Modes of knowing: Video research, multimodality and the problem of elusive knowledges

Article *in* Organizational Research Methods · February 2018

DOI: 10.1177/1094428116657394

CITATIONS
17

READS
102

3 authors, including:

Maria Laura Toraldo
University of Milan
35 PUBLICATIONS **58** CITATIONS
SEE PROFILE

Gazi Islam
Grenoble École de Management
104 PUBLICATIONS **834** CITATIONS
SEE PROFILE

MODES OF KNOWING: VIDEO RESEARCH AND THE PROBLEM OF ELUSIVE KNOWLEDGES

Maria Laura Toraldo
Post-doctoral Researcher
Institute of Marketing and Communication Management
Università della Svizzera italiana (USI)
Via G. Buffi 13
6900 Lugano, Switzerland
Phone: +41 58 666 4112
Email: maria.laura.toraldo@usi.ch

Gazi Islam
Associate Professor
Department of People, Organizations and Society
Grenoble Ecole de Management
12 Rue Pierre Semard
38000 Grenoble, France
Email: Gazi.ISLAM@grenoble-em.com

Gianluigi Mangia
Associate Professor
Department of Economics, Management and Institutions
University Federico II of Naples
Complesso Universitario di Monte Sant'Angelo - Via Cintia, 21
80126 Naples, Italy
Email: mangia@unina.it

BIOS

Maria Laura Toraldo is research assistant at the Università della Svizzera italiana, (USI), Switzerland. She worked as researcher at Grenoble Ecole de Management, France and she held visiting positions at Warwick Business School and Essex Business School (UK). Her current research interests include individual and collective work identities in organizations. She is interested in methodological issues surrounding the process of research, such as knowledge creation and interpretation during the practices of research.

Gazi Islam is associate professor at Grenoble Ecole de Management. He has served as faculty at Insper, Tulane University, and the University of New Orleans. His current research interests include the organizational antecedents and consequences of identity, and the relations between identity, group dynamics and the production of organizational cultures.

Gianluigi Mangia is associate professor of organization studies at the University of Naples Federico II, where he currently teaches organization theory and information systems. He was Visiting Scholar at Cardiff Business School and at London School of Economics (UK). Gianluigi's research interests are around power and resistance in organizations, and more recently on organizational research methods.

ABSTRACT

The current paper argues that video-based methodologies offer unique potential for multi-modal research applications. Multi-modal research, further, can respond to the problem of "elusive knowledges", that is, tacit, aesthetic, and embodied aspects of organizational life that are difficult to articulate in traditional methodological paradigms. We argue that the multi-modal qualities of video, including but not limited to its visual properties, provide a scaffold for translating embodied, tacit and aesthetic knowledge into discursive and textual forms, enabling the representation of organizational knowledge through academic discourse. First, we outline the problem of representation by comparing different forms of elusive knowledge, framing this problem as one of cross-modal translation. Second, we describe how video's unique affordances place it in an ideal position to address this problem. Third, we demonstrate how video-based solutions can contribute to research, providing examples both from the literature and from our own applied case work as models for video-based approaches. Finally, we discuss the implications and limitations of the proposed video approaches as a methodological support.

Organizational scholars increasingly recognize the importance of forms of knowledge that are difficult to articulate explicitly in language (Ewenstein & Whyte, 2007a; Hakanson, 2007; Taylor, 2002). Such 'elusive' knowledges are key to organizational functioning, yet are difficult to replicate, thus providing a source of sustained organizational advantage (e.g. Berman, Down & Hill, 2002). While scholars struggle to convert elusive knowledges from lived experience into academic concepts (Jarzabkowski, Bednarek & Le, 2014), organizations also rely on mastering elusive knowledges to understand, develop and adapt their own processes (Islam, 2015a).Working from different traditions, scholars have framed elusive knowledges in diverse ways, from research on tacit forms of knowledge (Ambrosini & Browmann, 2001; Lam, 2000) to aesthetic approaches emphasizing sensory ways of knowing (Warren 2008; Taylor & Hanson, 2005) to embodied and practice-based knowledge that stresses the materiality of cognition (Cunliffe & Copeland, 2012; Carlile, 2002).

Each of these approaches recognizes the importance of non-linguistic knowledge in organizational life (Meyer, Hollerer, Jancsary & van Leeuwen, 2013). Increasingly, organizational scholars have noted these difficult-to-articulate aspects of organizational knowledge at the micro-level, where subtle nuances and tacit features abound (e.g. Seidl & Whittington, 2014; Ambrosini & Browman, 2001). Such approaches question how to move beyond formulaic descriptions of organizational life to describe organizing reflexively, questioning the degree to which operationalization can be successful. If knowledge is embedded in material practices (cf. Nicolini, 2012; Nicolini, Gheradi & Yanow, 2003), it may be both empirically manifested and yet difficult to put into words. For such forms of knowledge, speech is both practically difficult, where actors themselves struggle to find words to represent aesthetic, tacit, or embodied knowledge, and epistemically problematic, where scholars' tools cannot fully represent elusive knowledges. The current paper focuses

on how specific methodological tools can explore knowledge that is difficult to articulate through verbal or textual accounts.

Increasingly, organizational scholars are exploring video-based methods, valuing video's ability to explore the interactions of humans with material settings, and to reveal facets of non-verbal communication (Fele, 2012; LeBaron, 2005; Dant, 2004). By recording facial expressions, habits, postures and gestures (Mondada, 2006), video seems to facilitate access to the 'habitualized knowledge implicit in social action' (Knoblauch & Tuma, 2011:12), promoting reflexivity around automatized or underlying bases of organizational action. This aspect has led scholars to use video in conjunction with interviews or ethnography (Erickson, 2011; Pink, 2004; Heath; 2010). As elaborated in our review of alternative uses of video in research, video offers a plurality of applications, from complementing and supporting fieldwork, to stimulated recall or sense-making among participants.

For example, as Pink (2004) argues, integrating videos into interviews combines visual and verbal knowledge, allowing interviewees to produce narratives while visually making sense of situated knowledge. In this way, 'video becomes an agent in the process by which knowledge is produced' (Pink, 2004:64).

While video has been valued for representing bodies and fine-grained actions by means of sequentially-linked images (e.g. Knoblauch & Schnettler, 2012), understanding how video promotes reflexivity and taps into elusive knowledges (e.g. Heath, Hindmarsh & Luff, 2010) remains less explored. In particular, only scant research has explored the translation of elusive knowledges from lived experiences into speech (e.g. during data collection) and text (and ultimately, in published scholarship) during the research process, and how multimedia tools such as video might facilitate this process.

We argue that video can support this process through leveraging its unique multimodal properties: its fusion of image, sound and movement. Beginning with the methodological/epistemological problem of how elusive (i.e., embodied, tacit, aesthetic) knowledges come to be studied in the first place, we explore how video assists researchers in a.) Accessing knowledge through the direct use of video as a research tool, either alone or in conjunction with other methods and b.) Promoting the articulation of elusive knowledge by organizational members themselves, thus developing organizational awareness and reflexivity. This dual motivation acknowledges that both researchers and organizational members struggle to ascribe words to their experiences and actions (e.g. McDermott, Gospodinoff, & Aron, 1978), and therefore producing discourse from embodied experience is not only a scholarly, but also a practical, problem.

This paper's contribution to organizational research methods is threefold. First, we advance knowledge on video-based methods for data collection, using the concept of multimodal affordances to show how three broad alternatives offer practical affordances for articulating elusive knowledge. Second, we connect currently dispersed threads among aesthetics, tacit knowledge, and embodiment literatures, by framing the articulability of organizational knowledge as a common conceptual concern around 'elusive knowledges'. Finally, we use video's multimodality to make inroads into the conceptual problem of articulability, arguing that crossing modal barriers facilitates knowledge translation across different modalities.

Specifically, we suggest that video methodologies facilitate access to embodied practical knowledge, not because video captures the 'real' of organizational life (cf. Jones and LeBaron, 2002), but because it promotes cross-modal translations that can be productive of new knowledge by promoting reflexivity. We propose that, in recording human actions, video can act as a modal mediator between discourse and lived experience. Video-based

methodologies provide a potential solution to an enduring dilemma: how to articulate in conceptual and abstract language knowledge which is immediate, embodied and aesthetic? Such articulation requires reflexivity around elusive knowledges, facilitated by working across modal boundaries.

Our argument is structured as follows. First, we describe the methodological problems posed by 'elusive knowledges', linking these to articulability, and then arguing that they share a common relation with multimodality. Then, we describe how video-based methods complement traditional discursive approaches, providing avenues for accessing knowledge by working across diverse modalities. After laying out this theoretical background, we illustrate three models through which videos can be integrated into qualitative research. The final section discusses the practical and epistemological issues raised by working across modalities in this way, with implications and future directions for video and, more generally, multimodal research methods.

BEYOND WORDS: TACIT, AESTHETIC AND EMBODIED FORMS OF ELUSIVE KNOWLEDGE

We term "elusive knowledges" those forms of knowledge that escape literal representation through discourse including alphanumeric symbols. While the term itself is new to organizational scholarship, it draws together existing strands emphasizing situatedness and embodiment, dispersed among tacit (Ambrosini & Bowmann, 2001; Lam, 2000), aesthetic (Warren, 2008; Taylor, 2002) and embodied knowledge (Kupers, 2013; Cunliffe and Coupland, 2012). Despite the overlap between these bodies of literature, each displays a distinct historical and theoretical trajectory, shares a common preoccupation with how knowledge is encoded and expressed and has been confronted with the elusiveness of knowledge-in-practice (e.g. Orlikowski, 2010), that is, knowledge that is enacted in ongoing activity. It is of no surprise that each has been recognized as important to practice-based

perspectives, nor is the close relationship of elusive knowledges with practice theories in organizations more generally surprising. We note this relation to practice perspectives and return to these in the discussion, although a full theoretical exploration of the issue is beyond the scope of our methodological focus here.

While mapping each of these distinct trajectories lies outside our current scope, we focus on the common *methodological* challenge that these literatures have faced, the problem of articulation. Our brief survey of each of these areas serves to foreground how this problem has resurfaced, thereby suggesting a common concern around multimodality.

In short, the unspeakability of beauty, the automaticity of habit and the visceral sense of embodiment represent distinct epistemological issues involving representation; however, they pose similar methodological difficulties regarding the production of scholarly knowledge and text out of the flux of "life as the experience of thrownness" (Weick, 2004, p.659). Central to, but not exclusive to, organizational practices perspectives (e.g. Nicolini, 2012), tacit, aesthetic and embodied knowledges are analytically distinct concepts, though they co-contribute to the elusive aspect of lived experience, and they may combine in practice. While organizational practices situate and deploy elusive knowledge (Nicolini et al, 2003), the concept is difficult to pin down. Indeed, except for a few contributions (e.g. Jarzabkowski, Bednarek & Le, 2014; Koschmann and LeBaron, 2002; Ambrosini & Browman, 2001), there is little academic research on how such knowledges are represented.

