

METHODOLOGY

Open Access



Analysing videos in educational research: an “Inquiry Graphics” approach for multimodal, Peircean semiotic coding of video data

Nataša Lacković 

Correspondence: n.lackovic@lancaster.ac.uk
Lancaster University, Educational Research, Lancaster LA1 4YD, United Kingdom

Abstract

This article introduces an “Inquiry Graphics” (IG) approach for multimodal, Peircean semiotic video analysis and coding. It builds on Charles Sanders Peirce’s core triadic interpretation of sign meaning-making. Multimodal methods offer analytical frameworks, templates and software to analyse video data. However, multimodal video analysis has been scarcely linked to semiotics in/of education (edusemiotics), for the purpose of exploring higher education teaching-learning and settings. This article addresses the mentioned gap by introducing the IG approach, which links multimodality and edusemiotics primarily via Peirce’s triadic sign. The article offers a step-by-step IG coding guide, examples and explanations. IG application can be expanded to video analysis across many fields, levels and subjects, within and beyond higher education research, nationally and internationally.

Keywords: Multimodality, Video analysis, Edusemiotics, Peirce, Higher education, Educational research, Inquiry graphics (IG)

Introduction

This article’s goal is to introduce an “Inquiry Graphics” (IG) approach for video data analysis and coding, situated within global educational research more broadly and higher education research more specifically. IG can be appropriated across fields and educational levels, as noted at the end of the article. Its novelty lies in merging the approach of multimodality (Jewitt et al., 2016) and Peirce (1974, 1991) ‘s pragmatic semiotics, based on his notion of the triadic sign. The IG’s analytical steps progress from micro focus on the video’s material affordances (Pikkarainen, 2014) to macro focus on research key concepts and theory, and the other way around, seeing these layers in analysis as a part of one relational and unifying system of video data interpretation. I consider that this method efficiently links to edusemiotics (Stables and Semetsky, 2014; Olteanu, 2016; Olteanu and Campbell, 2018), the semiotic theoretical framework for education. In a nutshell, the IG provides interpretative guidelines to support researchers in multimodal, edusemiotic coding and analysis of video data.

The multimodality of higher education and video communication

In the last two decades, the global movements of multimodality (Jewitt, 2014; Jewitt et al., 2016; Iedema, 2003; Jewitt, Kress, Ogborn and Tsatsarelis, 2001; Kress and van Leeuwen, 2001), new literacies (Freebody and Luke 1990; Knobel and Lankasher, 2006; Lankshear and Knobel 2007) and multiliteracies (Anstey and Bull 2006; Cope and Kalantzis, 2000; New London Group, 1996) have been paving the way for renewed understanding of communication and education processes, especially in relation to technology mediation. Such renewed understanding emphasises that those processes go beyond language. They include various modes of meaning-making, and, consequently, new definitions of “literacy”, “communication”, “learning” and methods for researching them. This means that communication acts, including communication in education and at educational institutions, are fundamentally multimodal (Bezemer and Kress, 2015).

The modes of “multimodality” include: body movement and posture, gestures, gaze, print or computer screen layout, design, sound, tactile senses, material resources such as diagrams, photographs, illustrations, 3D models, liquids, video, any material thing that mediates teaching-learning interactions. Therefore, to understand educational processes beyond language, it is useful to apply approaches and methods that consider more modes than just language (Metcalf, 2015; Norris, 2004; Bezemer and Kress, 2015; Jewitt et al., 2016; Breuer and Archer, 2016), such as video.

Higher education research has commonly applied methods that are language-driven (Metcalf, 2015), such as interviews. Furthermore, it seems that when video analyses are applied, they are “often conducted as verbal conversations, taking none of the other modalities into account (Buhl, 2010, 116)”. For example, if we consider “video analysis” research in teacher training, it has been defined via an analytical focus on language and reflection (see for example Nagro and Cornelius, 2013 and Schieble, Vetter and Meacham, 2015). Such an approach has merits as it clearly focuses on the video as a reflection tool. However, it can be enriched with considerations of the materiality (of objects and the human body) that is brought into relation with social action and reaction in videos, as a part of an analytical procedure, and as a reflection trigger. Recent review of a large body of articles that utilised videos as teacher development tools (Nagro and Cornelius, 2013) defined “video analysis” in this context as video recording teachers’ lessons for “the purpose of analyzing and reflecting on their own teaching performance (ibid, 320)”. This use is what Jewitt (2012, 3) calls “video elicitation”. The mentioned analysis can be enriched by distinctively accounting for the relations between reflection and material affordances of the recorded teaching performance, and how the analysis itself was performed. This is not to say that video should not to be used primarily as a resource to elicit language-based reflection and feedback. Rather, it is suggested that an ongoing tradition of “video analysis” in teacher training that “has been researched in education for almost fifty years” (Nagro and Cornelius, 2013, 313) can be enhanced with more focused considerations of multimodal interactions and what this means for teaching. What is then the value of using videos and accounting for their multimodal interconnectedness of the social and material?

The tactile and visual aspects of learning can be seen as prominent in some disciplines (e.g. medicine, engineering, applied arts, media and communication), but all disciplines include such aspects. If we accept that teaching-learning acts and interactions are multimodal, then capturing them in their multimodal nature, via a video recording, provides an opportunity to understand these teaching-learning interactions more fully.

Such multimodal nature of communication is complex and layered. It can take much time to do an encompassing analysis. Therefore, more often than not, multimodal research focuses on particular modes or specific combinations of and relations among modes, including language. This focus needs to be clearly acknowledged. For example, the foci of multimodality studies have been, among others: hand gestures/movement (Sakr, Jewitt and Price, 2014), Power Point features (Zhao, Djonov & van Leeuwen, 2014; Kress and van Leeuwen, 2001), and the relationship between speech utterance and photographs in lectures at a range of UK universities (Hallewell and Lackovic, 2017), to mention but a few. In general, when applying any method, it is useful to consider and acknowledge what the analysis covers exactly. A video research method in education, as it is the case with any research method, needs to be fit for purpose.

Video recording of any communication in education can serve the purpose of capturing and exploring *the nature, characteristics and features* of educational events (a seminar, lecture, project presentation), for example as an exploration of the interplay between the spoken and material (e.g. learning resources and body movements), and as a trigger for pedagogical feedback, or research participants' reflection. Sakr, Jewitt and Price (2016) provide an example of video analysis that explored emotional engagement in the context of primary school history learning. To understand emotional engagement, visual cues such as facial expression, gesturing, gaze, body posture and proximity, the sound of voice, and movements can be critical, in addition to any spoken word. In another classroom-based research, Sakr et al., 2014 show how gestures form an important part of students' engagement with touch tables when learning scientific concepts. Jewitt's (2012) National Centre for Research Methods's working paper offers a comprehensive overview of why and how to use video in research, including associated disadvantages, hence it is a useful place to start when planning a video research. In addition, "Video research in the learning sciences" (Goldman et al., 2014) offers diverse information on doing video in educational context. In higher education settings, Otrell-Cass (2018) provides examples of applying video ethnography with students in relation to their algorithmic thinking, as they document how scientists in different fields find solutions to problems they are facing in their daily work and research.

A great scope of multimodality studies has been done in the context of media and communication and to analyse publicly available artefacts, such as advertisements (e.g. Thibault, 2000). However, there is a paucity of multimodal studies about higher education teaching-learning. Some exceptions include the work of Archer (2010) and the recent edited collection on multimodality in higher education by Breuer and Archer (2016). This article addresses the stated gap by providing a multimodal semiotic analysis and coding *approach*, to support video analysis in the context of Higher education studies, building on Peircean semiotics and the related emerging educational theory of edusemiotics. The approach is termed "Inquiry Graphics" (IG) due to its focus on inquiring pictorial information in a triadic interpretative manner, in relation to other modes (e.g. language), and theoretical research concepts. The article proceeds to consider multimodal approaches for analysing and transcribing video data that the IG approach is related to. This is followed by an introduction of a triadic Peircean model of meaning-making, as one of the key element in the emerging field of edusemiotics, and the model's adaptation into the mentioned IG approach.

