine Work of Showing and Knowing the Film in the Edit Suite." *Journal of Social Semiotics* 21(2):239–257.

Lohr, Lenox R. 1940. *Television Broadcasting: Production, Economics, Technique*. New York: McGraw-Hill.

Luff, Paul, Jonathan Hindmarsh, and Christian Heath. 2000. *Workplace Studies: Recovering Work Practice and Informing Systems Design*. Cambridge: CUP.

Macbeth, Douglas. 1999. "Glances, Trances, and their Relevance for a Visual Sociology." Pp. 135–170 in *Media Studies: Ethnomethodological Approaches*, edited by P.L. Jalbert. Lanham, NY: University Press of America.

MacNeill, Margaret. 1996. "Networks: Producing Olympic Ice Hockey for a National Television Audience." *Sociology of Sport Journal* 13:103–124.

Marriott, Stephanie. 2007. *Live Television. Time, Space and the Broadcast Event*. London: Sage.

Mayer, Vicky, Miranda J. Banks, and John T. Caldwell. 2009. *Production Studies: Cultural Studies of Media Industries*. London: Routledge.

Merleau-Ponty, Maurice. 2012. *Phenomenology of Perception*. Translated by D. Landes. London: Routledge.

Mondada, Lorenza. 2003. "Working with Video: How Surgeons Produce Video Records of their Actions." *Visual Studies* 18(1):58–73.

Mondada, Lorenza. 2009. "Video Recording Practices and the Reflexive Constitution of the Interactional Order: Some Systematic Uses of the Split-Screen Technique." *Human Studies* 32(1):67–99.

Morris, Jeffrey M. 2008. "Structure in the Dimension of Liveness and Mediation." *Leonardo Music Journal* 18:59–61.

Perry, Mark, Oskar Juhlin, Mattias Esbjörnsson, and Arvid Engström. 2009. "Lean Collaboration through Video Gestures: Co-Ordinating the Production of Live Televised Sport." Pp. 2279–2288 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*. New York: ACM Press.

Perry, Mark, Oskar Juhlin, and Arvid Engström. 2014. "Dealing with Time, Just in Time: Sense-Making and Clip Allocation in Multi-Person, Multi-Stream, Live Replay TV Production." Pp. 262–286 in *Studies of Video Practices: Video at Work*, edited by M. Broth, E. Laurier, and L. Mondada. New York: Routledge.

Popova, Kristina. 2018. "Ethnomethodological Studies of Visuality." *Ethnographic Studies* 15:23–37. https://doi.org/10.5281/zenodo.1475767.

Rabiger, Michael. 1998. *Directing the Documentary*. 3rd ed. Oxford: Focal Press.

Silk, Michael, Trevor Slack, and John Amis. 2000. "Bread, Butter and Gravy: An Institutional Approach to Televised Sport Production." *Sport in Society* 3(1):1–21.

Stoddart, Brian. 2006. "Sport, Television, Interpretation and Practice Reconsidered: Televised Golf and Analytical Orthodoxies." *Sport in Society* 9(5):865–878.

Verna, Tony. 1987. *Live TV: An Inside Look at Directing and Producing*. London: Focal Press.

vom Lehn, Dirk, Christian Heath, and Jon Hindmarsh. 2001. "Exhibiting Interaction: Conduct and Collaboration in Museums and Galleries." *Symbolic Interaction* 24(2):189–216.

Williams, Brien R. 1977. "The Structure of Televised Football." *Journal of Communication* 27(3):133–139.

ABOUT THE CONTRIBUTOR(S)

Mark Perry is a Reader in Interactive Systems at Brunel University London. He studies the design and use of social and collaborative media combining ethnomethodological, conversation analytic, and ethnographic approaches with participatory design techniques. Mark's work is highly interdisciplinary, with research interests ranging across technology use in video production, financial interactions, social media, family life, and mobile settings. He holds a Royal Society Kan Tong Po International Fellowship exploring perspectives on "moneywork" in China.

Mathias Broth is a Professor of Language and Culture at Linköping University. He studies human interaction from an ethnomethodological and conversation analytic perspective, taking a particular interest in multimodal aspects of language and embodiment, technologically mediated interaction, and interaction in mobility. Activities studied include TV-production of studio interviews, news production in newsrooms, and guided walks. He is co-editor of several edited collections, for example, *Studies of Video Practices: Video at Work* (Routledge 2014) and *On the Road: Communicating Traffic* (special issue of *Language & Communication* 2018).

Arvid Engström is CEO and Co-Founder of Liveling AB. He holds a Ph.D. in Human-Computer Interaction from Stockholm University and an M.Sc. in Media Technology from the Royal Institute of Technology in Stockholm. His academic work includes ethnographic studies, user research, and design of technologies for media production.

Oskar Juhlin is Professor at Stockholm University at the Department of Computing and Systems Sciences. He is also Guest Professor at the Department of Information Science and Media Studies at Bergen University. He has an interdisciplinary background with an interest spanning technology and social science in areas such as fashion, road traffic, animal interaction, and video interaction.