Bringing knowledge into articulability is a process that is accomplished interactionally (Koschmann and LeBaron, 2002), where material and social contexts act as resources that enable embodied knowledge to be articulated. Material environments provide communicative resources, located in people's verbal and non-verbal interactions, such that jointly studying verbal and non-verbal messages is key for insight into human action (Jones & LeBaron, 2002).

The idea that representing action involves both verbal and non-verbal expression (Jones & LeBaron, 2002) suggests, furthermore, that interactive processes occur multimodally (Koschmann and LeBaron, 2002). Articulation into verbal expression is one possible way to represent knowledge; however, diverse modes – such as gestures, body movements, images or drawings – provide opportunities to complement speech or verbal communication, representing diverse possibilities for representation. Combining diverse modalities informs the research process, insofar as multimodality depicts how the material environment, bodies and gestures are integrated with conversational interactions. Thus, to grasp multidimensional interaction, multimodality becomes crucial. By presenting different forms of knowledge – embodied, aesthetic, tacit – in diverse ways, a space of reflexivity can facilitate the passage from elusive knowledge to articulability.

The tacit knowledge literature (Shamsie & Mannor, 2012; Hakanson, 2007; Tsoukas, 2003) has explicitly noted this focus on 'articulation'. Polanyi (1962), and later Nelson and Winter (1982) contrasted tacit knowledge with knowledge that is expressed. Knowledge can be tacit because it escapes conscious awareness, and yet it is internalised and expressed in practices (e.g. Collins, 2007). Tacit knowledge, an individual and collective resource impacting on performance (Shamsie & Mannor, 2012; Berman, Down & Hill, 2002), is particularly useful for studying the micro-interactions that give rise to individual and collective performance (e.g. Fele, 2012; Hindmarsh & Pilnick, 2007; Koschmann & LeBaron, 2002).

As Ambrosini & Browmann (2001) emphasize, tacit knowledge is both critical to organizational functioning and strategic advantage, as well as being notoriously difficult to operationalize or measure. Polanyi (1962: 4) describes articulation as mobilizing speech and other symbolic forms, arguing that "language is the instrument for articulation", yet that "we can know more than we can tell" (p.80). Similarly, Hakanson (2007:61) defined articulation

as "the process of expressing tacit knowledge into some socially shared code or symbolic representation – in the simplest case, ordinary natural language". Tacit knowledge scholars have acknowledged that making tacit knowledge explicit involves a process of transformation (e.g. Nonaka & Krough, 2009), although differences of opinion exist as to the extent to which the conversion to explicit knowledge is possible (cf. Donaldson, 2001). Some research emphasizes video's ability to access moments of articulation in microethnographic methods (e.g. LeBaron, 2005), for instance, as actors attempt to understand each other in emergency operator interactions (Fele, 2012), or as medical students come to articulate problem-based knowledge (Koschmann & LeBaron, 2002). Video methods therefore seem particularly suited to exploring tacit knowledge in organizations.

While tacit knowledge represents knowledge that is elusive due to its embeddedness in practices and routines (Collins, 2007), aesthetic knowledge involves the "felt meanings" (Warren, 2008) of organizational environments, including sensory knowledge and judgments of beauty or ugliness (Taylor & Hanson, 2005; Strati & Montoux, 2002; Strati, 1992). Aesthetic knowledge, or judgments of "taste", develop from sensory intuitions into judgments of form that become linked to aesthetic pleasure/displeasure (Gagliardi, 1996; Strati, 1992). Aesthetic judgments involve both subjective intuitions and normative force (Martin, 2002), marked by both personal connection and an implicit suggestion that they can be generalized (Kant, 1789/1911). From this arises the idea that "taste cannot be discussed", although organizational processes frequently depend on creating shared aesthetic visions, and discussing taste is both common and often essential to organizational functioning (Endrissat et al, 2016). Such communication is instrumental to organizational performance, particularly around team dynamics (Yaniv, Choshen-Hillel, & Milyavsky, 2010), design processes (Endrissat et al, 2016) and organizational culture and power dynamics (Wasserman & Frankel, 2011). Such processes have been scrutinized in craft studies of flute makers (Cook &

Yanow, 1993), perfumers (Islam et al, 2015) and restaurant chefs (Fine, 1992), among others. In each of these cases, special forms of communication and socialization are necessary to transmit this type of knowledge.

The importance of articulating and coordinating aesthetic knowledge sits uncomfortably alongside the idea that that aesthetics requires bracketing analysis and "being in the present moment" (Taylor, 2013:31). As Taylor and Hanson (2005) argue, discursive representations require a division into the signifier and signified, in contrast to the holism and unity of aesthetic experience, leading to difficulty when articulating aesthetic experience in organizations, a problem known as "aesthetic muteness" (Taylor, 2002). This difficulty of discussing aesthetic knowledge creates deep methodological problems (Warren, 2008), and is not only an issue for organizational members but also for observers. Such difficulties include the impossibility of breaking holistic experiences into measurable units, the necessity of using "elusive, poetic language" (Gagliardi, 1996: 576) to access aesthetic experience and the importance of researcher empathy in aesthetic matters. Recent applications of aesthetic approaches to deal with epistemic and other "objects" in organizations (Ewenstein & Whyte, 2007b), acknowledge the importance of cross-modal issues and material artefacts when studying aesthetic knowledge. Ewenstein & Whyte's (2007b) study of visual objects in architectural design illustrates how the aesthetic is embodied in tools, such as videos or other research instruments that mediate action. Following this conception, videos combine properties of technical and aesthetic objects, storing and transmitting knowledge while creating a material context wherein actions are made sensible and intelligible. Such studies suggest that video can be a key tool in exploring aesthetic aspects of organizing.

While both tacit and aesthetic knowledges are deeply embodied, the notion of embodied cognition is analytically important in its own right. Embodied cognitive perspectives (e.g. Clarke and Cornelissen, 2011; Mitchell & Mitchell, 2011; Sutton, 2006)

stress that knowledge is largely dependent on a material scaffold, including bodily orientations and analogy, as well as being dependent on tools and artefacts, physical spaces and interactional systems (Smith & Semin, 2004; Shore, 1996). Individuals and collectives use these structures to "lean on the world" (Smith & Semin, 2004), framing cognition as structured around both internal bodily and external material structures (Islam, 2015b; Sutton, 2006). Embodiment perspectives vary in their focus on the body as a locus of "internalization" or a horizon for lived experience (e.g. Cregan, 2006), and the 'external' embodiment of knowledge in technologies, tools and environments (e.g. Hutchins, 1995). Both perspectives stress that capturing the *materiality* of thought, rather than simply its content or meaning, is fundamental to understanding knowing as a situated phenomenon (Smith & Semin, 2004).

Embodied knowledge is an important resource in understanding organizational functioning, particularly around the ways in which bodies act so as to provide a resource for knowing and transmitting knowledge (e.g. Hindmarsh & Pilnick, 2007). Embodied knowledge has appeared in a plethora of diverse organizational approaches, and it is linked to organizational sensemaking processes (Kupers, 2013; Cunliffe & Couplan, 2012), leadership (Ladkin & Taylor) and organizational diversity (Thanem, 2006) processes. Techniques allowing researchers to explore embodiment, accessing both bodily expression and the material supports for knowing, would thus be useful to these literatures. This suggests that video would be well suited for studying embodied organizational knowledge.

Tacit, aesthetic and embodiment perspectives illustrate the varied challenges that researchers face around elusive knowledges. The unspoken side of internalized knowledge (tacit), the ineffable nature of judgements of taste, beauty and appropriateness (aesthetic) and the implicit embodied metaphors (Lakoff and Johnson, 1980) that underlie much of cognition (embodied) together relegate a large part of experience to being opaque to articulation,

begging the question of how researchers can theorize such phenomena using a variety of methods while doing justice to lived forms of knowledge.

It thus becomes vital to develop representational strategies to articulate aesthetic and embodied experiences, while avoiding what Melcher & Schooler (1996) term 'verbal overshadowing', where verbal representation dominates and occludes lived experience. In analysing the case of wine tasting, where sensory and perceptual experiences – such as taste and olfaction – are fundamental, Melcher & Schooler (1996) show that verbalization regarding smell or taste memory, together with the difficulty of describing these perceptual experiences, limit one's ability to fully appreciate the complexity of experience. This provides a methodological warning and a call to protect the perceptual texture and richness of elusive knowledge. Below, we argue that, through its multimodal features, video can offer opportunities to explain and articulate elusive knowledges.

RESEARCH, TEXTUALITY, AND TRANSLATING ACROSS MODALITIES

Translating lived experience into discursive and textual communications is central to organizational research, given that written texts are the primary means and product of our research projects (cf. Jazarbkowski et al, 2014). Research relies largely on the use of speech (e.g. interviews and conversations) and texts (e.g. archival and other written material) as the primary means of investigation, as these take advantage of natural language's ability to abstract and generalise (e.g. MacDougall, 2011), making discourse "theory-friendly". Nevertheless, discursive accounts may be limited in their ability to describe tacit aspects of actions, involving practical knowledge in embodied routines and directly pertaining to aesthetic experiences that are only imperfectly articulable. The situation of translating elusive knowledge into research thus involves "both rescue and irretrievable loss—a kind of death in life – in the making of texts from events" (Clifford, 1986; 115).

The limits of discursive accounts to represent experiences is thus an important methodological issue. Aesthetic muteness leads to difficulty translating experiences through traditional methods and tools – such as interviewing techniques, content analysis and the like, blocking the translation of 'feeling' into 'thinking' (Taylor, 2002). Aesthetic muteness is particularly acute in organizations because such knowledge may be considered illegitimate or without value within business organizations specifically (Taylor, 2002). One effect of muteness could be the inability to generate accounts from organizational members by asking them merely to describe their experiences and knowledge through verbal communication.

The difficulty in the translation from elusive knowledge into articulability is twofold. First, there is the difficulty of articulating organizational members' elusive knowledge by researchers; researchers may need to record and combine different sources of information, using videos to provide a support in this endeavour. The second difficulty concerns organizational members themselves, in terms of their difficulty in articulating their own tacit or elusive knowledge in verbal form. Organizational members may know '*how* to do' specific actions, but not necessarily how to translate a holistic experience into words. This may be true when knowledge is internalized, embodied in habits, routines or skills, and can be translated into language only at the expenses of a loss of richness (Hakanson, 2007). Similarly, with embodied and aesthetic experiences, the issue arises of how the 'felt sense' can be made understandable (Taylor, 2002). In these circumstances, researchers can facilitate articulation, for example, by creating a critical distance between members and their actions, objectifying these actions on video. The medium of video thus becomes a material support, mediating between the unarticulated embodied knowledge of organizational members, their articulation of this knowledge and the researchers' activity.