Multimodal approaches to analysing and transcribing video data

Multimodality movement has offered analytical frameworks, templates and software to perform a video analysis (e.g. Bezemer, 2014; O'Halloran et al., 2011; Norris, 2004; Harter and Otrell-Cass, 2017; Otrell-Cass, 2018). The multimodal approaches to video transcript and coding that have informed this model are Bezemer's (2014) video analysis building on Charles Goodwin's (interactional) conversation analysis and social semiotics, and Sigrid Norris (2004)' multimodal interaction analysis of videos.

In his analytical approach to video analysis, Jeff Bezemer (2014), often working on multimodal analysis collaboratively with institutional colleagues and scholars in multimodality - Carey Jewitt, Gunther Kress and Diane Mavers - builds on the tradition of conversation analysis (CA) and social semiotics (Kress, 2009). Charles Goodwin's CA work is notable in this area (Jewitt and Bezemer, 2016), inspired by interactionism in sociology research (of Blumer, Goffman and Garfinkel). Goodwin's analysis of video recordings of everyday interactions was some of the first ones to take into account the materiality that shapes those interactions such as bodies, gestures and movement, an embodied action approach (Jewitt and Bezemer, 2016). Bezemer's (2014) approach to multimodal video analysis is at the level of fine grain microanalysis focused on the gesture, body and gaze as salient semiotic modes, with occasional speech utterance in the context of an operating theatre. The video transcript provided by Bezemer (2014, 162) consists of vertical lines to signal temporality of action and allow for noting what happens in the video, accounting for the movements of key people participating in the recorded action, noting spatial directions of how body, gesture/hand and head/gaze moves. Similarly, the IG approach introduced here accounts for the movements of different body parts, linking them to their social meanings.

Another author who has developed video analysis and related templates is Sigrid Norris. Norris (2004) offers a system of codes that researchers can use for coding various types of interactions in videos. The main unit of analysis is *action* distinguished as lower and higher level action. Lower-level action is defined as "the smallest interactional unit" (Norris, 2004, 11). If we consider gesture, a complete gesture from beginning to end is a lower-level action (Norris, 2004). Higher-level actions are constructed via chains of lower-level actions by social actors (who are video-recorded) (ibid.). To illustrate these definitions following the example given by Norris (2004), in a video of a business coaching session, requesting to make notes is a higher-level action consisting of a number of lower-level actions such as: making a gesture to reach out for an ipad, taking the ipad, and commenting on that intention. Norris (2004) pays particular attention to gaze and gesture, also posture and proxemics, the focus also prominent in the work of Bezemer (2014). "Proxemics" (Norris, 2004) is synonymous with "proximity" that is relevant to the IG approach introduced here (the proximity or distance between and among objects in a video). "Proxemics" focuses on the distance the individuals take in relation to other individuals and relevant objects. The distance among people suggests formality or informality of the relationships and encounters. With regards to "posture", people's postures signal their level of involvement. This can also be applied for "gaze" (where the look is directed). "Gestures" (Norris, 2004, 28) are identified as:

- iconic = mimicking concrete concepts expressed verbally, e.g. describing objects' shape to make them vivid,
- metaphoric = abstract concepts given some metaphorical shape,
- deictic = pointing to objects, people or abstract ideas, and
- beat = like a music beat, up-down, in-out hand movements.

In the IG model proposed here, the researcher starts with the materiality where actions are embedded, but in the process of analysis would identify salient actions to focus on, including considerations of gestures, proximity, gaze and so on, as relevant to the research undertaken.

The rationale behind introducing a new multimodal analytical template is to specifically link the robust enough micro and meso analysis approach of multimodality with Peircean triadic reasoning. Many consider that Peircean semiotics is the core approach in the emerging theory of edusemiotics (Stables and Semetsky, 2014; Nöth, 2014, b; Semetsky and Stables, 2014; Olteanu, 2016; 2014). Therefore, the model contributes to the needed body of work to link educational philosophy and theory with multimodal methods (Breuer and Archer, 2016).

Edusemiotics and Peirce's sign for an "inquiry graphics" video approach

To understand the semiotic account adopted in the article, "semiotics" is first briefly defined. The essential character of communication is semiosis, or meaning making, "the action of signs" (Sebeok, 2001, 1991; Semetsky, 2005, 230). This means that semiosis deals with the *interpretation* of the perceived communication acts and distinct communicational units, which are called "signs" (such as a textbook text, a road sign, a blog content, a slideware presentation content, a photograph, a gesture, a chemistry model, and so on). This "interpretation" of signs is influenced by various factors, such as communicator's motivation, socio-cultural power, ideology, communicators' status, the most immediate and less visible environment, including the socio-economic and layered ecological system that it is embedded into, questions of class, race, economics, biosphere, and so on. The field that researches how signs make (/produce and evoke) meaning (semiosis) is semiotics.

Semiotics draws attention to the deep understanding of what is communicated, anything that that might be (e.g. social semiotics, developed by Robert Hodge and Gunther Kress building on M.A.K. Halliday, is mainly concerned with how signs are motivated in society and what intentions drive sign exchange). Different sub-fields of semiotics also include: biosemiotics (e.g., Sebeok 1991) animal (zoosemiotics) (e.g., Martinelli, 2010) and other specifically formulated types of and views on communication, such as geosemiotics and ecosemiotics.

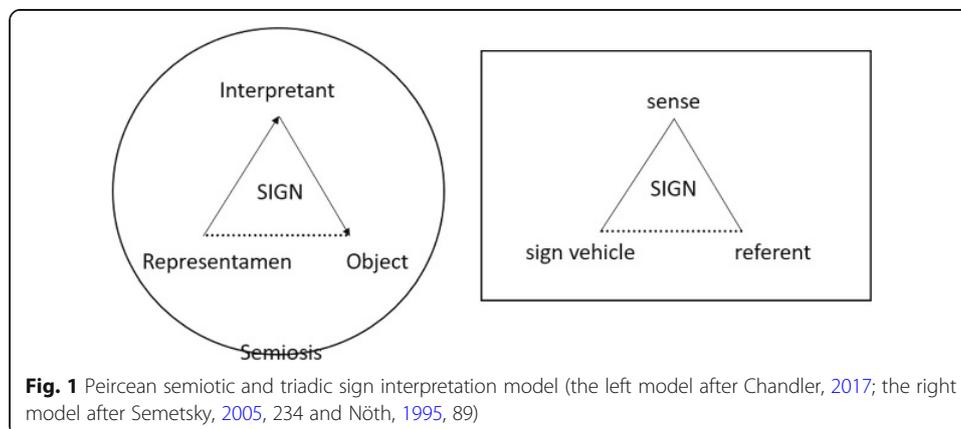
Edusemiotics is an emerging philosophical and theoretical approach to learning, knowledge and education. It recommends semiotics as providing the core conceptualization for a philosophy of education liberated from the rather rigid assumptions of analytical philosophy, which have dominated this area for some decades (Stables and Semetsky, 2014; Semetsky and Stables, 2014; Olteanu, 2014, 2016; Stables, 2012, Stables, 2006; Olteanu and Campbell, 2018). It could be seen to bridge American pragmatism, European semiotics (e.g. Nordic), Vygotskian constructivism and continental post-structuralism, building on a range of thinkers besides C.S. Peirce, J. Dewey, G. Deleuze and J. Kristeva. Other

schools of thought can be associated with edusemiotics, in relation to observing the relationship between humans and others (other humans, biosphere and artefacts). Here I build on this developing field and philosophy, alongside multimodality, *particularly focusing on Peirce's (1991) sign triad*. Via Peirce's triad, I adopt the edusemiotic view of relatedness between material and conceptual/abstract entities in communication (Sebeok, 1991, 2001; Stjernfelt, 2011). To illustrate this approach, Semetsky (2017, 704) argues that:

“Charles Sanders Peirce's philosophy did not limit signs to verbal utterances (...) Peirce's perspective (...) emphasised the process of sign growth and change called semiosis, representing the action, transformation, and evolution of signs across nature, culture, and the human mind. In contrast to isolated substances such as body and mind in philosophy of Descartes, a Peircean genuine sign as a minimal unit of description is a tri-relative entity”.

