



Reactivity beyond contamination. An integrative literature review of video studies in educational research

Leif Christian Lahn & Kirsti Klette

To cite this article: Leif Christian Lahn & Kirsti Klette (2022): Reactivity beyond contamination. An integrative literature review of video studies in educational research, International Journal of Research & Method in Education, DOI: [10.1080/1743727X.2022.2094356](https://doi.org/10.1080/1743727X.2022.2094356)

To link to this article: <https://doi.org/10.1080/1743727X.2022.2094356>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 01 Jul 2022.



Submit your article to this journal [↗](#)



Article views: 67



View related articles [↗](#)



View Crossmark data [↗](#)

Reactivity beyond contamination. An integrative literature review of video studies in educational research

Leif Christian Lahn ^a and Kirsti Klette^b

^aDepartment of Education, University of Oslo, Oslo, Norway; ^bDepartment of Teacher Education and School Research, University of Oslo, Oslo, Norway

ABSTRACT

The growing interest in video research and new technologies for recording human interaction has stirred debates about intrusiveness and 'reactivity' understood as researcher-derived changes in subjects. In addition to a plethora of concepts referring to such effects in extant literature, different ontological and epistemological positions provide contrasting frameworks for interpreting and deciding on methodological guidelines. In this article we discuss these elements, that we have called 'meta-methodological', from the standpoints of experimental research, social-constructivism and scientific realism. We combine conceptual analysis and a literature review of video-studies in teaching in order to identify both possible traces of contesting beliefs and to provide a glance at different aspects of 'reactivity' that needs to be systematized in the ongoing debates. Whereas the methodological literature underline the importance of such effects, these are rarely reported in the reviewed video studies. Moreover, reactivity is seen as a minor problem in the latter, and we found few instances that validated the effects on the field and on the empirical conclusions. Our article ask for more transparency in field researchers' judgment about reactivity and mitigating measures.

ARTICLE HISTORY

Received 18 March 2021
Accepted 26 April 2022

KEYWORDS

Reactivity; video studies; integrative literature review; meta-methodology; research on teaching and learning

Introduction

The development of video technology in recent decades has changed how we communicate and learn (Xu and Clarke 2019) and has provided unprecedented opportunities for research by enabling the capture of rich data on complex social interactions over time (Klette 2009, Derry *et al.* 2010, Heath *et al.* 2010). A common but often implicit assumption by researchers is that the observed events are natural as if the presence of an observer or a video camera makes no difference. An apparently competing view would be that this presence 'contaminate' a pure social environment under study (Kazdin 1982).

Different versions of these assumptions appear in the social and educational sciences literature on reactivity or observer effects. This theme boomed during the 1980s, then declined, until a recent period of several newer contributions promises to be a new turn. However, important reviews are somewhat entrenched in specific methodological paradigms, for example, Hazel (2016) in social constructivism and Praetorius *et al.* (2017) in experimental logic. There is a need to broaden the framework and introduce contributions that we have labelled 'meta-methodological' positions (Newman 2010): A naturalist with standards from the natural sciences, a social constructivist, a realist approach

CONTACT Leif Christian Lahn  l.c.lahn@iped.uio.no  Department of Education, University of Oslo, Oslo, Norway

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

and a mixed view referred to as interventional or transformational. Given this framing, our focus on video research and teaching quality is pertinent since it includes both small-scale (Derry *et al.* 2010) and large-scale studies (Jacobs *et al.* 2007, Praetorius and Charalambous 2018, Klette *et al.* 2021) – with research designs that cross the qualitative / quantitative divide (Seidel and Thiel 2017). The selection of video study designs is also justified by their growing popularity in studies of social interaction and methodologists' renewed interest in such aspects of videography etc. (Heath *et al.* 2010, Knoblauch and Tuma 2011). In addition to the general considerations in doing field observations, the video and audio recording layouts provide a variety of mitigating possibilities in terms of remote observation, self-recording etc. (MacMahon *et al.* 2019). As 'natural data' (Knoblauch *et al.* 2009) the captured classroom activities could be archived for specific reanalysis of observer effects (Andersson and Sørvik 2013). This feature of turning private realities into more public may also be a reactive factor (Aarsand 2016).