Such material supports function because of their modal affordances (e.g. Demir, 2015), as we elaborate below. The problem of articulation is, in part, the problem that elusive

knowledge involves different modalities of expression from the discursive representations that seek to communicate them. Subjective realities are 'packaged' and made understandable according to sensory modalities that afford different bases for experience and conceptualization (e.g. Islam et al, 2015). Video, in combining different modalities such as images, sounds, texts, becomes relevant in this context.

Video, as a *conjunction* of different modalities, integrates diverse communicative resources multimodally (e.g. Mondada, 2006). By recording sounds, bodies and movements, video captures gestures, facial expression, social interaction and embodied actions, providing tacit and embodied aspects of knowledge for later analysis. Researchers can appreciate embodied knowledge, for example, both by the direct analysis of multimodal data or by using multimodal data as a stimulus to invoke participant reflexivity (cf. Mondada, 2006). If multiple modalities facilitate the integration of diverse modes of expression, it becomes crucial to understand how research methods can bridge different modalities of experience and elicit knowledges.

The concept of multimodality stems from the idea that experience is synergistically composed of different sensory modalities which can be separated through analysis, but coexist in a field of experience (Pink, 2011). Different modalities of experience (e.g. movement, texture, temperature, tone) realize different potential meanings and forms of representation (Stivers and Sidnell, 2005). As observed by Price, Jewitt and Brown (2013) meaning arises within multimodal representations, from the unique contributions of each modality as it is enmeshed with the others. Writing, images, sound, speech, posture, embodied interaction and the like each "afford" different possibilities for the expression of meanings. The affordance notion (Gibson, 1986) has recently been used to describe human interaction from a materiality perspective (Demir, 2015; Jarzabkowski & Pinch; 2013). When applied to communication and representation, 'modal affordances' are used to explain

'what is possible to express and represent in a mode' (Bezemer and Jewitt, 2010: p. 6). Possibilities of representations presented by a particular mode can be explained through these affordances, which emerge from the material properties of the medium in a given action situation (Gibson, 1986).

Framing the question of articulability through multimodality recognizes that elusive knowledges resist theorization, in part, because the affordances of written and spoken language may not easily represent experiences lived under diverse modalities. While not completely "mute" (Taylor, 2002), participants' speech and researchers' texts represent such knowledge only with great difficulty. This is not to say that discourse is useless for elusive knowledges, but that invoking other modalities within the research process may facilitate or augment the awkwardness of discourse, allowing it to more closely approximate the lived conditions of research participants. Recent work has given credence to the idea that individuals work not only within modalities but across modalities to achieve a collective outcome. For instance, Jarzabkowski, Burke & Spree (2015) demonstrated how speech works together with material artefacts in strategic coordination. Islam et al (2015) showed how design teams relied on visual artefacts to design scent products, using visual affordances where scent-knowledge proved difficult to codify. Demir (2015) demonstrates how 'bundled' affordances complement each other in strategic events. In each case, action is achieved by working across the affordances of each modality, suggesting the importance of harnessing multimodal affordances in organizational activities. Methods that capture the multimodal affordances of data promote access to these aspects of organizational life.

Multimodal affordances describe the potentialities or constraints of different modes – what it is possible to represent in a straightforward way with a given mode, and what is difficult or even impossible to represent, possibilities that depend both on material properties and social action around a given medium (Price et al, 2013). Multimodal affordances signal

that information is materially anchored, and that modes of presentation matter (Streeck et al., 2011). For example, the textual or discursive mode may impose a flow-like structure on phenomena as it 'describes the linear progression of discursive objects in a narrative form' (Saussure, 1974). Written texts embody structural elements of knowledge sequentially, where language unfolds diachronically to reveal unfolding discursive elements (McCanles 1982). The logic of sequence in time is therefore crucial in texts, and the mode of writing offers potential for modes of representation that are linear in nature (Bezemer and Jewitt, 2010). By contrast, visual images are synchronic, freezing a set of related visual elements in a "paradigmatic" relation or snapshot in frozen time (Barthes, 1982). Visual language has its peculiar form of communication, and it is valued for its immediacy and ability to convey information holistically, in contrast to the linearity and sequentiality of verbal language (Meyer, et al., 2013).

It follows from this that images and textual accounts can both be used to perform the same 'fundamental systems of meaning'; however, each can provide access to different knowledge, whether independently or in a multimodal combination (Meyer et al, 2013: 4). Importantly, video involves both a diachronic flow and a synchronic structure, being both temporally extended and composed of visual images, and thereby providing a crucible for cross-modal translations of meaning.

Video takes advantage of the fact that different modal combinations offer unique insights into lived experiences. Cross-modal studies have shown how multimodality in interview techniques (telephone versus face-to-face modalities; Silvester, Anderson, Haddleton Cunningham-Snell & Gibb, 2000), and in representing affective relations (visual versus olfactory modalities; Islam et al., 2015) can provide a more holistic understanding of an object of study, articulating tacit knowledge by triangulating across modalities. Because video can provide both synchronic, image-based understandings and diachronic, narrative

unfoldings of events (cf. Mondada, 2012), it can illustrate both organizational processes and structures. As a visual depiction, video encodes the complex aesthetic and embodied relations within practices and interactions, revealing more than words alone. Yet, by recording participant interaction over time, video captures temporally-extended, unfolding actions that allow translation into a written narrative.

VIDEO-BASED METHODS

Video-based methodologies have been receiving growing scholarly interest (e.g. Luff & Heath, 2012; Knoblauch & Schnettler, 2012), and growing recognition of video's relevance in studying social behaviour (Fiele, 2012; Knoblauch, H. & Schnettler, 2012; Gylfe et al., 2016). As illustrated by Dant (2004), video can reveal shared social worlds, constituted by everyday embodied knowledge that is contained within the relatively unconscious, ordinary "ways of doing things" (Dant, 2004, p. 43). Similarly, microethnographic perspectives have emphasised the use of video in representing tacit and non-verbal behaviours (e.g. Streeck & Mehus, 2005; LeBaron, 2005). This work acknowledges the aesthetic and embodied possibilities of sensory methodologies, while focussing less on the catalytic function of alternative methods for bridging diverse modes of knowledge.

Video-based methods can address the problem of elusive knowledges through video's unique multimodal affordances, as deployed in diverse ways in the research process. As argued above, multimodality opens up reflexive spaces by allowing comparison between modes, rendering visible phenomena that are opaque to a single mode. Underlying meaning systems are constituted between modalities that are enmeshed in a holistic interaction situation (Pink, 2011). Accessing these interstitial spaces between modalities is a strategy for reflecting on the meaning of such situations.

In addition, because the problem of elusive knowledge is a problem both of researcher access to participant experience and of participants' reflection on their own embodied

experience, video-based methods allow different forms of reflexive work in articulating knowledge. Because video is internally multimodal, such work may be done via the medium of video; however, modal comparison may also work externally by comparing video with other data sources whose modalities are not captured by video. Thus, taking account the agent of reflection (researcher and participant) and the source of modal comparison (internal to video and triangulation with other sources) provides an array of distinct possibilities for reflexive work around elusive knowledges via video.

ALTERNATIVE USES OF VIDEO IN RESEARCH

Below, we describe three approaches to exploring different aspects of elusive knowledges via video. Each approach configures, in distinct ways, the relationship between researcher, participant and practice, such that cross-modality offers different possibilities for each type. We present the components, constraints and possibilities of each model, showing how each model addresses a particular type of research problem. Table 1 summarizes the main features of each of these approaches.

_____

Table 1 about here

_____

To illustrate how these models of analysis would work in practice, we provide a real-world example of a video-based study carried out by the second author (Islam & Benameur, 2016). The study concerned collective cognition and tacit knowledge in a professional volleyball team, where video was used extensively in conjunction with other methods for data collection. In using a professional sports case to illustrate micro-organizational phenomena, we follow on recent uses of sports teams in the organizational literature (e.g. Lok & deRond, 2013), and particularly around tacit knowledge (Shamsie & Manor, 2012; Berman et al, 2002). As Wolfe et al (2005) argue, sports organizations provide ideal settings for

organizational studies, "mirroring" the world of work in ways that are observable, dynamic and vivid. We use this example as a common thread running through the three cases in order to examine how the data may be approached differently depending on the approach, or used collaboratively to reach different aspects of the field.

Video as raw data. The most common use of video is as a source of raw data (Jewitt, 2013). Video is used to represent real-time behaviours by actors, as well as to represent social interactions (Knoblauch, 2012), to document tacit and explicit phenomena that are difficult to capture with texts or interviews (Fele, 2012), and to access 'naturally-occurring' aspects of social phenomena (Mondada, 2012). Video data provide rich material for subsequent qualitative coding, while freeing the researcher's attention for other data collecting activities. Because of this plethora of uses, video-based approaches have become a popular and recommended form of data collection (Heath, Hindmarsh and Luff, 2010; Mondada, 2006).

From the perspective of cross-modality, video data offer not only a new modality of data, but are themselves multimodal (Mondada, 2006). Video data involve different modes of presentation, integrating image, movements and facial expressions, all unfolding contiguously over time. The versatility of this combination allows researchers to explore elusive features of social interaction revealed by modal interactions in a particular situated action. Such interactions are complex mixtures of speech, gesture and tool use and thus demand research instruments that are able to detect these elements in their co-occurrence. By comparing insights across modalities, insights can be gleaned from the complementary affordances of each modality, insights that would be elusive under a single modality under traditional content-based coding (e.g. St. Pierre & Jackson, 2014).

To illustrate using the above mentioned sports project, tacit knowledge and communication were embedded in seconds-long episodes of play, moments difficult to describe in words alone. Indeed, even if players, as interviewees, had clear cognitive schemas

of their play episodes, translating this embodied cognition into language via post hoc interviews would be difficult and methodologically problematic. By videoing these episodes directly, we noted patterns in individual and collective action, such as who interacts with whom and in what form (language, gesture, proximity), as well as how such patterns evolve over time.

Figure 1 illustrates a short action sequence taken from one of the videos. Temporal and audio elements are missing from the illustration, but some basic interactional elements are nevertheless present. For example, as the two players lunge for the ball in the first instant, a tacit division of labour is established, with each player subsequently falling back into pre-established positions. While the 'system' of pre-established positions was openly articulated in interviews, the 'elusive' instantaneous breaking and re-establishing of order would have been difficult to access without video. Parallel to this, eye movements turn from opponents to each other, then to the ball, implying a collective coordination though object-anchoring material supports. This is a coordination that is not consciously registered by the participants (at least during the interviews), but is nevertheless easily identifiable from the video. This situation, which was a routine play, was contrasted with more 'crisis' type situations, with distinctly identifiable attentional, gestural and vocal elements.

———

Figure 1 about here

———

Data of this sort, providing temporal flow, sound and image, offer analytical supports that would otherwise be unavailable through discursive techniques. The direction of a gaze, the intonation of voices, postures, type of gestures and bodily movement all widen the types of semiotic resources available to interpret material contexts. By being naturally cross-modal,

video supports reflection on the material nature of representations, facilitating the study of elusive knowledges.