Peirce developed a rather complex and elaborate system of semiotics rooted in pragmatist relations between the mind, the world of concrete existence and representation (Peirce, 1991). This entire section and the remainder of the article builds both on edusemiotics and Peirce, since the two are inseparable (Olteanu, 2014). Philosophers of education who develop the semiotic approach to education as edusemiotics, for example Andrew Stables and Alin Olteanu, have written about edusemiotics foundations. Specifically, this article links Peirce's triadic meaning-making model of how humans interpret signs, central to edusemiotics, to an approach (IG) that is also encompassing of research on multimodality in society and education. A video represents one sign, in this case, related to human interpretation: a video *represents* what it refers to, its Object, something that happened and was video-captured at one point in the past. It manifests meaning via an interpreter (who is necessary for the sense making or, using Peirce's term, for the Interpretant to occur). The core model and structure of triadic sign meaning making introduced by Peirce is presented in the Fig. 1 below.

It is important to note that representing Peirce's sign as triangle is not perfect and ideal (Olteanu, 2015), but it works for the methodological purpose of this article. Peirce's sign interpretation (semiosis) model on the left in Fig. 1 consists of: Representamen (R), Interpretant (I) and Object (O) (Peirce, 1976, 1991;). All elements of semiosis always happen *simultaneously*. The rectangular frame in Fig. 1 on the right uses



different terminology for the same model sides in the circled model: Interpretant = sense, Representamen = sign vehicle and Object = referent.

Peirce's sign is a holistic and relational entity that does not subscribe to the ontology of separateness and Cartesian dualism. It suggests a synergy of "the material" and "the conceptual". Of course, Peirce's semiotics also has a *schematic* understanding: while meaning is of one piece, within it, *the three inseparable components can be noticed and analysed in relation to each other*. For empirical and practical reasons, *coding in the IG approach introduced here* is led by focusing on the *researcher's interpretation* of the three components of Peirce's semiotic sign, its distinct material (physical, sensed) affordances (Pikkarainen, 2014) and how they are related to and therefore melded with socio-cultural and historical particularities of the context.

The first step is to explain how the proposed IG approach relates to the video as Peircean sign in further detail. Within IG, *Representamen* is what is represented in a video, or in a photograph, a frozen or moving video moment, a representation of some space and time that was captured in the real world. If a tree is in a video, the features that make it look like something most humans recognise as a tree *form* Representamen. An understanding that the material quality of the tree and its context make it a tree corresponds to *Interpretant*. Representamen refers to the *Object* it represents. A video showing a tree *refers* to a particular plant species that exists in the world - its *Object* - both a generic idea and the real world existence of a tree in the world. In a meaning-making process, Interpretant connects Representamen to its Object. This triadic whole is meaning. The word or a video representation of something is not that thing, but refers to it. What kind of generic ideas of objects, concepts and phenomena people have varies. On different occasions, there are evoked different ideas of what something means. Context plays a crucial role in defining meanings. That is why it is essential to observe educational communication as an act of *contextualised semiosis*, within which personal experiences of the world can differ, either to a small or, perhaps more often in increasingly intercultural learning environments, large extent. Yet, it is important to share meaning-making commonalities, and establish what brings humans and cultures together via interpretation, aside many unique differences in interpretation of the same sign.

Representamen requires a mind's *interpretation*, just as Interpretant does. The concept of mind in Peirce is complex and non-Cartesian (Pietarinen, 2006). All three key elements in the IG approach are interpretative. However, Representamen's role in the IG approach is to focus attention on representational elements' manifestations (in a video) and their identification, which will then be positioned in relation to other levels of interpretation in the analytical meaning making of a video. Triadic elements are introduced in a hyphenated form "-led" to signify that the codes are not endorsing separation of entities and *are* not those entities (Representamen-Interpretant-Object). Rather, three sides of the Peircean sign support analytical coding. A translation of the Peircean triad into an analytical transcription and coding models is introduced in Table 1, after the explanation of all individual coding elements as follows.

Representation focus (Representamen-led): Video elements identification

For the purpose of video analysis, Representamen is named "Representation focus". It identifies individual key Elements in the video.

Elements

Under Representamen, researchers list every individual object (Element) they see, animate or inanimate as a list of nouns. *Element* categories represent more generic and commonly shared interpretations of objects: a man, a woman, a table, a face, a book, a paper. These are elements and some can be sub-elements, for example, a face is an element of a body. To further illustrate this, in a video that shows a seminar room, students might talk in groups at desks with laptops, notebooks, and many other things, such as mobile phones, pens, bags, umbrellas or sunglasses. All these objects, including participants' bodies and what constitutes those bodies, are the represented materiality in the holistic frame of lecture meaning-making.

The manner in which Representamen is adopted in IG aligns with how Nöth's (2014) description of the linguistic sign type *rheme*, Peirce's semiotic concept of a logical predicate. As Nöth explains a singular rheme cannot convey any particular information per se, unless it is brought into relationship with other things (2014, 16). A single (linguistic) noun, as one of a list of individual elements, does not convey any particular information per se (a man or a pen could be any man or any pen in any possible world). That is why Representamen, as they come into an interpreter's interpretation, suppose a rhematic syntax: this means that the elements in the video are simply listed.

The level of fine grain detail or focus in identifying elements varies depending on the researcher and research. A researcher-analyst has to make practical choices of what level of nuance and detail she should adopt in identifying an element unit (sometimes it is important to focus on the eyes or hands, but sometimes the level of analysis can be at a less micro level and focused on particular aspects of Representamen and action. This is further explained in the final paragraphs of the next section.

Sense focus (Interpretant-led): Action denotation and connotation

In the next step of interpreting the seen (Representamen), the focus moves to the Interpretant or Sense vehicle. Interpretant further branches the analytical reasoning into Denotation and Connotation of represented elements and of represented composition, adapted from, but *not* neatly aligned with, Roland Barthes (1973, 1977) and Peirce's notion of "dicent" (Nöth, 2014, b). When nouns (a woman, a laptop) are paired with verbal information (a woman is sitting on a chair) in the next step of coding interpretation (Interpretant), this "pairing" provides a descriptive information.

Nöth, 2014, b, 16) exemplifies how a linguistic rheme/rhematic symbol "rock" turns into dicent:

"When we want to communicate the information that "this rock is grey" we need to combine the rhematic symbol with indices (this, present tense) and icons (the mental image of color grey), and this combination results in a sign (as a sentence) that is not a rheme anymore but a dicent." (Nöth, 2014, b, 16).

Information is conveyed by rhemes being interpreted in conjunction to indexical signs, forming propositions, or *dicents* (*sign types*) or *dicisigns* in Peirce's terms (Stjernfelt 2011). Therefore, this step in IG analysis aligns with Peirce's *dicent*. Interpretant branches into two levels of interpretation: Denotation and Connotation.

Denotation in the analysis means to move analytic reasoning from identifying and stating the represented elements (Representamen) to *elements' description in the action they are a part of*, that is, what is shown during the chosen minutes or in a frozen moment, *without getting deeper into the socio-cultural context of the situation*. Denotation does have socio-cultural embeddedness like Connotation though, but this is at a more *basic interpretation level*. Within Denotation, particular attention is on the embodiment of an action, when observing people: it is useful to define the type of “gesture” (as linked to the element “hand (and arm)”), “posture” as linked to the element body, and “gaze”, as linked to the element “eyes” (and “face”). This means that when humans are identified within Representamen, this identification would commonly include elements of the embodied action: hand, eyes, body (torso).

Connotation level demands to observe *the context*, understand it, and watch the unravelling of the action. It is linked to socio-cultural knowledge of the represented elements as they are brought into relationships. Connotation does not yet involve interpretation of the video in relation to research goals and questions – this is the final analytical step of the Research Object. *Connotation provides contextual meaning of Denotation via connotation question*: How does the context given and what is happening determine the more specific socio-cultural meaning of Denotation?