The methodology literature abounds with terms that has some family resemblance with the term 'reactivity', such as 'observer effect', 'camera effect', 'expectancy effect' (Rosenthal 1966), 'demand characteristics' (Orne 1962), 'Hawthorne effect' (McCambridge *et al.* 2014b) and others (Jimenez-Buedo 2021). Their common reference is a change in social behaviour attributed to the presence of observations, an observer and / or participation in research. As a point of departure we stick to a generic understanding of 'reactivity' or 'research participation effects' (McCambridge *et al.* 2014a), but below we provide an overview of some key distinctions in different fields of research. We needed to explore this vocabulary from discourses on research methodology when designing the protocol for our literature review on video studies in classrooms. A differentiation of terms partly reflects the above contrasting epistemological perspectives and research designs. In our review we have asked how these frameworks influence the researchers' sensitivity to reactivity, the reporting of such effects, and their efforts in finding evidence for the occurrence of reactivity. The guidelines for preventing or mitigating those processes may likewise depend on the adopted research strategies.

Methodologists such as Patton (2015) would advise researchers to be attentive to reactivity effects and include them as caveats in their field notes (Monahan and Fisher 2010). Others ask for a stronger commitment to a systematic inquiry that may strengthen an understanding of the process and help mitigation (Spano 2006). A common finding in video studies reporting reactivity, is that the effect fades out quickly (Blikstad-Balas 2017). In our literature review we will have a focus on a number of conditional factors that may elicit and sustain the researcher-induced changes: For example, the use of fixed versus mobile cameras, social and personal characteristics of the observer, the gravity of the situation. To what extent does these factors influence groups such as younger and older children, differentially? An answer to this and similar questions may provide tentative guidelines for minimizing effects that potentially invalidate the conclusions drawn from the study. One of the themes covered in our review, is measures taken by researchers to prevent and mitigate reactivity, for example, making the recording less visible or familiarize the class familiar with the observer and the camera.

There is no simple relationship between the use of specific data collection methods, such as video observation, and reactive effects. Thus efforts to understand such processes need to be done empirically. A step further is then to ask to what extent and how researchers justify that empirical inferences are not biased by reactivity. Not all departure from naturalness may be consequential for the analysis of data, and steps taken to minimize intrusiveness may be unnecessary from an epistemic point of view. Jimenez-Buedo's (2021, p. 14) makes a distinction between benign reactivity and malignant reactivity. In our interpretation, the former refers to observer effects that do not influence the researchers' conclusions from the field material. The latter occurs when reactivity impacts and biases such inferences. Our questions then are to what extent video studies make explicit issues of reactivity and if so, what evidence is provided for making conclusions about reactivity effects, and further – about consequential biases in the interpretation and explanation of the field material. We will call for a reappraisal of triangulation designs as an instrument for making such inferences transparent.

A framework for integrative literature reviews and a conceptual clarification of 'reactivity'

In this section, we will present a way of combining literature reviews and conceptual analysis and then proceed to a discussion of key terminological problems associated with the use of 'reactivity.' As pointed out above, a basic clarification of defining features needs to consider the influence from meta-methodological positions. Our framework on these issues is mainly extracted from the methodological literature on 'reactivity'.

A short note on Torraco's integrative literature review

Torraco (2016) makes a distinction between two types of integrative literature reviews – ones that address mature topics and ones that address emerging topics. Our version belongs to the second category which, according to the author, would 'benefit from a holistic conceptualization and synthesis of the literature' (410) that should generate a preliminary conceptualization of the topic. Our inquiry is aligned with the principles of 'theoretical sampling' (Wolfswinkel *et al.* 2013) and a configurational logic of systematic reviews (Newman and Gough 2020) in subjecting the procedure to an iterative combination of conceptual analysis and literature searches where the qualification criteria (for inclusion) are revised and reframed during the review process. Such an approach and structure will provide coherence and clarity about the relationship between the main concepts of the topic. This approach to literature reviews also puts a strong emphasis on synthesis, where concepts and perspectives are recast, combined, reorganized and integrated (Torraco 2016, p. 420). Tentative models need to be justified through the explication of the underlying logic, such as meta-methodological foundations.