In this sense, video could be described as a 'multimodal ensemble' (Price et al, 2013), where 'several modes are involved in a communicative event and all of these modes combine to represent a message's meaning'. The valuable aspect of multimodal ensembles is that each mode performs a distinct, yet complementary, display of meaning. Although each mode offers a certain range of modal affordances, multimodal ensembles allow for distributing meaning across diverse modes and therefore demonstrating different aspects of the material environment.

Nonetheless, treating video as raw data may have drawbacks, the foremost being the 'illusion of objectivity' generated by using video. The act of recording implies choosing a point of observation (e.g. LeBaron, 2005); the invisible point of the camera gaze, not brought into evidence at this level of analysis, risks obscuring the production choices made in filming. Relatedly, video captures a limited horizon within their gaze, recording partial slices of the 'reality' sampled while offering an illusion of objectivity. In the volleyball example, this aspect may be less evident due to the highly-circumscribed nature of the action within an official rectangle. Nevertheless, choices of where and when to film may be paramount, and should be acknowledged by the researcher.

Moreover, video as raw data has the limitation of not including organizational members' first-hand interpretations of their experiences as data. Their actions, movement and behaviour are interpreted by the researcher from a purely third-person perspective. As the example of volleyball players shows, when they reflect on their unfolding actions, they do provide an interpretation and it is not purely third-person.

The reflexivity gained is one of modal comparison, aiding interpretation through multimodality. The limitations of video as a modal ensemble, and those of the unique point of

view of the camera and researcher, remain important issues. Thus, researchers may wish to combine video with other data sources, as described below.

Video as triangulation. The second approach uses video in conjunction with other data forms to compare, extend and cross-verify data sources. Part of a multi-pronged strategy of data collection, the goal of triangulation (e.g. Jick, 1979) presumes that a.) Different sources of data provide unique insights into organizational phenomena and b.) Despite these differences, sufficient commensurability exists between sources to use these sources together to build a coherent, empirical story.

The first premise around diverse affordances is based on compensating for the partiality of each modality though a plurality of modalities. The second premise assumes that different modalities, although ontologically different (i.e. it is impossible to faithfully translate an image into a sound, a movement into a sentence, etc.), possess sufficient *structural similarity* to illuminate an underlying structural relation. Because finding commonalities requires abstraction and searching for underlying structures characterizing phenomena, cross-modal triangulation across sources can supplement video analysis to support theory-building. Particularly where knowledge is tacit, triangulation can help infer relations that are less accessible directly.

Returning to the illustrative example, volleyball team videos were used to greater effect when combined with other data types. For instance, player interviews (detailed in approach three), participant observation by one of the authors (a professional trainer), non-obtrusive observation and archived player statistics and historical data were all used as factors in understanding the ongoing set of tacit practices that we considered to be an elusive form of knowledge. The collective knowledge enacted on court, although elusive and embodied in rapid moments of intuitive action, had roots in a longer history, which was explored by triangulating data sources. Why a player would choose one teammate over another one for a

pass, or why one would become flustered or annoyed at a certain juncture in the game, resulted from chains of events occurring in other games, in practice and off the court. While the materiality of practices thus required modalities that were able to grasp the presence of action, any interpretation required triangulation across modalities to show different temporal scales and contexts of information.

In Figure 2, we illustrate three different modes in which the chain of decisions within a given play unfolds. Comparing the video sequence with a schematic sketch drawn by the researcher to summarize each play, and compared with an interview with the player about how, in general, decisions are taken (i.e. chain or circle? Who decides which actions?) thus allows three different perspectives on the same action situation.

Figure 2 about here

Triangulation thus offers both a form of verification across modalities, and of extension beyond a given modality. Video data can build on emerging categories from interviews, observations, field notes and other techniques. As a verification tool, triangulation can be used to establish cross-source consensus. As an extension of findings, other data sources (such as interviews) can be contextualised along with video data and the interpretation of one source of data can be aided by the other, which is particularly beneficial where knowledge is elusive and difficult to articulate.

Conversely, constraints associated with triangulation involve the possible incommensurability between different modalities. While modal differences create the possibility for reflexive cross-modal insights, incommensurability shuts off the translation of insights across modes, or the possibility for structural inference at a higher level of abstraction. Incommensurability creates interpretive difficulties in cases of disagreement

between diverse data sources. Legitimate variations among different modalities, however, can generate discrepancies among data without necessarily indicating a lack of validity. The meanings of convergence and deference across modalities, thus, cannot be taken for granted, and may vary across cases.

Video as reflective artefact. The previous two approaches framed articulating elusive knowledges around reflexivity deriving from multimodal affordances. Comparison across modalities opens a kind of 'second-order' abstraction of concepts across sensory representations. Both approaches located reflexivity in the analyst; that is, researchers gain spaces of analysis by working across media, thereby increasing their own understandings of data.

A third approach places reflexivity in participants' and in researcher-participant exchanges, using video to promote participant reflection on individual and collective actions. Video is used here to facilitate the articulation of organizational actors' internalised knowledge by constructing, remembering and articulating experiences. As an 'artefact', video provides opportunities to study the situated nature of actions (Suchman & Trigg, 1991). Video is used to promote sensemaking around embodied actions, where members reflect on the meanings of what they are doing on the screen. Some contemporary methods, such as video elicitation (e.g. Henry & Fetters, 2012; Jarrett & Lui, forthcoming) and stimulated recall methods (Dempsey, 2010) use stimuli to elicit participant memories, for instance by watching their own recorded actions. Video serves as a mnemonic device to explain and elaborate memories; other uses focus more on sensemaking, creative storytelling or joint theorizing around the meaning of events. In this sense, such methods may resemble 'shadowing' (McDonald & Simpson, 2014) techniques, where shadowed participants provide running commentaries around their own actions. Showing videos to actors while interviewing them enables researchers to simultaneously investigate their intentions and experiences while

directly weighing these against observed practices without interruption. Furthermore, in situations where knowledge is embodied, looking at videos can help to bring this elusive knowledge to articulation.

Using video as a reflective artefact offers organizational members the opportunity to engage in discursive work around elusive knowledges, translating and transforming knowledge in the process. Rather than 'capturing' elusive knowledges, such work recasts them in the mode of observation and reflection, providing new insights rather than objective recall. However, the absence of an absolute perspective on elusive knowledges still allows them to be studied by promoting a plurality of angles and representational forms. Further, this method has the advantage that organizational members themselves can make sense of their knowledge, reflecting on the origins and development of practices and their complex motives, rather than relying on researcher-imposed interpretations.

By making participants analysts of their own action, this method has the advantage of establishing a bridge between theorization and practice. Participants are encouraged to take the position of the analyst, becoming partners in their own theorization. Simultaneously, by scrutinizing recordings of one's own actions, opportunities for self-reflection and critique are opened up. Such possibilities may feed back into organizational practice itself, as the data collection process becomes a reflexive opportunity for organizational members.

In the volleyball team example, the authors used participant-centred video methods as a central data collection technique, based on the premise that, in a highly-specialized and personally engaged field such as professional sports, participants could provide more illuminating accounts than via a researcher-imposed paradigm. Conversely, the deeply embodied and only partially verbal nature of the work made interviews an awkward solution matching poorly with the work environment. By selecting play episodes where players had central roles, we attempted to mediate the language-action gap. Each video was watched

several times, selectively pausing at decisive moments (chosen by both player and researcher) to explore the significance, reasoning and effects of particular actions by the members. By watching each embodied gesture across the different players and at successive moments, participants made inferences around the other members' states, and communication processes more generally. Figure 3 illustrates an extract of this collection.

_____

Figure 3 about here

_____

The constraints of this model include the promotion of post-hoc rationalization by participants, as individuals are motivated to provide 'explanations' of their behaviour independently of their lived experience in the moment. Accordingly, videos should not encourage the justification or defence of behaviour on the video, but rather encourage the use of video in an exploratory manner as an object for reflexion more generally. A further risk involves how member knowledge may itself be affected through repositioning the subjective gaze during the conversation with researchers. The third-person, distant subject position subtly changes the nature of informant knowledge. When 'subjectively' lived experience is displaced into 'objective' discourse by participants, they may lose the phenomenological standpoint from which elusive knowledge originates. Thus, the focus should not be on verification but on comparison between participants' experiences of 'doing' and 'looking'.

Despite these limitations, this third model is appropriate for situations in which individuals struggle to ascribe words to their internalised and embodied knowledge. In these occurrences, video can play an important role in aiding organizational members to articulate meanings and to verbalize elusive forms of knowledge.

APPLYING VIDEO APPROACHES

Applying the above framework, we can discern the extent to which current studies make use of the diverse aspects of video data collection and how these aspects support specific theoretical contributions, strengthening the link between video-based methods and theoretical innovation. The three approaches provided above summarize and give order to the existing literature, ranging from video as raw data to triangulation and to reflexive articulation. Nevertheless, because the multimodality issue has not been central to most of this literature, studies may fail to take advantage of the cross-modal properties of video to make full use of their data. Table 2 provides examples of existing studies which implement aspects of each approach.

Table 2 about here

Video as Data. The first case features the use of video as raw data. Video research of this type is used in particular in close studies of workplace dynamics (Heath and Hindmarsh, 2002; Luff and Heath, 2011). Microethnographic, close analyses of video data are frequently analysed both visually and discursively (e.g. Hindmarsh & Pilnick, 2007; Koschmann & LeBaron, 2002).Videos are here used to film work dynamics, practices or behaviours within specific working contexts, such as architects' offices, control rooms or call centres, and conceptual points are drawn directly from these videoed scenes.

For example, Mondada (2012) documents the work of architects in an office space, focusing on social interactions, with the conceptual objective of illuminating the use of tools and discourses in interaction. Actions such as architectural drawing, discussion, hesitations and contrasts are recorded and micro-analysed to discern how architects interact in 'the discovery of a solution' (p.310). Video data provide multiple types of insights: associating conversational fragments with images and movements, showing interactions and looking at

drawings of sketches, freeze frames and objects. These are analysed as multimodal resources, mutually configuring action, with each offering a distinct point of view. These multimodal data lend insights into the micro-constituents of architects' elusive knowledge.

Procedurally, Mondada (2012) interprets video data from meetings, where architects comment and produce drawings. She listens to informal talks and conversation between them, coding bodily interactions and postures. Artefacts such as plastic models and drawings are analysed from the video as resources used by architects to explain their opinions to other colleagues and to translate their ideas into words. To increase the perspectival richness of the video data, 4 cameras are positioned at different focal points and distances, capturing diverse actors and allowing within-medium triangulation. In particular, a bird's-eye view camera captures interactions that would be virtually impossible to capture via direct observation. The authors concede, however, that even with multiple views it is difficult to capture documentary details from video, suggesting that triangulation would be useful above and beyond video work.

