Visual research in clinical education
Jeff Bezemer, UCL Institute of Education

Abstract

Aim
The aim of this paper is to explore what might be gained from collecting and analysing visual data, such as photographs, scans, drawings, video and screen recordings, in clinical educational research. Its focus is on visual research that looks at teaching and learning ‘as it naturally occurs’ in the workplace, in simulation centres and other sites, and also involves the collection and analysis of visual learning materials circulating in these sites.

Background
With the ubiquity of digital recording devices, video data and visual learning materials are now relatively cheap to collect. Compared to other domains of education research visual materials are not widely used in clinical education research. The paper sets out to identify and reflect on the possibilities for visual research using examples from an ethnographic study on surgical and interprofessional learning in the operating theatres of a London hospital.

Main contribution
The paper shows how visual research enables recognition, analysis and critical evaluation of (1) the hidden curriculum, such as the meanings implied by embodied, visible actions of clinicians; (2) the ways in which clinical teachers
design multimodal learning environments using a range of modes of communication available to them, combining, for instance, gesture and speech; (3) the informal assessment of clinical skills, and the intricate relation between trainee performance and supervisor feedback; (4) the potentialities and limitations of different visual learning materials, such as textbooks and videos, for representing medical knowledge.

Discussion and conclusion
The paper concludes with theoretical and methodological reflections on what can be made visible, and therefore available for analysis, explanation and evaluation if visual materials are used for clinical education research, and what remains unaccounted for if written language remains the dominant mode in the research cycle. Opportunities for quantitative analysis and ethical implications are also discussed.

Introduction
Visual research describes a rapidly expanding field focused on the study of visual phenomena and the use of visual data. Cutting across the arts and humanities and social sciences, it is theoretically and methodologically highly diverse, with education being one of its many domains of application. In this paper I explore the possibilities, significance and relevance for clinical education of a growing body of visual research that sits in the tradition of naturalistic inquiry. It looks at teaching and learning ‘as it naturally occurs’ in the work place, in simulation
centres and other clinical educational settings through a visual lens, and also involves the collection and analysis of visual materials circulating in these sites. Thus the approach draws on visual data—photographs, scans, drawings, video and screen recordings—of two kinds: visual recordings made by researchers to document what happens in clinical learning environments; and materials produced by health professionals for use in those learning environments, and collected by the researcher. In practice, the former refers largely to video recordings of learning activity—e.g. of a lecture, bedside teaching, immersive simulation; and the latter to visual records of resources, tools, technologies, documents, and objects designed and used for learning, such as a textbook, instructional video, or simulator.

With the ubiquity of digital recording devices, video data are now relatively cheap to create and collect by researchers. For instance, in modern operating theatres, often three different cameras are available: the laparoscope, light handle cameras and wall mounted cameras. They show, from different angles, both intra- and inter-professional learning and teaching activity. These cameras are frequently used to give demonstrations to remote audiences, or to consult colleagues based elsewhere. Yet, they are seldom used for educational research purposes. This is typical of clinical educational research more widely. According to Rees, writing in 2010, only a “handful” of articles in Medical Education referred to video as data. Since then, only a few more video-based studies have been published in this journal, with each adopting distinctly different methodological approaches. This stands in sharp contrast to mainstream, non-
clinical education research, where visual data have been widely used from the moment the technologies became affordable.\textsuperscript{6}

With the ubiquity of digital recording devices comes the widespread development of visual learning materials and its dissemination on the web. For instance, YouTube hosts thousands of ‘how-to videos’ for virtually every clinical procedural skill identified in the undergraduate medical curriculum for the UK and other countries (see, e.g., geekymedics.com), including videos produced by students and/or educators and those by ‘professional’ producers and/or ‘professional’, expensive equipment, involving complex forms of animation, for instance. These videos, which exist alongside, and potentially replace still image materials used in, e.g. anatomy atlases, attract hundreds of thousands of views every year, yet they have not yet attracted much research attention. Again, this stands in sharp contrast to educational research outside the clinical domain, where the changing role of image in learning environments has been well researched.\textsuperscript{7}