This approach to literature studies bears a resemblance to conceptual analysis as an empirical method (Soltis 1978, Levering 2002). In framing our review, we have only made cursory visits to such maxims for conceptual clarification and synthesis. We do not engage in a historical excavation of the term 'reactivity' (Vogelsang 2012), even though a literature span of 50–60 years has been chosen for a rough tracing of trends. Its metaphorical origins in chemistry and neurology and a parallel terminological lineage in psychology ('emotional reactivity', Shapero *et al.* 2016) are not included, thereby indicating that the focus here is on 'observer reactivity' (Baum *et al.* 1979) as an instance of the more general 'social reactivity'/'social facilitation' (Steinmetz and Pfattheicher 2017). Our approach to 'holistic conceptualization' will include three principles derived from Soltis' (1978) systematic version of conceptual analysis: (1) generic analysis, which consists of inquiries into the essential features of concepts often defined by counterexamples with other terms, such as 'reactivity' versus 'reflexivity'; (2) differential analysis of the uses of concepts as a basis for clustering and typologies, such as fixed versus mobile recording technology and (3) conditional analysis, focusing on the context of use, for example, if 'reactivity' is identified in the early phases of a field work as a 'novelty effect' (Bracht and Glass 1968) versus later as a 'habituation effect' (Redman *et al.* 1989).

The many faces of 'reactivity' in field and video observations

We take field observation to be a very pertinent case when addressing issues of reactivity. As succinctly pointed out by Adler and Adler (1994, p. 378); 'The naturalness of the observer role, coupled with its non-direction, makes it the least noticeably intrusive of all research techniques'. Field researchers may be extensively blind in taking it for granted that they describe, when in fact they intervene. Such instances made the sociolinguist William Labov coin the 'observer's paradox' when he realized that members of English worker communities turned natural speech into a formal genre in the presence of the researcher. Naturalness could only be achieved by being absent (Labov 1972, p. 209). There is, however, more to reactivity than the presence and absence

of an observer. Based on assumptions about underlying mechanisms, these could be clustered into factors related to observers' presence, participants' beliefs about research purpose etc., and social norms about desirability in field observation. In addition, we distinguish between such effects in research with a mainly descriptive goal, and research that has an interventional design and is explicitly aiming at changing the object of research. Given the variety of sources for reactivity, it is not surprising that terms referring to such phenomena, have proliferated in the literature.

The term 'observer effect' is often used synonymously with 'reactivity', but authors in the qualitative methodology tradition may prefer the former given the latter's origins in psychological laboratory-based research (Knoblauch *et al.* 2009). Their focus is then departures from naturalness as a function of the observers' presence. It is intrinsically meshed up with the basic methodological concerns when doing field work; to be participant or non-participant, close-distant, active-passive, new-familiar, hidden agenda-open agenda, etc. Such precautions may be justified by reference to material factors such as the layout of video recording instruments or to elements of socio-cognitive elicitation expounded in the following well-known concepts. The (observer) 'expectancy effect' (Rosenthal 1966) is identified when participants' behaviour change as a result of expectations conveyed inadvertently by the observer or experimenter, whereas 'demand characteristics of experimentation' (Orne 1962) refers more clearly to participants' adjustments to their perceived role in research and assumptions about the observation / experimentation. Reviews of literature on the 'Hawthorne effect' (McCambridge *et al.* 2014b) conclude that this term should be reserved for biasing changes that are attributable to the mere presence of external experts.