In a second illustration, Fele (2012) documents cooperative work and tacit participation in an emergency cooperation centre. Specifically, the study contributes to exploring how given situational features are made meaningful from within work situations, requiring an analysis of the temporal unfolding of work situations. In this case, 45 hours of video recordings were made in an operation centre handling emergency calls, with calls lasting about one minute each. The authors explored micro-interactions between two people working on different tasks whilst sharing the same workspace. In contrast with the previous example, Fele's videos show how apparently disjointed actions, performed by two participants, are actually highly intertwined. The researcher gains access to ephemeral and subtle knowledge by recording the sequential actions of the emergency workers. Specifically, Fele (2012) strategically focuses on the interactions of operators sitting at each 'box' for a short time, to

highlight subtle cues during collaboration and cooperation episodes that would be lost with more macro approaches. Data are presented in the form of sequential frames taken from the video (see also Heath et al, 2010). By focusing on talk, bodily display, actions, gestures and movement, video features actions that lie beyond what is being said or could be later recounted by members. It also permits an understanding of tacit exchanges between operators, an aspect which might otherwise not be captured in texts or audio recordings.

As these examples illustrate, video as raw data facilitates access to elusive knowledges by augmenting textual data with gestural cues, temporal sequences and nuances that would be obscured in conventional data. Discourse is itself interpreted in the light of such cues, and vice versa, leading to a holistic picture of situational interactions (Goodwin, 1994).

Video as Triangulation. The second model consists of video within a triangulation approach. We focus on studies that use video as a central data analysis strategy, for instance in co-coding video data with interviews and surveys into conceptual categories (e.g. Oliver & Roos, 2007). Numerous studies invoke triangulation, while mentioning video only briefly (cf. video literature review, forthcoming this volume). A typical methodology section might state "A total of 38 hours of videotape were also taken and were used as a backup to the written notes" (Fayard & Weeks, 2007:63), while not explicitly integrating video for its own unique affordances. This suggests that many researchers already collect video data, but how to use video beyond simply "backing up" interviews remains obscure. In some studies, however, triangulation actively involves video by comparing video sequences to fieldnotes and interviews without epistemically privileging one over the other. Here, triangulation implies that theory should be built from a synthesis of data sources, taking advantage both of their overlapping conclusions and their unique insights around a site.

For instance, Paroutis, Franco & Papadopoulous' (2015) analysis of strategic knowledge workshops combined video recordings with researcher diaries, contribution logs and meeting transcripts, used together in data coding. Their theoretical objective was to show how "shift, inertia and assembly" patterns marked how actors structured workshops, and revealed the interactivity between tools, actions and discourses in a micro setting. By paying attention to both visual and discursive material, the authors were able to match the forms of data collection with the specific theoretical contributions of the study.

In terms of data collection, two independent actors (researcher and assistant) took observational notes and noted team contributions in real time. These notes were combined with transcripts that were then read against the video itself. Video segments were divided into 20 discrete segments based on meaning. First and second order coding categories were presented and were then complemented by video snapshot sequences to illustrate each category. The result was to provide a picture of tool use and interactivity that would otherwise have been difficult to depict with words alone. After presenting the visual depiction, transcript excerpts fleshed out the details of the interactions, so that both visual images and flow of discourse were accessible in the article format, facilitating triangulation.

While Paroutis et al (2015) gain new angles on their data through multiple sources, Smets et al (2014) use video data to problematize ethnographic notions of "being there" and to explore the possibilities and problems around different observational positions. Their team-based video ethnography of reinsurance trading in London involved around 400 hours of video over the course of one year, used to support observation by 'producing more detailed, rigorous and defensible insights' (p. 20). Video was used to confirm or counter observations, supporting interpretations of events observed during fieldwork, thus serving a triangulation purpose. By examining the material recorded, the ethnographer remembers events long after the end of the fieldwork.

Beyond verification, however, the ethnographers used video to compare classical issues related to subject positions and observations. For instance, building trust around the ethnographer's physical presence was compared with the complications of video recording, and the observational affordances of capturing movement and dynamics were highlighted by the difficulties of doing this with video. Thus, the research goals of understanding micro-interactions in a trading room was facilitated by forcing reflexivity when switching media and using different observational tools. In this way, the researchers addressed elusive knowledges by complementing traditional fieldwork with the layered, multimodal description afforded by video analysis. Video analysis complements traditional written ethnographic narratives and observation with novel affordances of zooming or freeze-framing details, enabling researchers to become aware of their own effects as an observer in the process. The visual aids afforded by video data bring the gaze itself into the scene and thus highlight the problematic notion of ethnographic "presence". Reflexivity emerges from these multimodal affordances, providing insights into elusive aspects of ethnographic fieldwork. Thus, video's versatility across modes creates conditions for exploring elusive knowledge on the part of participants.

Video as Reflective Artefact. The third approach entails using video as a reflective artefact. Here, video is used to assist participants in articulating their experiences and in making sense of their elusive knowledges. Approaches such as video elicitation techniques (e.g. Henry & Fetters, 2012; Griefenhagen, 2008) could be considered as one such approach. While some elicitation techniques are geared toward memory recovery, often immediately after actions are performed (e.g. Henry & Fetters, 2012), in others reflexivity is highlighted, where video can "make these practices amenable to more detailed scrutiny than in more traditional forms of ethnographic observation" (Griefenhagen, 2008:7). Furthermore, while elicitation

techniques can stimulate memories, video can also be used to create new ideas or sensemaking, while simultaneously emphasizing participation.

For instance, Marotto, Roos & Victor (2007), in a grounded theory study of peak performance among conservatory musicians, combined video with participant observation and interviews. Moving beyond triangulation, however, the video was used as a tool during note taking, and then to promote participant reflection around minute gestures during the performance. For instance, video footage revealed "first violinists physically turning around and looking at the flute at that very moment. From that moment on, the magic somehow was broken – the rest of the overture did not exceed anything we had played in rehearsals." (Marotto et al, 2007: 399). This moment was then presented to the violinists, who amplified the point that their concentration had been broken, that they were trying to communicate with the flautist and that the unity of the performance had suffered from this distraction.

The authors establish their contribution by using video beyond recording and triangulation, to promote participant reflexivity. They focus on "collective virtuosity", an original concept that integrates verbal and non-verbal displays to understand the tacit aspects of automatized behaviours. The authors reinforce participant reflexivity, which in turn feeds back into the quality of the interviews. As Oliver & Roos (2007) describe in an earlier article, when aided by multimodal stimuli, 'participants move beyond their discursive consciousness to a more practical consciousness, engaging in reflexive monitoring more difficult to verbalize' (p 354).

Alongside this object-mediated inquiry, a further useful example from the literature is the work of Mitchell and de Lange (2011), representative of what may be termed 'participatory video research'. The research objective was to use video to build reflexivity in a rural community around shared narratives of sensitive social topics such as HIV. The analysis was centred on a South African community, where participants took part in video

making under the researchers' guidance and in post-screening discussions. Video production was conceived as a group process. After a one-day video making workshop, small groups of community members recorded and viewed each other's videos, and were then asked to reflect on their work using open-ended questions. The technique stimulated "collective reflexivity on issues about which people are silent" (p. 179). Differently than in Marotto et al (2007), participants were involved in both production and reflection on videos, but the central aspect being the use of video to alternate between the gaze of the researcher and the participant to promote reflexivity, thereby achieving the de-centering of the theoretical gaze that is key to this third approach.

The Mitchell and de Lange (2011) study reveals how videos may be used to stimulate talk about sensitive issues – in this case poverty, disease, child abuse and crime – by means of reflection on the recorded image. Video recordings allow community members to engage in a form of collective reflexivity, where 'unspeakable' issues are articulated by means of communitarian awareness. In its pertaining to knowledge whose articulation is taboo or sensitive, this study touched upon a form of 'elusivity' not mentioned above, and provided an interesting variation of the concept of articulation. In both cases, reflection by the participants sheds light on how video can act as a 'catalyst or trigger in post-screening discussions' (p.4), making participants themselves reflexive about their knowledge. Again, video promotes reflexivity around elusive knowledge by presenting knowledge in its diverse modalities.

DISCUSSION

Beginning with the concept of elusive knowledge as an epistemological and methodological problem in organizational scholarship, the above approaches highlight video-based methods as one way of gaining traction on this problem. Based on the argument that the problem of elusive knowledges is at least in part a problem of modal limitations in knowledge, we

propose that multimodality is central to gaining a pluralistic vision of a research site. Video, with its multimodal possibilities, becomes an important support to study elusive knowledges.

Video is ideal for studying communicative interactions that present knowledge in diverse formats: spoken language, emotive expressions, interpersonal positioning, temporal sequences and the like (Silvester et al., 2000). Meaning-making in such interactions is distributed across different modes (Price, Jewitt & Brown, 2013). Understanding the diversity of modes and their relationships is pivotal to understanding a given research context. Working across these diverse modalities led us to examine video as a technical device, illustrating three distinct and progressively complex uses of video-based methods that allow exploration of these relationships.

Using video as raw data treats video as semiotically rich, multi-layered, and intrinsically cross-modal, featuring images, sounds and temporal continuity. By recording participant interaction over time, videos capture temporally-extended, unfolding actions, while providing a crucible for cross-modal translations of meaning. This method may be extended by seeking out and comparing alternate modalities of an event or phenomenon, complementing video data with participant observation, organizational archives or historical data. The third approach directly involves participants in the research process, using video as an artefact for reflection. Organizational members' internalised knowledge is made explicit by eliciting commentaries on relived experiences and attributions of meaning.

We contextualized these three approaches by a.) Briefly describing their current use within the organizational literature b.) Illustrating their mechanics within the empirical illustration of an ongoing video study of volleyball professionals and c.) Highlighting showcase studies where each method led to a conceptual contribution to the literature. The three sets of techniques should not be taken as independent approaches, but as complementary methods based on different forms of multimodal exploration.

Our study thus offers a methodological contribution to a range of organizational concerns around what we term 'elusive knowledges' by focusing on the methodological issues raised by multimodality, a concept increasingly surfacing in micro studies of social interaction (e.g. Pink, 2011; LeBaron, 2005). Previous studies provide an excellent overview of video production techniques from the perspective of multimodality (e.g. Luff & Heath, 2012; Heath, Hindmarsh & Luff, 2010), outlining the ability of video to take multiple perspectives. By theorizing multimodality from the perspective of multiple affordances, we address the articulation of elusive knowledges. Drawing together tacit, aesthetic and embodied knowledges as avenues for multimodel research, we engage in a "consensus-creating" (LeBreton, 2014; Hollenbeck, 2008) exercise, showing that the common methodological dilemma of elusiveness can be arbitrated through a sensitivity to multimodality. We thus lay out an array of organizational phenomena around which video research can be practically applied, outlining the broad approaches this application can take. The three modes build upon one another in a progressively reflexive view of data collection.