So what might be gained from collecting and analysing visual data? We might rephrase this question and ask, if our aim is to describe and understand what clinical learning environments actually look like, i.e. what really happens in day-to-day education practice, thus complementing research aimed at establishing ‘what works’, or ought to happen, then what would remain unaccounted for if we ignored the visual? In addressing that question, I point to four interrelated areas for clinical educational research: the ‘hidden’ curriculum, teaching skills, assessment of clinical skills, and visual learning materials.
I will discuss and illustrate what visual research draws attention to in each of the areas with examples from ethnographic research on surgical and inter-professional learning in the operating theatres of a London hospital. The examples were selected from a corpus of audio and video recordings of 74 hours capturing 44 operations recorded in the period between 2010 and 2013, and from learning materials collected in the course of the project. All staff and patients involved consented to taking part and NHS research ethics approval was obtained prior to data collection. A detailed methodological account of the project can be found elsewhere.8

All data extracts were subjected to detailed semiotic analysis. Unlike more general approaches to qualitative research such as thematic content analysis, semiotic analysis attends to the fine grain detail provided by visual records. Such details are often overlooked the first time a piece of data is looked at. For instance, a semiotic analysis of learning materials might involve attention to the pictorial detail and degree of colour saturation used in a diagram. The semiotic analysis of video materials typically proceeds through detailed transcription. Transcription is the systematic registration, documentation, and re-presentation of video recordings: speech is written out, gesture is depicted in image, movement is graphically represented, et cetera. Software packages such as Transana and ELAN support the detailed transcription and annotation of video data. Transcription draws attention to details that are difficult to spot in real time and record in written field notes, such as a pointing gesture or
momentary shift in the direction of gaze. Detailed guidance on these analytical procedures is given elsewhere.\textsuperscript{9}

**The ‘hidden’ curriculum**

The notion of the ‘hidden curriculum’ suggests that much of what is taught and learned is not made explicit in official descriptions of the knowledge and ‘competencies’ that ought to be taught and learned, and that they are not easily recognizable, noticeable, observable and describable. A detailed visual analysis is one essential means of making components of the hidden curriculum explicit, informing discussion about what ought to be included in the official curriculum; and how learning beyond the official curriculum might be facilitated.

Take the following situation.\textsuperscript{10} A surgeon is tying knots. Meanwhile, one assistant turns to the scrub nurse, asking for scissors. Moments later, she receives the scissors, and cuts the surgeon’s tie. Without a word being said, the assistant had understood that soon the tie would need to be cut, and she made the necessary preparations to achieve that task. She then made the cut in response not to a verbal request from the surgeon, but in response to the surgeon stretching the tie tightly between his two hands and raising his arms slightly in the direction of the assistant, with the mid-point of the part of the tie that is stretched out marking the point where he wanted the assistant to cut. All of this happens entirely seamlessly and, seemingly, effortlessly. Yet underpinning this seamlessness is a significant body of ‘practical’ knowledge, enabling the assistant
to interpret the surgeon’s actions and plan her own supporting actions accordingly (‘he’s already done two throws, he’ll want me to cut the tie in a moment – so I need to get hold of a pair of stitch scissors right away’). None of this knowledge is made explicit in ‘revision notes’ for junior doctors on a surgical rotation, let alone in descriptions of the official surgical curriculum. Yet it is an essential part of becoming a surgeon; and none of it can be noticed and described unless we have a detailed visual record of trainees participating in actual operations.

Put differently, visual research draws attention to the situated meanings of ‘embodied’ action. In one study I looked at gall bladder removal operations and found that operating surgeons make a distinctive movement with their instrument just before they clip the cystic duct, sweeping up and down behind what they take to be the cystic duct and squeezing it with a grasper. These movements can be read by other surgeons as gestures; as signs – indeed explications- of the operating surgeon’s clinical judgement at an important point in this procedure, as if saying, ‘Look, I’ve mobilised this structure, it’s completely freed, it’s definitely the cystic duct and not the common bile duct, and I’ve checked that there are no stones in it, so I can now safely proceed to clip it; if you disagree, speak up now.’ We found that at our research site both consultants and surgical trainees made these gestures, without ever commenting on it: they were ‘passed on’ without any explicit teaching.

Examples such as these suggest that as trainees participate in clinical activity they recognize and interpret the actions of others. All members of a team, both
experienced and novices do so, yet attach differing meanings to them, depending on their professional role and individual experiences. For instance, like surgical trainees, nurses learn to read the surgeon’s actions as signs of upcoming requests,\(^\text{13}\) and at the same time look for visual signs of the ‘mood’ they are in.\(^\text{14}\) Thus, clinical work can be seen as an unceasing process of interpreting and improving the capacity to ‘read’ the silent, visible bodily conduct of others.\(^\text{15}\) As with language (e.g. names used to refer to surgical instruments) some of the actions recognized and given meaning are typical of one individual only; others are shared locally, nationally or internationally. They are rarely made explicit in instructional materials: they are learned by watching, observing, noticing, and engaging, and they can be made explicit (only) through detailed analysis of visual data.