These and similar concepts underlining various aspects of social desirability, may be understood as bringing to the fore potential threat to ecological validity in observational studies (Heath *et al.* 2010). However, the more voluminous literature on experimental studies provide a different explanation of confounding effects (McCambridge *et al.* 2014b). Since the interventional design is meant to enable inferences about the 'true effect' of experimentation, any contributions by the subjects to construe or complete this effect is a potentially biasing factor (Jimenez-Buedo 2021). To what extent this type of reactivity represents a less or more conspicuous processes compared to the loss of naturalness in field observation, has to our knowledge not been addressed in the methodological literature on experimental designs or field interventions.

To repeat, we will use 'reactivity' in a generic way that includes the variants discussed above, such as the expectancy and the Hawthorne effects. With some hesitation we have also decided to keep it in light of the meta-methodological skirmishes that we present and discuss in the next section.

Theoretical and meta-methodological approaches to studies of 'reactivity'

Since video data generate a material for both quantitative and qualitative studies, the literature of video-based social research is split into more or less clearly defined research paradigms (Knoblauch *et al.* 2018) that are also echoed in contrasting views on reactivity and its severity for empirical research. In this area, four positions could be identified: (1) an eliminative stance rooted in experimental psychology that treats reactivity as a 'contaminating' effect by representing a possible threat to the validity of the studies, (2) a social-constructivist position that sees such influence as unavoidable and contingent on local factors or 'resources,' (3) a loosely-defined realist approach that urges researchers to investigate if their presence makes a difference to the participants and (4) research strategies that have a transformative or interventional purpose and are deliberately making a change in their study object, which as pointed out above, may raise specific issues when it comes to 'reactivity'.

Reactivity and contamination effects

Experimental studies of reactivity in educational contexts had several booming decades in the 1960s, 1970s and 1980s (Sechrest 1982, Blease 1983, Praetorius *et al.* 2017) and were summarized as a

critique of research that asked teachers to retrospectively report on their behaviour during the presence of a non-participant observer (Harris and Lahey 1982). In addition, research designs with systematic alternation between periods of observation and non-observation gained ground – inspired by the principles of operant analysis and an ABAB-logic (baseline-intervention-baseline-intervention; Kazdin 1982). Video cameras were used in deception studies as an alternative to one-way windows by having them installed non-obtrusively in the classroom for extended periods – letting the teachers believe the video only recorded with the observer present (Mercators and Craighead 1974). Such studies generated inconclusive results when it comes to reactivity (Blease 1983). Weinrott *et al.* (1978, p. 909) suggested that findings could be attributed to ‘observation sophisticated’ subjects who suspected data would be gathered covertly during the observer-absent condition, but were relieved when the observers turned the situation into a more normal one, thus minimizing reactivity. Although Kazdin (1982) advocated the use of unobtrusive measures and special apparatus to reduce the likelihood of reactive effects, more recently, authors have questioned the naturalness and underlined the ethical issues with surreptitious strategies (Lee 2000, Kucirkova and Falloon 2018, Kleckner *et al.* 2020).

The miniaturization of video technology has made this type of recording less conspicuous (Muench *et al.* 2013), and recent advances in mobile eye-tracking instruments have introduced an alternative design for recording visual data of classroom interaction (Webb *et al.* 2000). The claim is made that automatic behaviour such as gaze is likely to be controlled by the targeted task and to a lesser extent by voluntary alignment with social demands (Goodwin and Velicer 2008). Although teachers and students reported some discomfort wearing ‘eye-tracking’ glasses (Haataja *et al.* 2019), their attention on the technology very quickly faded (for teachers a couple of minutes, Praetorius *et al.* 2017) and switched to the tasks at hand (McIntyre *et al.* 2020). This method for tracing the attention of classroom participants is described as a minimally distracting design (Magnussen *et al.* 2017). However, such body-based measures are usually correlated with more intrusive data collection procedures such as video recording (Prieto *et al.* 2018) and interviews/questionnaires (Haataja *et al.* 2019).