In this way, we contribute to a growing interest around the use of videos in organizational research methods scholarship, but also to using video to demonstrate a deeper methodological point that is not limited to video. This is illustrated by exploring the methodological challenges and potentialities of video to generate deep accounts of organizational life (Luff and Heath, 2011; Knobaluch, 2012; Knoblauch & Schnettler, 2012). Thus, one of the key lessons of the 'visual turn' in organizational studies is that data contain material properties that can be leveraged to draw new insights. Such insights can inform areas in which materiality is a key aspect of theorizing human activity, from sociomateriality (e.g. Jarzabkowski & Pinch, 2013) to human-technology communication (e.g. Suchman, 2007)

In highlighting different ways of using videos, we make an additional contribution by linking the concept of elusive knowledges with methodological issues of data collection. By

framing articulability as a general problem, we sought to create a common vocabulary dealing with the issue of elusive knowledges. As argued above, numerous studies underscore the difficulty of articulating embodied knowledge, yet few studies thematise this aspect of knowledge as a central methodological problem (Hakanson, 2007; Taylor, 2002). The current paper thus contributes to laying a basis for empirical research around elusive knowledges.

Nevertheless, our attempts to approach elusive knowledges must be acknowledged as approximative. The difficulty of articulation into language can result from a lack of modal diversity, but even such diversity, given current academic publishing norms, remains tied to the written word. Notably, attempts to not only gather and analyse but also to *present* results via video have gained some traction and may prove to be promising. For instance, some scholars have included CDs with their work (e.g. Jones & LeBaron, 2002), and outlets such as the *Academy of Management Discoveries* have themselves explored the use of media tools. The possibilities for multimedia academic production raise the issue of the epistemology of discursive modes and the codification of academic knowledge, issues that are beyond our scope here.

Furthermore, in dealing primarily with video data as an 'artefact' or tool to promote reflexivity, we leave for future research those issues involving constructing video images themselves. Indeed, all video images, through their production choices, selectively include, exclude and frame aspects of environments (MacDougall, 2011). Heath et al (2010), for instance, describe how different camera positions and angles allow different data configurations. In short, video images are not epistemically 'objective', but using video to take advantage of cross-modal properties can promote reflexivity around elusive knowledges. Such reflexivity operates at three progressively complex levels: that of the inner relations between modalities themselves within a video, that of the comparison between video and other data and the subjective processes of research participants as they reflect on videos.

Likewise, we can imagine scenarios where reflexivity is promoted not through viewing but through producing videos (Lomax, 1998; Pink, 2001). Such uses might parallel, for example, Warren's (2002) technique of eliciting participant-created images to stimulate reflection. Producer-driven techniques would use the construction of videos to show how selective perception goes into the creation of representations. Because our approach focuses on how viewers (scholars or participants) struggle to make sense of representations based on their cross-modal aspects, such producer-driven approaches are beyond our current scope, although they would suitably complement the current discussion.

In keeping with the focus on data collection as opposed to production, we treat videos as 'simple' recording devices and ignore the vast array of montage elements available in video-editing, such as perspective-taking, temporal manipulation and cutting, backlighting and angles, ellipses, and manipulation of focus, among others. A wide array of cinematic and filmic production techniques are known about and have been discussed elsewhere (Henley, 1998), and this is not the focus here. It is the case that researchers have a wide array of techniques for selecting and framing video data beyond treating the camera as an 'amateur home video'. Each of these techniques changes the modal and semiotic characteristics of the resulting data, thus shifting the ways in which each of the three above approaches might unfold in practice. However, the three approaches are sufficiently general that they could be applied across particular techniques, techniques which could be chosen based on the given research focus. As such, the diverse cinematic practices available for producing video data are not opposed to, but subsumed within, the current wide body of approaches.

Further, both including participants in the research process, and stressing reflexivity through textuality, are laden with power implications that merit further exploration. For instance, as compared to raw data or triangulation approaches, using video as reflective artefact involves participants in the interpretive process, giving them authorship in the

theoretical process. This may be particularly important when dealing with elusive knowledges, when researchers might interpret gestures, tool use or other subtle cues in ways quite different from the participants. In such cases, researcher-participant dialogue may help avoid imposing meanings alien to those of the actors. Yet, for the same reason, bringing participants "outside" of actions to adopt the role of "theorists" may bring its own dangers, where translating actions into text may alter the ways participants view their own actions. While promoting participant reflexivity may be an important 'performative' aspect of research (Spicer, Alvesson & Karreman, 2009), it also risks imposing an "unsolicited reflexivity" (Tobin & Davidson, 1990), wherein people's lived experiences seem to acquire legitimation through textualization. As cited in Clifford (1986) above, the "death in life" of textualization carries ethical implications that merit further exploration.

CONCLUSION

Conceptually, we have discussed the intersections of tacit, aesthetic and embodied knoweldges in their 'elusive' aspect, focusing on articulability as a common theme rather than the differences between each literature. Such treatment should not be regarded as equating the approaches, but highlighting a particular problematic. Additionally, the methodological focus on articulating forms of knowledge that are situated, material and embodied implies strong parallels with practice-based theories (e.g. Seidl & Whittington, 2014; Nicolini, 2012; Orlikowski, 2010; Nicolini et al, 2003). The relation of articulation with the situatedness of theory-in-practice raises important questions about how, and under what conditions, situated knowledges emerge as 'elusive', and when, by contrast, they appear more amenable to reflexive contemplation. Our focus on methods tackles only one side of this question, that of articulation in theory, but largely takes for granted the initial status of "elusiveness". We thus achieve a merely partial inroad into the problem of elusive

knowledges by posing it as a methodological obstacle, presenting the concept only in its general contours and leaving it to be fleshed out in further research.

In sum, we have focused on one small slice of the research endeavour, that of exploring elusive knowledges via video. The emergent nature of this mode of research means that much work remains to be done; even within this limited scope, myriad foundational and practical questions remain. Such questions are best addressed through regarding video in terms of its capacity to bring together modes of experience, thus linking video to the emerging organizational study of multimodality. Bringing in issues of both content and medium, image and movement, video methods provide an inroad into foundational questions that are increasingly occupying organizational researchers.

REFERENCES

Ambrosini, V., Bowman, C.(2001). Tacit knowledge: Some suggestions for operationalization. *Journal of Management Studies,* 38(6), 811–829.

Barthes, R.(1982). The photographic message. In S. Sontag (Ed.) A Barthes Reader. New York: Hill & Wang.

Berman, S., Down, J., & Hill, C.(2002). Tacit knowledge as a source of competitive advantage in the National Basketball Association. *Academy of Management Journal*, *45*(1), 13–32.

Bezemer, J., Jewitt, C.(2010). Multimodal analysis: key issues. *Research Methods in Linguistics* 180-197. London: Continuum.

Carlile, P.R.(2002). A pragmatic view of knowledge and boundaries: Boundary objects in new product development. *Organization Science,* 13(4), 442-455.

Clarke J., Cornelissen J.(2011) Language, communication, and socially situated cognition in entrepreneurship. *Academy of Management Review,* 36(4), 776–778.

Clifford, J. (1986) On ethnographic allegory. In J.Clifford & G.Markus (Eds.) *Writing Culture: The poetics and politics of ethnography*, 98-121. Berkeley: University of California Press.

Collins, H.(2007). Bicycling on the moon: collective tacit knowledge and somatic-limit tacit knowledge. *Organization Studies*, *28*(2), 257–262.

Cook, S.D., & Yanow, D.(1993). Culture and organizational learning. *Journal of Management Inquiry,* 2(4), 373-390.

Cunliffe, A., Coupland, C.(2012). From hero to villain to hero: Making experience sensible through embodied narrative sensemaking. *Human Relations,* 65(1), 63–88.

Dant, T.(2004). Recording the habitus. In Pole (Ed.) *Seeing is believing? Approaches to visual research*. New York: Elsevier.

Demir, R.(2015). Strategic activity as bundled affordances. *British Journal of Management*, *26*, S125–S141.

Dempsey, N.P.(2010). Stimulated recall interviews in ethnography. *Qualitative Sociology*, *33*(3), 349–367.

Donaldson, L.(2001). Reflections on knowledge and knowledge-intensive firms. *Human Relations*, 54(7), 955-63.

Erickson, F.(2011). Uses of video in social research: a brief history. *International Journal of Social Research Methodology*, 14(3), 179–189.

Endrissat, N., Islam, G., & Noppeney, C.(2016). Visual organizing: Balancing coordination and creative freedom via mood boards. *Journal of Business Research*, 69(7), 2353–2362.

Ewenstein, B., Whyte J.(2007a). Beyond words: Aesthetic Knowledge and knowing in organizations. *Organization Studies,* 28(5), 689-708.

Ewenstein, B., and Whyte J.(2007b). Visual representations as 'artefacts of knowing'. *Building Research & Information,* 35(1), 81–89.

Fele, G.(2012). The use of video to document tacit participation in an emergency operations centre. *Qualitative Research,* 12(3) 280–303.

Fine, G.A.(1992). The culture of production: Aesthetic choices and constraints in culinary work. *American Journal of Sociology,* 97(5), 1268–1294.

Gagliardi, P.(1996). Exploring the aesthetic side of organizational life. In Clegg, S.C. Hardy, N., Nord, W., (Eds), *Handbook of Organizational Studies*, 565-580. London: Sage

Gibson, J.J.(1986). *The ecological approach to visual perception.* Hillsdale, NJ: Erlbaum.

Greiffenhagen, C.(2008). Video analysis of mathematical practice? Different attempts to 'open up' mathematics for sociological investigation. FQS 9(3) http://www.qualitative-research.net/index.php/fqs/article/view/1172/2585#g3

Goodwin, C.(1994). Professional vision. *American Anthropologist,* 96, 606-633.

Gylfe, P., Franck, H., LeBaron, C., and Mantere, S.(2016). Video methods in strategy research: Focusing on embodied cognition. *Strategic Management Journal*, 37(1), 133–148.

Hakanson, L.(2007). Creating knowledge: The power and logic of articulation. *Industrial and Corporate Change,* 16(1), 51-88.

Heath, C., Hindmarsh, J.(2002). Analysing interaction. In T. May(Ed) *Qualitative Research in Practice*, 99 – 121. London: Sage.

Heath, C., Hindmarsh, J., Luff, P.(2010). *Video in qualitative research. Analysing social interaction in everyday life.* London: Sage.

Henley P.(1998). Film-Making and Ethnographic Research. In J. Prosser (Ed.) *Image-based Research: A Sourcebook for Qualitative Researchers,* 42-59. London: Routledge.

Henry, S. & Fetters, M.(2012). Video elicitation interviews: A qualitative research method for investigating physician-patient interactions. *Annals of Family Medicine,* 10(2), 118-125.