**Teaching skills**

Some teaching can be recognized by attending to what educators say to their trainees. For instance, one teaching practice that is more or less straightforwardly identifiable as ‘teaching’ is the ‘quizzing’ of trainees during case presentations, on the ward, during ‘bedside’ teaching, or in the operating theatre.\(^\text{16}\) This may be taken as an indication of the importance of the language skills of the educator (and trainee). Yet attention to the visual shows that clinical teaching involves the skilful use of a range of different means of communication, including but not exclusively speech.
As a simple example, take Figure 1. It shows some medical students standing around an operating table, looking at the screen to which the laparoscopic camera is connected. Figure 2 is a snapshot of the view of the camera projected on the screen. It shows the operating surgeon pushing something to one side with an instrument, at the same time as he says ‘That’s the liver’.

[Insert Figures 1 and 2 about here. Caption Figure 1: Medical students observing a key hole operation; Caption Figure 2: Consultant surgeon pointing at his patient’s liver].

In this example, the surgeon uses gesture to point at, and speech to name, an anatomical entity. He needs gesture as much as speech to draw the medical students’ attention to an organ in this patient, and to identify it as the liver. Without either ‘modes’ of gesture or speech, this brief teaching episode would have been rather different. Naming alone would leave the students with the job of establishing exactly what, in this ‘mess’ of ‘stuff’ inside the patient’s abdomen, counts as the liver – a medical category of which they have seen abstracted representations in textbooks that look different to what they see now on the screen. Pointing alone would leave the students with the job of establishing what the object is that is marked out by the pointing: that is, what anatomical category it is an instance of. The students, for their part, do not speak at all; they engage with the object at issue through gaze and through gaze communicate to the operating surgeon that they are attending to what he is highlighting. All are positioned so as to have sight of the screen, looking at the liver.
We might say that the consultant uses a range of modes to design a multimodal learning environment. His pointing gesture and naming suggest that he has designed an environment for the students to engage with, in which the surgeon performs, and the students watch. This social/pedagogic relation is realised in other modes, too, such as dress: the surgeon wears a gown, allowing him to enter the sterile zone around the patient, while the students do not; they only wear scrubs. The relation is evident also in room layout and in the positioning of participants: the surgeon is positioned at the centre of the operating theatre and stands centrally at the operating table; the students are placed somewhat more peripherally to the operating table. And there is a differential use of communication resources: the surgeon speaking and pointing, and the students remaining silent and holding their hands behind their backs.

The multimodal tool box of the clinical educator enables them to respond to complex and differing needs from different groups of trainees. For instance, I observed one surgeon who asked medical students in the room about the complications of stomas, while at the same time holding up a piece of tissue wrapped around his finger, indicating to the assistant when and where to apply the diathermy hook. Thus different learning environments were created at the same time: one for the medical students, which was audible, highlighting 'propositional' knowledge; and one with the assistant, which was visible only, highlighting 'procedural' knowledge. Had I attended to the spoken only, I would have been left to believe that the surgeon was dealing with one type of learner only, while failing to notice the entirely common, taken-for-granted practice of quietly offering instructions; for instance, by manually guiding the hands of the
trainee, or by demonstrating gesturally, mimicking the actions involved in, say, administering an injection.

**Assessment of clinical skills**

Visual analysis provides insight in the informal assessment of clinical skills. Take the example represented in Figure 3. It details the spoken interaction between a supervising consultant surgeon and an operating trainee and makes visible how the trainee applies the scalpel as he is dissecting connecting tissue between the colon and the abdominal wall.

[Insert Figure 3 about here. Caption: Surgical trainee applying scalpel and supervising consultant giving instructions.]

In the transcript, the application of the scalpel is shown on vertical time lines. Every second is marked with a short hairline, and every 10 seconds with a long hairline. Four types of actions are rendered visible, in separate columns: 1) getting into contact with tissue; 2) lifting tissue up; 3) grasping tissue; and 4) cutting tissue. The transcript shows that every time the operating trainee makes a cut, he goes through this cycle of steps, one by one. Each step is preparing him for the next cut, and to the supervising consultant signifies increasing commitment to cut where the scalpel is in contact with —indeed, pointing at—tissue. Like the trainee interpreting the actions of the consultant, the consultant
interprets the actions of the trainee, allowing him to anticipate the trainee’s course of action and provide ‘tuned’, fitting feedback.