To exemplify, an **Element Denotation** of an element “a male person/man/a person who looks like a male person” (or in smaller element units: a man/hand/face) could be “a young adult male is raising his hand, eyes and eyebrows in an expression signalling emotional discomfort”. The focus is on describing the element action in *more generic terms*. When interpretation moves to **Element Connotation**, the researcher considers the context of the action, its place, space, actors, and identifies contextually what happens in the video over some time (seconds or few minutes). Such consideration moves the analysis further to connote the meaning of the video. In the case of the mentioned description of the raised hand and facial expression “Denotation”, its connotation would be “a male student is seeking permission from the lecturer to speak and ask to go to the lavatory”. The male person in the video is now assigned its contextual role: a male *student*. In a different context the same person would have a different socio-cultural role, for example, a gamer/game player, a boyfriend, a son, a customer, and so on. The action of raising hand is now contextualised as “seeking permission”. Hearing what the student says together with student action, if there is speech in the video, creates a holistic meaning of that action.

Composition

Composition moves the focus from individual elements to relationships amongst elements, what such relationships create in design and compositional terms. This positioning and relations could be in terms of proximity, layout, interactions.

In essence, **Composition Denotation** would consider what is happening with and around a few elements, or consider all of the represented elements (at the level of the whole video scene). For example, composition denotation of a video could be “a large number of mostly younger looking adults of mixed gender who appear to be looking at the screen, which is displaying bullet points and text in front of a male person”. From this example, by the virtues and characteristics of the environment and context

observed, **Composition Connotation** would be “a large number of *students are following the lecture content* displayed in the form of *textual data, the lecture being led by a male lecturer*”. Distinguishing between more generic and contextually defined action can help researchers understand in greater depth how interpretation links material senses and conceptual understanding, how things make meaning in human communication and how they depend on the context, how action (together with spoken or written word) happens, under what conditions and, importantly, inclusive of what other material objects and movements as mediators of that action.

Where and how exactly to draw the line between denotation and connotation, and between representational and connoted element might not be always clear. This, in itself, is telling. Connotation assigns *contextual* meaning for denotation. As mentioned above, *it is the context that helps identify a particular (changing and fluid) role of a person* – e.g. “*a man*” - and hence the meaning in a given context: a student, a father, a son, a book shop customer, a film enthusiast, a patient, a cook, etc. Certainly, interpretation is always dependent on the interpreter and her/his knowledge. Some things might not be what they seem to be to an interpreter (a perceived man might be a woman or transgender). These are the known limits of observational qualitative data coding. We can only make interpretations within our limited and contextualised knowledge.

A researcher-analyst has to make practical choices about the level of nuance and detail she should adopt. “Sad looking” could be a connotation element linked to the element “human (man, woman, person)” or linked to a smaller element “face”, if the researcher wants to focus on nuances of facial expression and emotions. The body is the communication resource of humans, as well as of any animals (Stjernfelt 2006). Hence, the researcher can choose any small, composing units of elements – e.g. face, hands, legs, torso and further into lips, eyes, eyebrows, fingers, shoulders, and feet. If we consider “sad looking” as element connotation for element “face”, “mouth corners pointing downwards and inner eyebrow corners pointing slightly upwards” would be an element denotation of that same element face, and “mouth corners pointing downwards”, element denotation for an even smaller element unit: a mouth. This is to illustrate how fine grain detail the research can choose to apply or *not to* apply. It depends on the level of analysis, the time available for the analysis, research goals and research questions. If the researcher is particularly interested in gaze and facial features, gestures and/or emotions, this requires finer grain analysis and focus on associated body elements. The finer grain the analytical need, the smaller the element unit is. Yet, researchers can do a type of more meso or selective analysis where the focus is, for example, on only selected composition only, and particular action focus.

Action focus

Action focus specifically highlights and extracts video actions which are most salient for the analysis in question, to help researchers focus on particular action and look closely to all the elements linked to that action. This action could be for example identified either as a description of an action, or by writing a gerund form of action (e.g. singing, pointing, notebook giving). Action can also be assigned to non-animate elements, such as “the pen falling” (e.g. it falls off the table rather than being dropped by a

hand). Once the action focus is defined, the researcher looks into the elements related to the action as well as key conceptual framework and research questions to guide interpretation of the video.

Speech/voice/sound transcript

Chunks of text are transcribed and embedded in the template as they happen within the chosen duration of the video that is to be analysed. Text is then related to the representational/visual side of video action: two meaning systems or modes of different kinds (speech and the visual) will merge to produce meaning. The researcher brings together speech and analysed materiality of the event (Representamen and Interpretant), identifies thematic units, and relates the material video action to the meaning of the spoken words to establish relationships. The researcher would code the speech text at an intersection of all aspects of the model so far, thinking about the following questions:

- What level of the represented (representation and sense focus – element, element denotation/connotation, composition denotation/connotation) does the speech tackle, if any, and what does this mean? How do the material setting and objects provide the context for the speech?
- If the speech does not refer explicitly to any of the observable actions, what can be concluded about the content of the speech on its own (e.g. by applying theme analysis in relation to research questions)?
- If the speech happens simultaneously with another action, with or without explicit reference to that action, how do the action and speech relate?
- What can be concluded about the relationships between the content of the speech and the represented (Representamen + Interpretant), in relation to Research Object (research aims/ key concepts/theory/ questions)?

The spoken word will reveal environmental/contextual meaning of the material and the material will expand the spoken word by providing material situatedness for where the speech takes place. These things are one in an educational event. Therefore, researchers can think about the relationships, to what extent they happen, how, why do they happen or not. This moves the article to the final analytical step of Research Object.

Research object-led: Interpretation via research theory, questions and aims

The true Object in Peircean terms is the environment or setting where the video was recorded in real life. A video format is only a representation of something that once was video recorded, its Object. When the video is used for analytical purposes, as proposed in this article and IG, the IG Object has a very different role. It becomes Research Object. This means that all the previously analysed video elements and compositional characteristics (Representation focus, Denotation and Connotation) are now considered in relation to the accompanying speech in a video excerpt chosen for analysis (speech transcript), *guided by* research questions, aims and theoretical or conceptual framework of the research. It is important to interpret these in relation to key concepts relevant to research questions.

Research Object helps to limit the wheel of interpretative associations (endless semiosis (Peirce, 1991) and focus the researcher on the meanings of individual elements, levels of denotation and connotation in relation to research questions or key concepts in a theory (e.g. critical theory, theories concerning motivation, identity, historical analysis, and so on) or any conceptual framework the researcher has adopted. One example of a research question can be: “In what ways do the visual data (videos) of video lectures convey aspects of teacher’s identity and the nature of the lecture?” Such a research question focuses the interpretation of the observed elements, composition and action. The Object-led side of the model functions as the “linking glue” for the Representamen-Interpretant (–led) side of the transcript and the transcribed speech. It brings together all parts of the IG interpretation and transcription.

In Peircean terms, as used in edusemiotics, Research Object represents a type of argument (Noth, 2014). It means that it takes all various levels of interpretation in question to come to some research arguments, via the process of *associative thinking* as the discovery of similarities among entities and things (Olteanu, 2014, 2015). This analytical and interpretational discovery is fixed via the research aim, objective and questions, including any theoretical and/or conceptual framework.

Units of analysis

A unit of analysis (in a video recording) is purposefully chosen by researchers. Commonly, these units would be particular moments in the video and a combination of particular modes (e.g. gestures, gaze, speech). Within IG, an analytical unit is “a video excerpt/scene showing modes present in some action or phenomenon, salient for research focus”. Within the scene, a researcher could focus on one mode or action. Generally, these scenes are identified and chosen for their action or mode salience, for example, choosing to focus on particular actors, or types of actions, modes, and combinations of those. Action salience is defined in relation to research focus. Ideally, an entire video is analysed if the project spans across a considerable timeline, but this is not the case in most research projects. Hence, choices are commonly made and explained. The researcher might want to identify *variations* of actions and then extract them as units of analysis, and/or identify *the same/repeated actions* in order to understand their nature in context.

For example, “a pen” is an element; “writing” could be an element action linked to an observed element pen and hand, which can be extracted in the list of “Focus actions” (Table 1), if found particularly relevant to research aims and questions. Such focus helps the researcher pay close attention to individual elements represented and what actions are related to them. The numbering of elements can sharpen researchers’ focus to make relational inferences between the number of elements and what this number means in the given research context. To illustrate this, if the video shows a lecture theatre with one projected slide screen and rows of many students, making relational inferences means to consider what the relationship between the number of students (many) and the number of displayed resources (one) can mean, i.e., what it means for a large number of students to look at one display. There are many other insights that can be inferred from such focused identification of elements: student positioning in relation to the screen, screen and student visibility, proximity between peers and in relation to the screen, how the lecture’s nature affect those, and vice versa, how any material configurations influence and shape the progress of the lecture.