Hindmarsh, J., & Pilnick, A.(2007). Knowing bodies at work: Embodiment and ephemeral teamwork in anesthesia. *Organization Studies*, *28*(9), 1395–1416.

Hollenbeck, J.R.(2008). The role of editing in knowledge development: Consensus shifting and consensus creation. In Y.Baruch, A.M. Konrad, H.Aguinis, & W.H.Starbuck (Eds.), *Opening the black box of editorship*, 16-26. New York: Palgrave- Macmillan.

Hutchins E.(1995). *Cognition in the Wild*. Cambridge, MA: MIT Press.

Islam, G.(2015a). Practitioners as theorists: Para-ethnography and the collaborative study of contemporary organizations *Organization Research Methods,* 18 (2), 231-251.

Islam, G.(2015b). Extending organizational cognition: A conceptual exploration of mental extension in organizations. *Human Relations,* 68(3), 463-487.

Islam, G. & Benameur, K. (2016). Eyes on the ball: Tacit knowledge and improvisation in a professional volleyball team. Working paper.

Islam, G., Endrissat, N. & Noppeney, C.(forthcoming) Beyond the eye of the beholder: Scent innovation through analogical reconfiguration. *Organization Studies.*

Jarrett, M., & Lui, F.(forthcoming). "Zooming with": A Participatory Approach to the Use of Video Ethnography in Organizational Studies. *Organizational Research Methods*.

Jarzabkowski, P., Burke, G., & Spee, P.(2015). Constructing spaces for strategic work: A multimodal perspective. *British Journal of Management*, *26*, S26–S47.

Jarzabkowski, P., Bednarek, R., & Le, J.K.(2014). Producing persuasive findings: Demystifying ethnographic textwork in strategy and organization research. *Strategic Organization*, *12*(4), 274–287.

Jarzabkowski, P., Pinch, T.(2013). Sociomateriality is 'the New Black': Accomplishing repurposing, reinscripting and repairing in context. *M@na@gement*, 16(5), 579-592.

Jewitt, C.(2013). Multimodal methods for researching digital technologies. In S.Price., C. Jewitt., & Brown, B. (Eds.) *The Sage Handbook of Digital Technology Research*. London: SAGE

Jick, T.D.(1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly,* 602-611.

Jones, S., LeBaron, C.(2002). Research on the relationship between verbal and nonverbal communication: Emerging integrations. *Journal of Communication,* 52(3), 499-521.

Knoblauch, H. (2012). Introduction to the special issue of Qualitative Research: Video-analysis and videography. *Qualitative Research*, 12(3), 251–254.

Knoblauch, C.H., Tuma, R.(2011). *Videography: An Interpretative Approach to Video-Recorded Micro-Social Interaction*. The SAGE Handbook Visual Research Methods, London: SAGE.

Knoblauch, H., Schnettler, B.(2012). Videography: Analysing video data as a "focused" ethnographic and hermeneutical exercise. *Qualitative Research,* 12(3), 334–356.

Koschmann, T., LeBaron, C.(2002). Learner Articulation as Interactional Achievement: Studying the Conversation of Gesture. *Cognition and Instruction*, 20(2), 249-282.

Kupers, W.(2013). A phenomenology of embodied senses: The 'making' of sense in organisational culture. *International Journal of Work, Organizations and Emotions*, 5(4), 324-341.

Ladkin, D., & Taylor, S.S.(2010). Enacting the "true self": Towards a theory of embodied authentic leadership. *The Leadership Quarterly*, *21*(1), 64–74.

Lakoff, G., & Johnson, M.(1980). Metaphors We Live By. Chicago: The University of Chicago Press.

Lam, A.(2000). Tacit Knowledge, Organizational Learning and Societal Institutions: An Integrated Framework. *Organization Studies, 21*(3), 487-513.

LeBaron, C.(2005). Considering the social and material surround: Toward micro-ethnographic understandings of nonverbal behavior. In V. Manusov (Ed.), *The Sourcebook of Nonverbal Measures,* 493-506. Mahwah, NJ: Erlbaum.

LeBreton, J.(2014). Editorial. *Organizational Research Methods*,17(2), 113-117.

Lok, J., & deRond, M.(2013). On the plasticity of institutions: Containing and restoring practice breakdowns at the Cambridge University Boat Club. *Academy of Management Journal*, *56*(1), 185–207.

Lomax, H. Casey, N.(1998). Recording social Life: Reflexivity and video methodology. *Sociological Research Online*, 3(2), 1-30.

Luff, P., Heath, C.(2011). Some 'technical challenges' of video analysis: Social actions, objects, material realities and the problems of perspective. *Qualitative Research,* 12(3), 255-279.

McCanles, M. (1982). The dialectical structure of discourse. Poetics Today, 3 (4), 21–37.

McDonald, S., & Simpson, B.(2014). Shadowing research in organizations: The methodological debates. *Qualitative Research in Organizations and Management: An International Journal*, *9*(1), 3–20.

McDermott, R., Gospodinoff, K., & Aron, J.(1978). Criteria for an ethnographically adequate description of concerted activities and their contexts. *Semiotica,* 24(3/4), 245-276.

MacDougall, D.(2011). Anthropological filmmaking: An empirical art. In E.Margolis & Pauwels, L.(Eds.) *SAGE Handbook Visual Research Methods*,99-113. London: Sage

Melcher, J.M., Schooler, J.W.(1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language,* 35, 231-245.

Marotto, M., Roos, J., & Victor, B.(2007). Collective virtuosity in organizations: A study of peak performance in an orchestra. *Journal of Management Studies*, *44*(3), 388–413.

Martin, P.Y.(2002). Sensations, bodies, and the "spirit of a place": Aesthetics in residential organizations for the elderly. *Human Relations,* 55(7), 861-885.

Meyer, R.E., Hollerer, M.A., Jancsary, D., van Leeuwen, T.(2013). The visual dimension in organizing, organization and organization research: Core ideas, current developments, and promising avenues. *The Academy of Management Annals*, 7(1), 489-555.

Mitchell, R., Mitchell, J.R.(2011). Socially situated cognition: Imagining new opportunities for entrepreneurship research. *The Academy of Management Review* 36(4), 774–776.

Mitchell, C., deLange, N.(2011). Community-based participatory video and social action in rural South Africa. In E.Margolis and L.Pauwels (Eds) *The SAGE Handbook Visual Research Methods* 171-185. London: Sage

Mondada, L.(2012). Video analysis and the temporality of inscriptions within social interaction: the case of architects at work. *Qualitative Research,* 12(3), 304–333.

Mondada, L.(2006). Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In B.Knoblauch, H.Raab, J.Soeffner, H.Schnettler, (Eds.) *Video Analysis,* 1–18. Bern: Lang.

Nelson, R. & Winter, S.(1982). *An evolutionary theory of economic change*. Harvard University Press: Cambridge, MA.

Nicolini, D.(2012) *Practice theory, work, and organization: An introduction.* Oxford: OUP.

Nicolini, D., Gherardi, S. & Yanow, D.(2003). *Knowing in organizations. A practice-based approach*. London: Routledge.

Nonaka, I., & Krogh, G.(2009). Tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. *Organization Science*, *20*(3), 635–652.

Oliver, D., Roos, J.(2007). Beyond text: Constructing organizational identity multimodally. *British Journal of Management*, 18, 342–358.

Orlikowski, W.J. (2010).Engaging practice in research: Phenomenon, perspective, and philosophy. in D.Golsorkhi, L.Rouleau, D.Seidl, and E.Vaara (Eds.) *The Cambridge Handbook on Strategy as Practice, 23-33.* Cambridge: Cambridge University Press.

Paroutis, S., Franco, L.A.and Papadopoulos, T.(2015). Visual interactions with strategy tools: Producing strategic knowledge in workshops. *British Journal of Management*, 26,S48–S66.

Pink, S.(2011). Multimodality, multisensoriality and ethnographic knowing: Social semiotics and the phenomenology of perception. *Qualitative Research,* 11(1), 261-276.

Pink, S.(2004). Performance, self-representation and narrative: Interviewing with video. In Pole, C.(ed.) *Seeing is believing? Approaches to visual research*, 61-77. New York: Elsevier.

Pink, S.(2001). More visualising, more methodologies: On video, reflexivity and qualitative research. *The Sociological Review,* 49(4), 586–599.

Polanyi, M.(1962). *Personal knowledge. Towards a post-critical philosophy.* London: Routledge.

Price, S., Jewitt, C., Brown, B.(2013). *The Sage Handbook of Digital Technology Research*. London: SAGE

Saussure, F.(1974). *General course in linguistics* (W. Baskin, Trans.). London: Fontana.

Seidl, D., & Whittington, R.(2014). Enlarging the strategy-as-practice research agenda: Towards taller and flatter ontologies. *Organization Studies*, *35*(10), 1407–1421.

Shamsie, J., & Mannor, M.J.(2012). Looking inside the dream team: Probing into the contributions of tacit knowledge as an organizational resource. *Organization Science*, 1-17.

Shore, B.(1996). *Culture in Mind: Cognition, Culture and the Problem of Meaning*. Oxford: Oxford University Press.

Silvester, J., Anderson, N., Haddleton, E., Cunningham S., Gibb, A.(2000). A cross-modal comparison of telephone and face-to-face selection interviews in graduate recruitment. *International Journal of Selection and Assessment,* 8, 16–21.

Smets, M., Burke, G., Jarzabkowski, P., Spree, P.(2014). Charting new territory for organizational ethnography. Insights from a team-based video ethnography. *Journal of Organizational Ethnography,* 3(1) 10-26.

Smith E.R. and Semin G. R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology 36*, 53-117.

Spicer A., Alvesson M. and Kärreman D.(2009). Critical performativity: The unfinished business of critical management studies. *Human Relations* 62(4), 537–560.

Stivers, T., and Sidnell, J.(2005) Multi-modal interaction. *Semiotica* 156 (1/4):1-20

St. Pierre, E.A., Jackson, A.Y.(2014). Qualitative data analysis after coding. *Qualitative Inquiry,* 20(6), 715-719.

Streeck, J., Goodwin, C., & LeBaron, C. (Eds.) (2011). *Embodied interaction: Language and body in the material world.* Cambridge, England: Cambridge University Press

Strati, A., De Montoux, P.G.(2002). Introduction: Organizing aesthetics. *Human Relations*, 55 (7), 755–66.

Strati, A.(1992). Aesthetic understanding of organizational life. *Academy of Management Review,* 17, 568-581.

Streeck, J., & Mehus, S.(2005). Microethnography: The study of practices. In Fitch, K.L., & Sanders, R.E. (Eds.), *Handbook of language and social interaction,* 381-404. Psychology Press.

Suchman, L.(2007). *Human-machine reconfigurations: Plans and situated actions*. Cambridge: Cambridge University Press.