By systematically and graphically representing the trainee’s actions and the consultant’s spoken feedback in one transcript we have rendered visible how each action is judged and what kind of feedback it prompts: e.g. to change course (“I’d go superficial to that actually”); flag up potential danger points (“Just be a bit careful with the planes there”); confirm that he is ‘on the right track’ (“That’s the right plane”); or ratify a cut (“go for it”). The transcript also highlights some of the challenges of giving feedback, e.g. the challenge of making references to specific points in the focal space (“I kind of meant that other bit down there”). In other words, visual analysis renders the object of assessment visible, and in doing so facilitates exploration of the interaction between instructor feedback on the one hand, and trainee action on the other: i.e. how one action leads to another; how educator and trainee respond to one another, given their understanding of the situation they are in.

Assessing a trainee and giving feedback is of course also a teaching skill; and exploring how that ‘gets done’ provides insight in the ‘hidden curriculum’ - e.g. about what counts as an acceptable level of certainty to justify an invasive action. In one instance, I observed a surgical trainee doing a gall bladder removal operation making the distinctive instrument movements described above repeatedly (the experienced surgeons I observed proceeded to clip straight after having made that gesture). It wasn’t until the supervising consultant commented that ‘I think you’re fine there, I would just take these structures’ that he
proceeded to clip the cystic duct. From this we can infer that to the educator, the trainee’s actions signified ‘hesitation’ and ‘uncertainty’ about how to read the anatomy: is the area sufficiently clear and is it therefore possible to claim with certainty that a structure within it is indeed the structure that needs to be clipped? Through visual analysis these processes of interpretation can be reconstructed, and the unspoken, tacit clinical judgments and norms identified.

**Visual learning materials**

If we consider line drawings in early anatomy atlases, full colour images in contemporary ‘e-textbooks’ and the visual appearance of contemporary (virtual) simulators we can say that visual representation is, and has always been central to clinical education; that we now have more and a wider range of different technologies for producing and disseminating them; and that these technologies are now widely available.

As new forms of visual representation are introduced to clinical education, the question arises as to what their distinct possibilities for learning are. Consider, for instance, two learning resources designed for medical students preparing for the OSCE skin suturing station. One is a set of revision notes; the other one is a short video published on YouTube ([https://youtu.be/bE8SEOXjTpo](https://youtu.be/bE8SEOXjTpo)). The ‘notes’ consist mainly of writing, with some still images. The video consists of a much wider range of means of representation, including speech, gesture and animation.
One of the implications of choosing one of these multimodal configurations over the other is the following. The written notes offer generalised, abstract instructions – ‘rules of thumb’ – which students can apply in the actual, concrete situations they will be confronted with. For students who revised using the notes, the cognitive work involved is largely that of deduction: they are to follow a set of general rules to deal with the concrete, unique instance given in the OSCE. The video, unlike the notes, is organised around a concrete situation, showing an actual (simulated) body part (and, occasionally, patient), an actual suture and skin pad, et cetera, from which generalisations are sometimes inferred and articulated in writing superimposed and in animations added in the post production, edited stage. For students watching, the cognitive work demanded is largely that of induction: i.e. to analyse and infer from the concrete, unique instance shown a set of general rules that they can follow. While in principle both writing and moving image can be used to design the two learning environments, the level of ‘concreteness’ achieved in the video would be difficult to match in the written notes; and the level of abstraction in the writing would be difficult to achieve in the video.

In each case, the modes used shape the representation of the suturing procedure. For instance, by default, the video shows all movements involved in suturing continuously. In writing, choices always need to be made about what movements to select and which ones to leave out. Movements that appear simultaneously on the video are to be presented sequentially, one after another, in writing. Each of these epistemological ‘losses’ may at the same time count as pedagogic gains: a
‘reduced’ account may be perfectly fitting to certain trainees, relative to their knowledge and skills.

Given the mediating effects of modes, no two learning resources or activities provide the same potential for learning, even if they were designed to ‘cover’ the ‘same’ body of knowledge. In the contemporary world, clinical trainees are rarely restricted to only one site. Instead they move between – and learn – across many different sites, each uniquely configured: workplace, simulation centre, online platforms, et cetera. Visual research provides means for exploring what is ‘unique’ in each, and how they are used by educators and trainees to facilitate learning.