Different code relations can bring different analytical insights. For example, the Representamen-Object relation sets a focus on the meanings of elements/objects historically and socio-culturally, not only the meaning of the interpreted action that these objects constitute. This can be an analytical focus in its own right: tracking historical development of particular learning spaces and their elements/objects. Thesen (2016) provides an example of historical lecture analysis, with the focus on lecture modes and gaze. Thinking about the meaning of both animate and inanimate objects constituting an action in this model opens up a new element of multimodal video analysis.

At this point, Table 1, with all the codes explained beforehand, is introduced below.

Tables 3 and 4 embed some of the aforementioned illustrative examples for the Sense vehicle/Interpretant in a grid template layout (Table 2), to exemplify how the mentioned codes can be applied in such a layout.

Table 1 Core analytical codes of the “Inquiry Graphics” model which embeds Edusemiotic principles of Peircean meaning making triad

Representation focus (Representamen-led, sign vehicle)	
Element (E)	Anything (any object) represented that the researcher can observe, including animate objects such as humans, is identified as a list of individual nouns (Elements). Starting from the Element, researcher decides the level of the analysis needed (how fine grained or not)
Numbering	Identified Elements are counted and numbered
Sense vehicle focus (Interpretant-led): Denotation and Connotation	
Element Denotation (ED)	Basic generic <i>description</i> of an action or state focusing on a particular Element (e.g. face, hand) is provided. This also refers to element positioning/states, what happens to them; e.g. “a hand is doing a gesture, moving up and down”. The quoted description is an ED of the element hand (of an element body).
Element Connotation (EC)	This means that ED moves to the next level of interpretation: this code assigns an environmental/socio-cultural/contextual meaning to ED (How does the element, context and socio-cultural particularities jointly make this action mean what it means?). E.g. the gesture’s meaning in the given context is considered and stated. Connotation is still a descriptive code.
Composition Denotation (CD)	A more basic description of an action linked to any identified composition/relationship of a few elements; any chosen composition or what is happening holistically in an entire scene (1 whole video scene or smaller “scenes” with a few elements)
Composition Connotation (CC)	CD is more strongly assigned an environmental/socio-cultural/contextual meaning (how the element, context and socio-cultural particularities jointly make this action mean what it means in the given context)
Action focus (Relating to ED-EC; CD-CC)	Extracting particular action focus from animate or inanimate element or compositions, extracted to be the focus of further analysis (e.g. for humans (body): sitting, gesturing, gazing; for objects: falling, showing content, running on batteries, etc.) – the researcher extracts the main “action” with element/composition focus for analysis, based on previous codes, and as relevant to research aims, questions and theory/concepts.
Speech/sound transcript	Speech is transcribed and thematically coded for meaning <i>in relation</i> to the visual video transcript and research aims/questions/key concepts.
Research Object (Object-led)	
Critical and analytical inferences, conclusions and connections made in relation to research goal, theory, and research questions. This means to do a multimodal thematic coding of data by linking holistically key theory terms/key research aims and concepts to an IG multimodal transcript (this includes speech transcript together with the focused “Representamen-Interpretant” interpretation). This is the point when the researcher asks: “So what? What does all the presented and described (e.g. Denotation and Connotation) mean in the context of my research?”	

Table 2 An IG analysis in a grid template layout (also Tables 3 and 4): columns show simultaneity of each coding unit within any chosen video scene. Codes could be presented in rows too. Research Object could be developed on its own by looking at the other codes in the table, ending with Speech

Video scene (optional: video image)	Representamen-led (Representation focus)	Interpretant-led Sense vehicle	Action focus	Speech	Research Object-led

Educational research context: Applying IG model in video research

The IG model has been named “Inquiry Graphics” since it *starts* with inquiring the graphical/pictorial in a video, or a still image, a frozen video action. The act of inquiry relates this graphical materiality to speech/sound, video action, research aims, questions, theory and key concepts. It will, as all models, theories and ideas do, evolve further, with the help of research community. In the context of educational research, there are many possible avenues where video can be a part of research. This is still an under-applied method in higher education, especially in relation to the notion of exploring multimodal, teaching-learning interactions holistically. Some possibilities and challenges are commented on below, and further in [Discussion](#).

Video can be a tool used to capture teaching-learning interactions, for example record a lecture, seminar, presentation. The researcher might want to focus on particular relations in a teaching-learning context. For example, relationships between the materiality of teaching-learning resources (often digitally presented), voiced content and the development of seminar’s and lecture’s conceptual unfolding and students’ actions. Video can illuminate how knowledge and interactions unravel in a lab or in a teacher-training context or any mentorship and community of practice situation. There is a scarcity of research to unpack teaching-learning unfolding at such a level of analysis; hence, there are exciting opportunities for exploration awaiting.

Furthermore, students can keep video diaries; researchers can video record spaces such as libraries and various facilities that library spaces provide, including learning, collaborative work and snack zones. However, it is acknowledged that recording in large spaces can be particularly challenging. While planning these and any other video related action, it is essential to devise the best strategy to assure an ethically sound project. Obtaining relevant information from ethical committees and external specialist organisations is important in the planning stage. A growing body of research that has utilised videos in various contexts outside of education can provide insights about approaches and solutions applied in those different contexts and lessons learnt. “An Introduction to Using Video for Research” (Jewitt, 2012), NCRM (National Centre for Research Methods) working paper offers an overview of video approaches such as videography or video elicitation. Today’s expansion of video equipment includes portable and mobile mini cameras, like GoPro. Such cameras provide opportunities to create “video trails”, for example of students’/participants’/researcher’s movements (e.g. around campus, when commuting, and so on). Of course this has to be handled carefully in terms of impinging on people’s privacy and anonymity. Today’s technology has advanced so that it can offer possibilities to protect the anonymity of video-recorded people, such as face or body blurring technique.

In addition, video is frequently used in the context of virtual learning environments where teachers-researchers upload their video lectures, e.g. captured by lecture- capture

Table 3 Illustrations of the coding examples provided in the article, related to elements, denotation, connotation, action focus, and speech, applied in a grid template

Video scene (optional: video image)	Interpretant Sense vehicle		Action focus	Speech	Research Object
	Representation (Representation focus)	Denotation (ED)			
Time duration	List	No.			
Elements	1) man/body// hand//face 2) screen 3) laptop	1) 1/1// 2//1 2) 1 3) many	1) a young adult male is raising his hand, eyes and eyebrows in an expression signalling emotional discomfort 2) a male student in an English speaking environment is seeking permission from the lecturer to speak, looking distressed and asks to leave the seminar room Hand: raising/seeking permission	"Can I please leave the session now?"	Interpretations depends on research questions and focus.