Praetorius *et al.* (2017), in a systematic investigation of reactivity in video-based classroom research, point to a methodological dilemma in ‘achieving the two intended experimental conditions’ observed/not observed (54) given the multidimensionalities of teaching. Their analyses and summary of large-scale studies (TIMSS Video and IPN Video Studies) indicated that such effects occurred more frequently with respect to teachers’ and students’ emotions compared with changes in behaviour. However, the authors provide clear evidence for a process of habituation, but not without exceptions. The earlier (Harris and Lahey 1982, Kazdin 1982) and more recent literature (McCambridge *et al.* 2014b, Praetorius *et al.* 2017) in this tradition have listed a number of contingencies that affect reactivity or ‘observer effect’: situational variables that generate specific demands on the observers such as assessment or activities for social integration, prior interactions between observer and participants, conspicuous character of the observation and individual differences between the subjects. Although an underlying logic of the experimental paradigm should be to go beyond reactivity as a contamination factor by substantiating how it represent a validity threat, there is scant research on the latter.

Video-elicited reactivity and ‘contamination’ as reconfiguration and local resources

During the last decade scholars identifying with varieties of social-constructivism, such as ethnomethodology/conversation analysis (Hazel 2016), praxeology (Mondada 2006) and ethnography/videography (Knoblauch and Tuma 2011), have radically redefined the eliminative concept of reactivity. Contaminating effects imposed by researchers on participants are understood as ‘resources’ to which subjects may orient and ‘stage performances’ (Speer and Hutchby 2003a) that deeply reveal how people perceive themselves and how they want to be perceived (Monahan and Fisher 2010). This dramaturgical metaphor redefines the video camera as a configuring device (Mondada 2006,

p. 3) that may elicit a reframing of the situation by the observees. They may strive to appear normal or use the opportunity to be on stage and present themselves in a socially desirable manner (Heath *et al.* 2010). Research inspired by social-constructivist thinking has provided a number of studies of family conversations (Aarsand and Forsberg 2010, Gordon 2010, Hazel 2016), classroom interaction (Bhatt 2017, Sahlström *et al.* 2019) and professional collaboration (Spano 2006) that display how couples, students and police officers, respectively, use the data collection instruments to create rich repertoires of both playful and non-playful communicational patterns, displaying distinctive identities (Speer and Hutchby 2003a).

Followers of conversation and interaction analysis (Jordan and Henderson 1995) take a critical stance at the concept of (participant) reactivity, which they claim represents a belief in a pristine and natural realm uncontaminated by research (Speer and Hutchby 2003a, Hazel 2016). Instead, they topicalize how the interaction between participants and researchers are co-produced (Mondada 2014). Reflexivity is constituted by the participants' retrospective verbal and non-verbal reactions to this ongoing communication (MacBeth 2001), and not by the researchers' methodological maneuvers to uncover their effects on the field. This strategy makes no a priori assumptions about specific consequences of for example video-recording on participants' behaviour in terms of social desirability effects etc. (Heath *et al.* 2010, p. 47). Thus it may provide detailed accounts of how subjects use recording devices as resources in their interactions with researchers, as shown in the previous section. However, this tradition of micro-level discourses is in principle skeptical of making generalizations about underlying patterns from case studies of co-produced communication (Hazel 2016). A related issue is to what extent conversation analytic researchers that do not explicitly topicalize the researcher-participant interaction in classrooms, make explicit their influence on the field. Are observed and reported discourses taken at face value?

Multi-method and realist approaches to studies of reactivity

In a debate with Speer and Hutchby (2003a, 2003b) Martyn Hammersley contested the rephrasing of reactivity as a 'staged performance' when participants are observed by researchers (Hammersley 2003). The latter acknowledges that such a focus may be relevant for specific empirical investigations, in line with our comment above. However, he adds that, in general, claims about the social world need to be studied by minimizing the effect of the researcher on the data and by investigating factors that produce challenges with participant reactivity. At the core lies a concern for '... identifying what are typical or habitual responses on the part of people' (Hammersley 2018, p. 14). This strategy Hammersley associates with ontological postulates about the nature of the world independent of any constructions of it. Researchers should concern themselves with the accuracy of informants' accounts beyond how such data are constructed. This includes a systematic investigation of possible reactive effects. Hammersley (2003, p. 344) contrasts such a position in scientific realism to eliminative approaches to reactivity in terms of same stimulus control (the positivist view).