Discussion

Throughout the paper I have indicated how visual research contributes to a better understanding of clinical education. In this concluding section I wish to draw attention to some more general theoretical, methodological and ethical issues arising from visual research.

First, what transpires from the examples discussed is a picture of clinical learning and teaching as an unceasing process of making meaning of the physical and social environment. This notion resonates with theories of, e.g. symbolic interactionism,16 phenomenology,15 ethnomethodology21 and social semiotics.17 Whether as ‘educator’ or ‘trainee’, clinicians notice and give meaning to
(interpret, ‘read’) selected features of that environment, and, in the process, their ability to make meaning changes – a process also described as transformative engagement, indeed, ‘learning’. That world is a multimodal world: becoming a clinician is about learning to see, feel and talk about the world through the lens of the profession whose practices they participate in. For instance, in the case of surgery, recognizing a structure as an ‘artery’ from its location, shape, colour pulsation, resistance to pressure, and so forth; an instrument movement as a sign of clinical judgment; or a shift in gaze as a sign of an upcoming request. Much, though not all, of these processes can be accounted for using visual data. Audio recordings are equally important, and yet neither record what things feel or smell like; ‘data’ are always only partial representations of reality.

Second, as is typical of the social sciences more generally, accounts of clinical education have, until recently, relied heavily on, indeed are frequently dominated by, language, notably writing alone. In that frame, in that world, sooner or later, at any one stage in a research project, whether qualitative or quantitative, the multimodal world is reduced to a written account. Stuff treated as ‘data’ is already written and if not, is bound to be transformed into a written text. Observation tools used to measure the performance of trainees or teams are written, as are surveys used to gauge the ‘face validity’ of a new simulation. Interviews, though often video recorded, are usually reduced to written transcripts that focus on what was said, while largely ignoring gestures, facial expressions, and so forth. And of course, the research report is primarily written. The visual research discussed in this paper invites researchers to reflect critically on the implications of such methodological choices, asking, What is
gained and what is lost as learning activity is re-presented? What is it that the chosen means of re-presentation does and does not make visible, and therefore available for analysis and assessment? And what, therefore, is the preferred means of producing and disseminating my account of clinical education?

Third, in-depth analysis of the kind illustrated in this paper can be and has to date only been applied to relatively small data samples. Yet insights gained from micro-analysis can be used to develop sophisticated coding schemes that can be applied to large samples of materials, enabling quantitative analysis. Using this approach, associations could be explored between the occurrence of certain actions on the one hand and conventional global ratings by independent observers/researchers, trainers and trainees evaluating practices on the other. Doing so would make explicit exactly what evaluators believe to be indicators of ‘good’ educational practice, providing an operationalization of what are currently relatively vague statements (e.g. ‘the instructions were clear’), enabling us to develop more detailed, concrete suggestions for educators. A similar advancement could be achieved in relation to the assessment of clinical skills: combined with evaluations measured on global skills, a detailed visual analysis could make explicit exactly what counts as ‘good tissue handling’, for instance – providing more concrete suggestions for practical action.

Fourth, as with any data, the use of visual data requires careful ethical consideration. For instance, when researching online videos permission requirements need to be considered. When the materials sourced are in the public domain, permission from those who produced and/or disseminated them
may not be required, unless the owners state otherwise. In general, creating visual records at the workplace raises more concerns among research participants (and their employers) than creating written records, and these concerns, along with legal considerations (litigation), means that it can be difficult to create and collect visual recordings. They can only be made available for research purposes if voluntary informed consent was given by all those involved, including staff and patient.

One point to consider is anonymity and confidentiality of the research participants. Anonymity cannot be ensured in video research, unless you edit the video such that people’s faces and voices can no longer be recognized (software such as Windows Movie Maker allows you to do that). In doing so, the video also loses some of its potential for making visible the phenomenon of interest. However, confidentiality can be protected, by omitting personal information and information about the research site. Some participants do not have any objections to being identifiable. It is therefore common practice to give participants different options: ask them whether they agree that anonymized clips and transcripts be used and/or un-anonymized clips and transcripts for research and training purposes.

Some studies suggest that the use of video for recording and researching clinical care has declined significantly since the introduction of more strict privacy legislation. However, the approval of ethics applications for video based research by the UK’s National Health Service in recent years does show that it is possible to obtain approval for visual research in clinical settings, provided that
clear arrangements for who accesses the data and for what purposes are in place, and that the need for visual data is justified. It is hoped that this paper has made that need plausible.

References