Table 4 Illustrations of the coding examples provided in the article, related to composition, denotation and connotation, applied in a grid template

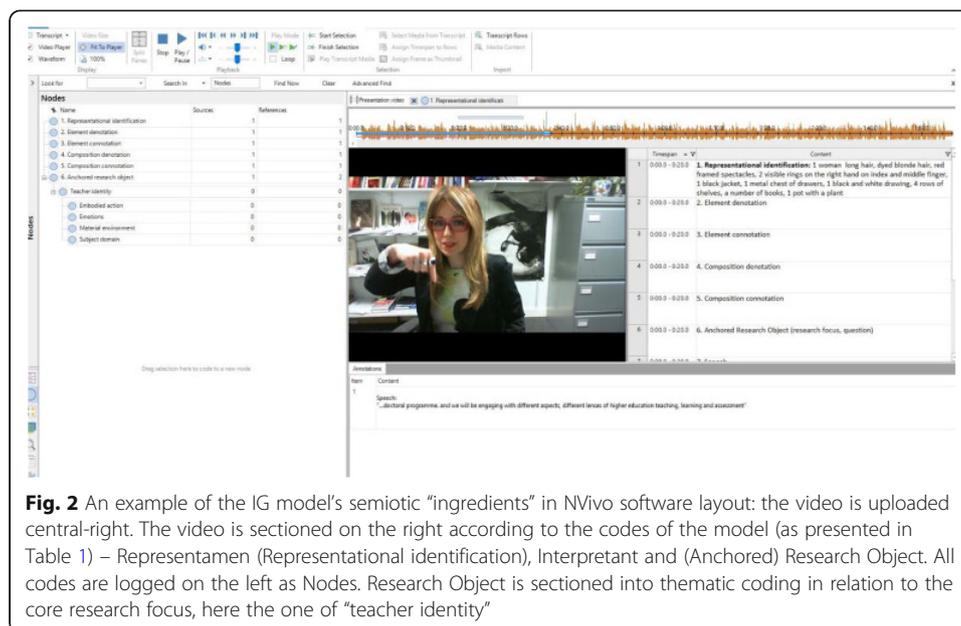
Video scene (optional: video image)	Representamen (Representation focus)		Interpretant Sense vehicle		Action focus	Speech	Research Object
			Denotation (CD)	Connotation (CC)			
Composition	1) people + Man + screen -----	Many + 1 + 1	1) a large number of mostly younger looking adults of mixed gender appear to be looking at the screen, which is displaying bullet points and text (in English language), in front of a male person	1) a large number of students are following the lecture content displayed in the form of sequential textual data in English language, the lecture being led by a male lecturer			Interpretations depends on research questions and focus.

technologies or created for online students using various software, such as Panopto or Screencast-O-Matic. Such videos could be a useful resource for exploration. Moreover, students and general population increasingly learn from You Tube and other similar platforms. Therefore, the researcher could ask: How does this process happen?

IG transcript choices: Layout, framing, selecting, highlighting

The proposed IG model for analysis and coding should be further elaborated by considering possibilities of its layout (how to represent the pictorial and linguistic modes of the video) and transduction, that is, the choices of transforming one mode into another (e.g. a video scene or still image into language or drawing) (Bezemer and Mavers, 2011). Researchers need to consider the following, when deciding on the form of the transcript (Bezemer and Mavers, 2011):

- **Layout:** decisions about how image and language will be set out on the page or screen, whether a transcript consists entirely of writing or entirely of images, or a multimodal mixture of the two; researchers use spatial organization to construct separation and cohesion, to disconnect certain parts of the writing and images and to show which parts belong together, e.g. a horizontal line can signify temporal unfolding (ibid, 202).
- **Framing:** video extracts that are selected for transcription are both framed by the communicational aims of the original interaction and by the purpose for which the graphic version is being made (ibid, 194). Framing is motivated by research aims and intentions (Research Object) that led the video recording.
- **Selection:** researchers choose to include a video extract that involves certain participants and excludes others (ibid, 195).
- **Highlighting/salience:** what is highlighted in the transcript, or which of the re-made features are given prominence (bigger font, circled object, a drawing outline to emphasise posture, and so on).



Bezemer and Mavers, (2011) provide further explanations of multimodal transcription subtleties. The present author leaves the choices of how to combine the core codes and mentioned transcript particularities to the readers-researchers. For example, a transcript might involve video stills in the form of images. "Frozen" video images can be shown in panels with added speech or speech bubbles, e.g. in a manner of comic-like graphic transcript with panels (Laurier, 2014). The analytical model proposed here could be "translated" into an analytical template grid, involving (or not) still video images (Tables 2, 3 and 4). It could also be "translated" into analytical software, such as NVivo, which allows for working with the video itself during the coding. A related illustrative example is shown above in Fig. 2.

Why use the IG approach?

The IG approach to video analysis is distinct from other approaches to video analysis in that it builds on Peirce's triadic model of meaning making, as the main approach to how the observed makes meaning. This particular school of semiotics is proving rather salient in contemporary social semiotic approaches to multimodality and education, which particularly recommends it for the analysis of video educational content. The IG approach highlights individual Element signification via the proposed appropriation of Representamen through Elements and Element Denotation, Connotation and how those Element features link to Research Object. IG helps the researcher to interpret individual elements and their signification, unpacking the meaning of the elements both historically (outside the video via their historical development and signification), socio-culturally, and contextually (as shown in the video). In that way, meanings and interpretations of videos are foregrounded as in constant flux of historical and socio-cultural developments and chains of significations that preceded any particular video moment. Research Object is also a novel element in an analytical video coding. It helps researchers to focus on what their research aims and questions are, in order to

move interpretations of the observed video to the next stage of research-focused interpretation. In that way, IG is unique in having three levels of interpretations. This triadic-level format has four layers as mentioned earlier in relation to Peircean edusemiotic: 1) the “rhemic” quality of Representamen, the “dicentric” quality of 2) Denotative Interpretant and 3) Connotative Interpretant, and the “argument” quality of 4) Research Object. While the IG approach contains useful elements of multimodal analysis at the level of micro analysis, it can also be scaled up to a more meso level (via focused Action or focus on particular Elements and Composition). Importantly, its analytical focus does not just privilege action, but what elements this action consists of. IG unpacks how the individual meanings of those elements – de-contextualised and contextualised can inform Research Object and the other way round. This brings a renewed understanding of a video, its signification and video analysis. The analysis clearly shows the interpretative nature of all research when Research Object (the positioning of the research and the researcher) frames the observed and its layers in particular ways to draw *associative* research insights. This quality of associative insight brings another unique feature of IG. It offers an incremental process of inferences from more generally recognised entity (Representamen, say, a person who looks like a woman, which the majority of interpreters (in one socio-cultural milieu) would agree is that - a person who looks like a woman) to least generally recognised entity (Research Object – always specific to particular research). To highlight, as compared to other video analysis models and approaches, the following features are distinct to IG:

- the meaning of individual represented elements (Representamen) plays an important part in understanding how the whole meaning of the video is made, and how different elements connect to and contribute to this meaning.
- different levels of signification (e.g. Denotation and Connotation) are distinguished in order to pay greater attention to socio-cultural and historical meanings of artefacts, as well as how the context influences the meaning of an action, the material elements and environment that action is embedded within. Simply put, what majority of interpreters with similar socio-cultural background would see is closer to Representamen (albeit at some core level what humans see with the sense of vision are shapes and the quality of colour and other qualitative sensations that informs vision). Then, interpretation starts becoming more and more specific, reaching a narrowly defined research scope in Research Object. Such analytical layering allows for discovering where and when in the process, if the framework is used with multiple interpreters, interpretations start to meander and differ, how this happens and why.
- the triadic associative interpretation and logic (Peirce’s rheme, dicent, and argument type signs) works from most general to least general analytical entity. IG directs the rhemic meaning of individual elements of the video (Representamen) to their dicentric meaning (Interpretant) all the way to the argument signified by the video as a whole within the frame of particular research questions (Research Object). This provides for insights and meaning dimensions of the observed via *associative thinking*, to link the signification of the video action and elements with research questions, research theory and/or conceptual framework and aims.

IG would not take more or less time for analysis than any other video analysis approach. The level of analysis depends on the researcher and their decision with regards to what level of element identification and action focus they want to work with. Someone might decide that they only want to focus on the elements of hands and eyes in a video, or on some salient action focus points. Time allowance and the scope of research will determine the level of analysis (more micro or more meso).

Discussion

In attempting to offer analytical models, one has to be mindful of possible challenges and model weaknesses, aside proposed benefits. I am aware that adapting one model into another is never straightforward. Proposing codes and definitions necessarily creates possibilities to challenge them. Peirce's triadic model of meaning making is adapted here to serve the purpose of *supporting analytical focus in the context of analysing educational interactions and environments where social and material are entwined*. It is not constructed to scale down Peirce's elaborate and profound philosophy to his triadic model, but offer an analytical approach that focuses on his semiotic meaning making triad.