Suchman, Lucy A., and Trigg, R.H.(1991). Understanding practice: Video as a medium for reflection and design. In J. Greenbaum and M. Kyng, Eds, *Design at Work,* 65–89. Hillsdale, NJ: Lawrence Erlbaum.

Sutton J.(2006). Introduction: Memory, embodied cognition and the extended mind. *Philosophical Psychology,* 19(3), 281–289.

Taylor, S.S.(2013). What is organizational aesthetics? *Organizational Aesthetics* (2)1, 30-32.

Taylor, S. S.(2002). Overcoming aesthetic muteness: Researching organizational members' aesthetic experience. *Human Relations,* 55(7) 821-840

Taylor, S. S., & Hansen, H.(2005). Finding form: Looking at the field of organizational aesthetics. *Journal of Management Studies,* 42(6), 1211-1231.

Thanem, T.(2006). Organization living on the edge: Towards a monstrous organization theory. *Organization, 13*, 163—193.

Tobin, J., & Davidson, D.(2006). The ethics of polyvocal ethnography: Empowering vs. textualizing children and teachers. *International Journal of Qualitative Studies in Education*, *3*(3), 271–283.

Tsoukas, H.(2003). Do we really understand tacit knowledge?' in Easterby-Smith and Lyles (Eds), *The Blackwell Handbook of Organizational Learning and Knowledge Management,* 411–427. Cambridge, MA: Blackwell.

Warren, S.(2008). Empirical challenges in organizational aesthetics research: Towards a sensual methodology, *Organization Studies,* 29(4), 559-580.

Warren, S.(2002). 'Show me how it feels to work here': Using photography to research organizational aesthetics. *Ephemera,* 2(3), 224-245.

Wasserman, V., & Frenkel, M.(2010). Organizational aesthetics: Caught between identity regulation and culture jamming. *Organization Science*, 22(2), 503-521.

Weick, K.E.(2004). Mundane Poetics: Searching for Wisdom in Organization Studies. *Organization Studies*, *25*(4), 653–668.

Wolfe R.A., Weick K.E., Usher J.M., Terborg J.R., Poppo L., Murrell A.J., Dukerich J.M., Core D.C., Dickson K.E., Jourdan J.S.(2005). Sport and organizational studies: Exploring synergy. *Journal of Management Inquiry* 14(2),182–210.

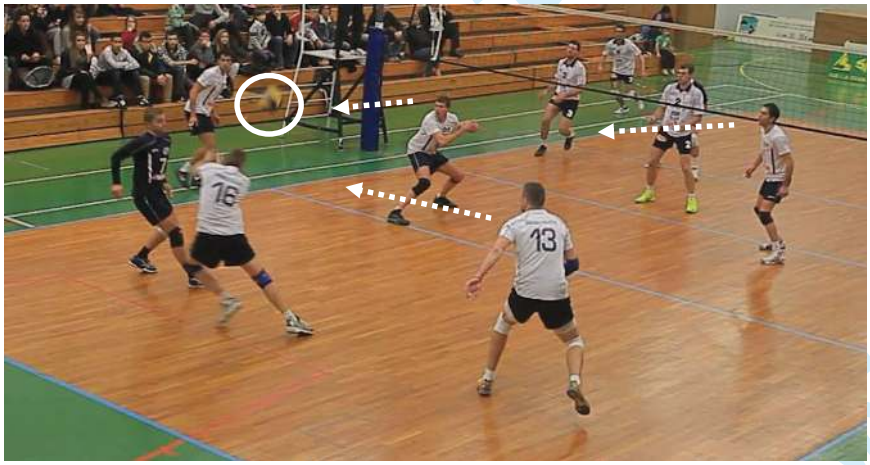
Yaniv, I., Choshen-Hillel, S., & Milyavsky, M. (2010). Receiving advice on matters of taste: Similarity, majority influence, and taste discrimination. *Organizational Behavior and Human Decision Processes*, 115(1), 111-120.

Table 1. *Alternative uses of videos in the research process*

Type of model	What it consists of	Reflexive Possibilities	Constraints	Situation in which to use it
Video as raw data	- Direct coding of video data - Content coding can be complemented by cross-modal information - Modalities compared by contents but also by differential mode of representation (e.g. sound, image)	- Inherently cross-modal as opposed to text-based techniques - Widen the semiotic resources available for interpretation - Maintain awareness of the material nature of representations	- Does not include organizational members' first-person experience as data - Limited by horizons of video 'gaze' -Video provides an 'illusion of objectivity'	- Where gestural or non-verbal elements are involved - Where knowledge is tacit and difficult to articulate - Where interviews are likely to be biased or self-censored
Video as triangulation	- Comparison across methods - Interviews, observations, field notes combined to approximate experiences. - Confirmation versus pluralistic use	- Comparative possibilities for verification and agreement - Interpretation of one source of data can be aided by the other - Reveals modal differences among data - Each mode contributes to more holistic understanding of contexts	- Which modalities to compare? - Ambiguity in cases of disagreement - Incommensurability between modalities	- Where different sources of data provide unique perspectives - Where diverse communicative resources are available and coherent
Video as reflective artefact	-Assisting participants in constructing, remembering and articulating experiences - Display and interview simultaneously - Sensemaking around embodied knowledge	- Offers participative opportunity in articulating elusive knowledges - Reduces biases of researcher-imposed interpretations. - Increase bridge between theorization and practice - Possibilities for self-reflection and critique	- Danger of post-hoc rationalization - Conflation of first-person experience and reconstruction - Third-person subject position may change nature of knowledge	- Where critical distance from practice is not routinized - Where elusive knowledges are not evident from direct cross-modal observation - Where verbal communication requires augmentation or reference

Figure 1. *Illustration of video used as Raw Data – Unfolding of Collective Action*

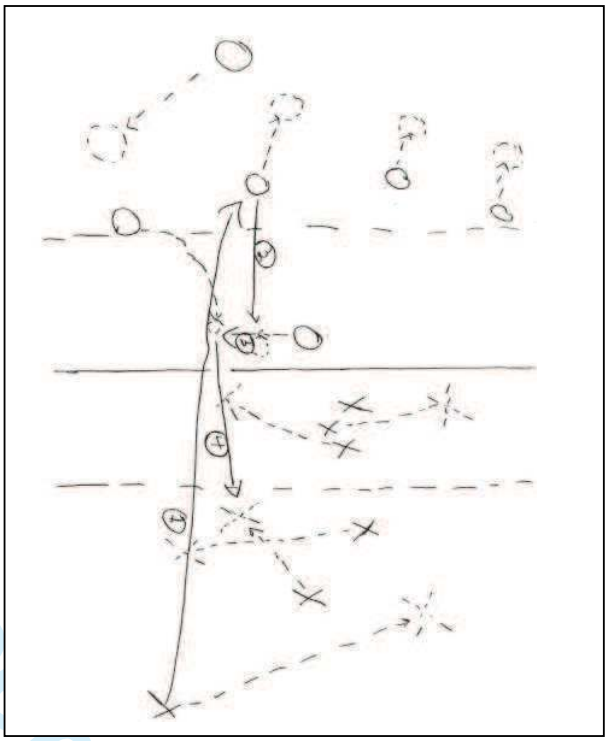
- Two players approach ball
- Number 16 calls for initiative



- Number 7 returns to position
- Teammates' gaze focuses on player/hit

- 16 moves out of way of play
- Gaze moves to ball in air (off screen)

Figure 2. *Illustration of video used as Triangulation – Stills versus Researcher Sketch of Action versus Interview Description (General)*



Server: Everyone makes the decision in the moment, but I don't say that my decision is most important, no. Maybe it's the decision who is the receiver of the volley that is my decision. For example, how we play it, if we play good or not good, maybe my decision is to change the player who attacks. But also, my decision is the decision of the reception player, who receives. And this is the one circle who comes between me, and that is the system. After that it's difficult to say if it's more important than the decision of the attacker.
Researcher: Do you see it as a circle, and you're the middle of the circle?
Server: Maybe it's like a circle – For example, my decision is harder than the decision of the receptor. More difficult - for example, he has the decision to give the ball to one player, but my decision is to give the ball. Which player? I have 4 players, which one to give the ball?

Figure 3. *Illustration of video used as a Reflective Stimulus*

Screenshot from Video Episode - Choice between two possible attacks	Player Commentary on Episode During Multiple Viewing
	Player: Yes, yes. For example, in this moment that I explained, I planning to play with the middle blocker and poster 4. But, we didn't make the point, we made the error, it was not what I wanted, after that I changed my plan, I played with another player, thinking "let's see what happens there". Interviewer 1: but you change…in the plan…because you cannot say to the players "we will change"? Player: No, no, no, no! No, me alone I make the decision. For example what I played, what I said for example. In this position, I am planning to play poster 4, poster 3. In this point we make the point, but imagine if we had made a fault, we had lost the point, and after that I have a second solution but I don't say… Interviewer 1: …For another situation… Player: … I don't tell my solution to anybody, but I have it. For example, if we stay in the same position five times, I must find some solution with another player, something else. But also, also the situation in the moment that I make it, you know the ball is very difficult to set perfectly. And you want to play with the poster 4, but it's impossible - it's very difficult to set the ball in poster 4. I make my decision for this ball to set, to come good on the position where I set. It's not that I don't see which player I have, what I am planning, but to make the solution out of the difficult situation with what I have it in this moment.

Study	Description of Cases	Research Problem	Data Sources	Approach it approximates	Lessons learned
Mondada (2012)	Architectural office	Social interaction among participants	Video recording and video transcription	Video as raw data	Multimodal data available from videos aid researchers to understand architects' interaction and how their ideas turn into a 'solution'
Fele (2012)	Operations centre handling emergency calls in Italy	Coordination and collaboration emerging from the interaction between workers	Video recording and telephone call recording	Video as raw data	Gestural and non-verbal recording permit an understanding of tacit exchanges between operators working on different tasks
Paroutis, Franco & Papadopoulous (2015)	Management team during a strategy workshop	How strategy tools are created during workshops	Video, Researcher diary, Research Assistant observations	Video as triangulation	Shift, inertia and assembly describe forms of visual interaction during tool creation
Smets, Burke, Jarzabkowski and Spree (2014)	Reinsurance trading in London	New forms of organizational ethnography.	Team-based video ethnography	Video as triangulation	Multimodal data available from videos can complement and support traditional fieldwork. They can confirm what emerge from observations and provide 'rigorous' insights.
Marotto, Roos & Victor (2007)	Musical conservatory orchestra	How collective talent relates to collective performance	Ethnography, video, written journal	Video as reflective artefact	Individual virtuosity can contribute to collective peak performance through reflexivity and shared emotional experience
Mitchell and de Lange (2011)	1 Rural southern African community	Stimulate community members to talks about 'sensible' issues	Video analysis, follow-up viewings, small group discussions, and community screening	Video as reflective artefact	Videos can be used by participants themselves and can function as a catalyst or trigger in post-screening discussions

Table 2. *Recent examples of cases from the literature*