Furthermore, multimodality and semiotics have been criticised as largely depending on the interpretational aptness and "verbosity" of the researcher-interpreter, and as individual interpretations that do not account for more interpretative voices. To address this challenge, researchers can work in groups and the IG model can serve research where multiple voices are included. In addition, Jewitt (2012, 8) lists further disadvantages of video in research, such as: "video data is partial: it includes and excludes elements; it usually provides one perspective on an event; it takes time to watch and review and can be difficult to be meaningfully summarized". In relation to further disadvantages of visual data in Higher education, Gourley (2016) argues that the weakness of the visual is due to it being commonly seen as more ambiguous than language interpretation. The IG model proposed here provides systematic support in visual and video interpretation, and involves speech as an integral part of analysing video in educational research. However, the focus is more on the pictorial, as the pictoriality of videos has been given less attention. It is for the research community to further apply, unpack, critique and adapt the model or link it to other models and theories.

Using particular terms can connect the model to some debates, for example, the ones surrounding the "sociomateriality" approach (Orlikowski, 2007; Leonardi et al., 2012). To open up this debate is a risky business. I can reflect on it briefly, since I use the word material and materiality here. The present IG model challenges the ontology of "separateness" between concept/society/culture and matter /materiality in the development of action and action systems, in the spirit of Peircean semiotics. However, this ontological oneness is *understood/interpreted by humans*. I do not literally *see* by using my vision how a mobile phone and a human are one entity (although the physical contact between the hand and mobile phone is certainly visible), but I can *know* and *conceptualise* that they are. To be precise, I have "trained" my eye to "see" social in material analytically via my mind – in that way my eye and my mind are intrinsically connected. Therefore, material affordances are useful to be identified (via an interpretation that involves both vision and mind) in order to understand how they come together with their situated socio-cultural meaning.

Technology and humans can have different compositional material “affordances” (humans have flesh and blood, technology does not, literally speaking, at least not of the same kind), but these affordances can and do get together to form hybrid (or entwined) and complex practices. I stop here since to develop these ideas further is already the scope of yet another paper. For the sake of clarity, I mention that what I mean by materiality in this model is *focusing analytical attention* and interpretation on the sense of vision by identifying any perceived object or human and contemplate its *distinct material characteristics*. Socio-cultural is not separate from it. Interpretation per se is always subjective, social and cultural, but in the proposed analysis, the social (cultural, historical, experiential) focus expands/exponentially grows from identified Representamen to Object, culminating into critical inferences via Research Object.

Further questions can be raised and discussed about why it is hard for multimodality and semiotics to penetrate higher education studies. Answers can be many and varied too, such as the slow pace of change in entrenched traditions and the higher analytical and research design demand posed by multimodality methods (as compared to for example a focus group or an interview). If the time is an issue in research, decision to go for more traditional and established approaches might win, since breaking new ground would take more time, but it could be more rewarding and bring novel and exciting insights. Educational research mostly builds on the traditions, theories and methods of longstanding, mainstream Psychology, Sociology, and Linguistics. More traditional methods can be positioned in opposition to the so-called “art-based methods, artistic endeavours, creative practice, visual methods, for visual learners” and so on. But this branding might be a symptom of a more patronising and exclusive, rather than endorsing attitude, in a sense that mainstream Psychology or Sociology studies are positioned as “serious business”, and there is all this “creative stuff”. Of course, any researcher in these fields would know that all disciplines have applied a variety of visual and “creative” methods, including Sociology, Psychology and so on. All research is creative and some research includes the visuality and multimodality of the world and education. Another reason for a slow pace of adoption might be the relatively new and recent status of multimodality, including new and multiliteracies, and most recently edusemiotics, albeit the field of semiotics has been around for quite a while. In addition, a lack of “training” and support when it comes to such new and emerging methods poses a challenge, but this can be resolved. Future research is needed to develop theoretical positions surrounding multimodality in Education, which has been only introduced here via edusemiotics and Peirce’s reasoning triad. Such research can consider the positioning of the IG as a method in relation to a number of approaches and theories that seem to promote a lot of similar but also rather different core views, such as: Activity Theory (Harter and Otrell-Cass, 2017) and Sociomaterialism (Orlikowski, 2007), including the approaches and debates surrounding the Post-anthropocene (Wallin, 2017) and Post-truth (Peters et al., 2018).

The IG coding can be applied to analyse videos beyond educational research. Educational research and Studies in Higher Education are cross-disciplinary fields that tackle education across disciplines. Researchers in any discipline who are interested in capturing different environments to understand those environments better and what happens within them can apply the IG approach to analyse the observed video data led by the Research Object pertinent to their area of interest, either disciplinary or cross-disciplinary. For

example, Sociologists can find IG useful for considering and unpacking renewed sociological relationships between artefacts, actors, action and their context. Psychologist can unpack further the embodied cognition of human action and its link to material environment. Researchers interested in the use of digital technologies can observe and understand the complex interpretative signification of how animate beings such as humans use and relate to any type of technology. Researchers working in the area of language learning can observe how language learning is entwined with contextualised signification of artefacts, movements, and signification in/through action and artefact mediation, linguist articulation and entire interpretative repertoire of the world that language learners carry around. Historians might want to apply IG approach to analyse and understand the meanings of historical places or objects. Researchers working in prominently tactile and visual fields such as Medicine or Engineering can explore various situations and engagement which require constant interpretation in conjunction between actors and material surroundings. To understand how environments and actions in those environments make meaning is one of the main goals of international research across disciplines, not only Educational research (what significations are at play in any given educational context).

If subjectivity in qualitative research is considered, every qualitative analysis is subjective and does not imply any universal interpretation, be it theme, discourse, phenomenographic or any other qualitative analysis, hence this subjectivity is present in this approach too. The researcher would endeavour to be consistent in their interpretation. However, the present codes can help a group of researchers check the reliability of some individual coding, and thus provide what is deemed to be some level of reliability in qualitative coding.

Last but not least, ethical procedure and questions of exclusion need to be carefully addressed in video research. This has been mentioned earlier, but deserves further emphasis. For example, obtaining consent to be video-recorded and handling video data in terms of anonymity and access is a challenging area in video research, but it could be carried out successfully with caution and full engagement with the ethical requirements. Furthermore, the methods that focus on visual senses would exclude visually impaired and blind participants. It is therefore the researcher's duty to think about and provide alternative methodological provisions for that community.

Conclusion

This article introduced an "Inquiry Graphics" (IG) analytical approach and its interpretative coding system. The IG approach contributes to the emerging literature that links multimodality and edusemiotics, here from an analytical perspective. Edusemiotics is a new field in educational philosophy, stressing the communicative, holistic, and interpretative character of education, distinctly building on Peirce's philosophy and semiotics. The presented IG approach focuses on Peirce's triadic meaning-making model. It is a multimodal Peircean edusemiotic approach for analysing and coding videos (including still video images) in Higher Education research and Studies of Higher Education. However, the approach can be appropriated for different types of research, projects and contexts, as reflected on in the article. The multimodality approach is yet to find its space in general and global Higher Education research and Studies, although times are changing and recent publications offer a promising step forward into this direction (Archer, 2010; Breuer and Archer, 2016). Time will tell how, to what extent and with what implication

the aforementioned approaches and the IG are applied in higher education research globally, especially concerning technology mediation. I hope that it will offer a valuable support to researchers in any higher education discipline and interdisciplinary teams.

Acknowledgements

I would like to thank many doctoral researchers that have participated at Lancaster University's doctoral research programme and the module I convene: thank you for helping the evolution of the analytical framework over time. I would like to thank Alin Olteanu and Michael Pearson for their support. Special thanks to Dr Olteanu for the conversations about Peirce.

Availability of data and materials

This is a methodological framework article.

Authors' contributions

I am a single author. The author read and approved the final manuscript.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 21 March 2018 Accepted: 1 June 2018

Published online: 19 June 2018

References

- Anstey, M., & Bull, G. (2006). *Teaching and Learning Multiliteracies: Changing Times, Changing Literacies*. International reading association, 800 Barksdale road, PO Box 8139, Newark, DE 19714-8139
- Archer A (2010) Multimodal texts in higher education and the implications for writing pedagogy. *Engl Educ* 44(3):201–213
- Barthes R (1973) *Mythologies*. Paladin, London
- Barthes, R. (1977). *Image, music, text*. London: Fontana Press
- Bezemer, J. (2014). Multimodal transcription: A case study. In Sigrid Norris and Carmen Daniela Maier (Eds.) *Interactions, images and texts. A Reader in Multimodality*. Berlin: De Gruyter
- Bezemer J, Kress G (2015) *Multimodality, learning and communication: a social semiotic frame*. Routledge
- Bezemer J, Mavers D (2011) Multimodal transcription as academic practice: a social semiotic perspective. *Int J Soc Res Methodol* 14(3):191–206.
- Breuer E, Archer A. (Eds.). (2016). *Multimodality in higher education*. Brill
- Buhl M (2010) Multimodality: on video mediated counselling for educational purposes. In *Lifelong Learning-Building a Learning Society*, ASEM LL Hub
- Chandler D (2017) *Semiotics: the basics*. In: Taylor & Francis
- Cope B, Kalantzis, M. (2000) *Multiliteracies: literacy learning and the design of social futures*. Psychology Press
- Freebody P, Luke A (1990) Literacies programs: debates and demands in cultural context. *Prospect: Australian Journal of TESOL* 5(7):7–16
- Goldman R, Pea R, Barron B, Derry, S. J. (2014) *Video research in the learning sciences*. Routledge
- Gourley L (2016) Multimodality, argument, and the persistence of WrittenText. In E. Breur and A. Archer (Eds) *Multimodality in*. High Educ:79–90
- Hallewell MJ, Lackovic N (2017) Do pictures 'tell' a thousand words in lectures? How lecturers vocalise photographs in their presentations. *Higher education Research & Development*:1–15
- Harter C, Otrell-Cass K (2017) Coding the complexity of ACTIVITY in video recordings: a proposal for constructing codes for video analysis using ACTIVITY theory. *Knowledge Cultures* 5(5)
- Iedema R (2003) Multimodality, resemiotization: extending the analysis of discourse as multi-semiotic practice. *Vis Commun* 2(1):29–57
- Jewitt C (2012) *An Introduction to Using Video for Research*. NCRM working paper. NCRM. (unpublished). In: <http://eprints.ncrm.ac.uk/2259/>
- Jewitt C (2014) *The Routledge handbook of multimodal analysis* (pp. 28–39), 2nd edn. Routledge, London
- Jewitt C, Bezemer J, O'Halloran K (eds) (2016) *Introducing multimodality*. Routledge
- Jewitt C, Kress G, Ogborn J, Tsatsarelis C (2001) Exploring learning through visual, actional and linguistic communication: the multimodal environment of a science classroom. *Educ Rev* 53(1):5–18
- Knobel M, Lankshear C (2006) Discussing new literacies. *Language Arts* 84(1):78
- Kress G (2009) *Multimodality: a social semiotic approach to contemporary communication*. Routledge
- Kress G, & van Leeuwen TV (2001). *Multimodal discourse: The modes and media of contemporary communication*
- Lankshear C, Knobel M (2007) Sampling "the new" in new literacies. In: Knobel M, Lankshear C (eds) *A new literacies sampler*. Peter Lang, NY, pp 1–24
- Laurier E (2014) The graphic transcript: poaching comic book grammar for inscribing the visual, spatial and temporal aspects of action. *Geography Compass* 8(4):235–248
- Leonardi PM, Nardi BA, Kallinikos J (2012) *Materiality and organizing: social interaction in a technological world*. In: Oxford university press on demand

- Martinelli D (2010) *A Critical Companion to Zoosemiotics: People, Paths, Ideas* (Vol. 5). In: Springer Science & Business Media
- Metcalfe, A. S. (2015). Visual methods in High Educ Research in the College Context: Approaches and Methods, 111
- Nagro SA, Cornelius KE (2013) Evaluating the evidence base of video analysis: a special education teacher development tool. *Teacher Education and Special Education* 36(4):312–329
- New London Group (1996) A pedagogy of multiliteracies: designing social futures. *Harv Educ Rev* 66:60–92
- Norris S (2004) *Analyzing multimodal interaction: a methodological framework*. Routledge
- Noth W (1995) *Handbook of semiotics (advances in semiotics)*. Indiana University Press, Bloomington
- Nöth, W. (2014). Signs as Educators: Peircean insights. In *Pedagogy and edusemiotics* (pp. 7–18). Sense Publishers, Rotterdam
- Nöth W (2014b) The semiotics of learning new words. *J Philos Educ* 48(3):446–456
- O'Halloran K, Tan S, Smith B, Podlasov A (2011) Multimodal analysis within an interactive software environment: critical discourse perspectives. *Critical Discourse Studies* 8(2):109–125
- Olteanu A (2014) The semiotic evolution of education. *J Philos Educ* 48(3):457–473
- Olteanu A (2015) Philosophy of education in the semiotics of Charles Peirce: a cosmology of learning and loving. Peter Lang, Oxford
- Olteanu A (2016) Review of edusemiotics. *Soc Semiot* 26(5):582–586
- Olteanu A, Campbell C (2018) A short introduction to Edusemiotics. *Chinese Semiotic Studies* 14(2):245–260
- Orlikowski WJ (2007) Sociomaterial practices: exploring technology at work. *Organ Stud* 28(9):1435–1448
- Otrell-Cass K (2018) Reflections on algorithmic thinking for video analysis: sorting out complex human activities. *Knowledge Cultures*
- Peirce CS (1974) *Collected papers of Charles Sanders Peirce*, vol 5. Harvard University Press
- Peirce CS (1991) In: Hooks J (ed) *Peirce on signs: Writings on semiotic*. UNC Press Books
- Peters MA, Rider S, Hyvönen M, Besley, T. (Eds.). (2018) *Post-truth, fake news: viral modernity & higher education*. Springer
- Pietarinen A-V (2006) *Signs of logic*. Springer, Dordrecht
- Pikkarainen E (2014) Competence as a key concept of educational theory: a semiotic point of view. *J Philos Educ* 62(1):621–636
- Sakr M, Jewitt C, Price S (2014) The semiotic work of the hands in scientific enquiry. *Classroom Discourse* 5(1):51–70
- Sakr M, Jewitt C, Price S (2016) Mobile experiences of historical place: a multimodal analysis of emotional engagement. *Journal of the Learning Sciences* 25(1):51–92
- Schieble M, Vetter A, Meacham M (2015) A discourse analytic approach to video analysis of teaching: aligning desired identities with practice. *J Teach Educ* 66(3):245–260
- Sebeok T (1991) *A Sign is Just a Sign: Advances in Semiotics*. Bloomington and Indianapolis. Indiana University Press
- Sebeok, T. (2001 [1994]). *Signs: An Introduction to Semiotics*. Toronto: University of Toronto Press
- Semetsky I (2005) Peirce's semiotics, subdoxastic aboutness, and the paradox of inquiry. *Educ Philos and Theory* 37(2): 227–238
- Semetsky I (2017) Edusemiotics to date, an introduction of. *Encyclopedia of Educ Philos and Theory*:1–6
- Semetsky, I., & Stables, A. (Eds.). (2014). *Pedagogy and edusemiotics: Theoretical challenges/practical opportunities*(Vol. 62). Springer
- Stables A, Semetsky I (2014) Edusemiotics: semiotic philosophy as educational foundation. Routledge
- Stjernfelt F (2006) The semiotic body. A semiotic concept of embodiment? In: Noth W (ed) *Semiotic Bodies, Aesthetic Embodiments, and Cyberbodies*. University Press, Kassel, pp 13–48
- Stjernfelt F (2011) Signs conveying information. On the Range of Peirce's Notion of Propositions: Dicsigns, *International Journal of Signs and Semiotic Systems* 1(2):40–52
- Thesen L (2016) The past in the present: modes, gaze and changing communicative practices in lectures. *Multimodality in Higher education* 31
- Thibault, P. (2000) 'The Multimodal Transcription of a Television Advertisement: Theory and Practice', in A.P. Baldry (ed.) *Multimodality and Multimediality in the Distance Learning Age*, pp. 311–385. Campobasso: Palladino Editore
- Zhao S, Djonov E, van Leeuwen T (2014) Semiotic technology and practice: a multimodal social semiotic approach to PowerPoint. *Text & Talk* 34(3):349–375

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