Under the heading of 'reflexivity' Hammersley (2003) insists on the imperative of monitoring the effects of the research activity on what is observed. Similar views are articulated by Erickson (2006), scholars in micro-ethnography (Heath *et al.* 2010, Haw and Hadfield 2011), multimodal discourse analysis (Jewitt 2012) and proponents of realist methodological platforms (Nassauer and Legewie 2019). More precisely this literature advocates methodological designs that triangulate data and methods in a strategic way and are supported by provisional knowledge of potential reactive effects. Will the research subjects' self-reports provide good evidence? Should the observer stay longer in order to investigate processes of familiarization, habituation, fading etc.? Or is it advisable to choose situations with different degrees of demand characteristics and evaluative elements? However, such approaches share with Hammersley's (2011) subtle approach to realism a belief in fallibilism in that empirical knowledge cannot be proven with certainty and is transient and subject to pragmatic concerns.

Zahle (2019) urges qualitative researchers to make 'reactivity assumptions' based on data transparency and the need to take into account the effects of research intrusiveness on inferences drawn from the data. The rationale being that no method is a priori more prone to generate reactivity. In line with realist interpretation of methodological triangulation (Hammersley 2008, Maxwell 2012, Patton 2015) the tenets of convergent confirmation are upheld since they reduce the risk that findings are reflective of systematic errors associated with each method (Zahle 2019). However, Zahle adds that the validation of the research questions should be paralleled by deliberations about the severity of researcher-induced changes to the field. The relevant factors could include whether the distortions are consequential for the justification of evidence, and if so, to what extent the design does provide data for assessing their epistemic effects. Again we need to insert that the version of scientific realism that has gained support in the debates addressed here, is fallibilistic in maintaining that our theories cannot be rationally supported in a conclusive way. They are only approximately true (Niiniluoto 2003), and scientific explanations should contemplate counterfactual evidence and assumptions about relevance, coherence, reactivity etc. (Psillos 2011).

Reactivity in intervention studies and transformative research

As pointed out above, we include the different concepts referring to experimenter effects in the generic term 'reactivity'. However, we have also suggested that intervention studies in social contexts, ranging from experimental studies to action research (Ilin *et al.* 2013) or design-based research (Anderson and Shattuck 2012), may represent specific issues in terms of transparency, mitigating measures and validity threats compared with descriptive research. In the quantitative research tradition, 'naturalness' (uncontaminated by the researcher) is a methodological concern when identifying the baseline conditions in pre- and post-test logic (Baldassari and Abascal 2017). Interpretive positions tend not to be so explicit about these stages when advocating 'multi-layered reflexivity' (Nolan *et al.* 2018), 'relational reflexivity' (Collier and Wyer 2016) and memoing as ways of heightening the researchers' awareness of their presence and influence in the field. A major challenge to all variants of intervention and transformative research is to disentangle the differential influence of the research project, the intervention, the researchers' presence, the involvement of the participants, the naturalness of the site, etc. In the experimental tradition, the design protocol, implementation designs (Ogden and Fixsen 2014) and fidelity data are meant to be effective instruments that facilitate an 'isolation' of field effects attributable to the research activities. In qualitative studies, various techniques for generating an analytical distance to the transformation are advocated (Gilbert 2002). However, a complicating factor is when the interventional research is understood and designed as 'participative' or 'collaborative' (Bergold and Thomas 2012). In many cases, video recording in classrooms activate elements of interventional designs, especially if teachers or students are involved as 'co-researchers' in the sense recommended by Flyvbjerg (2002, p. 132) when '... turning the camera literally around in your hand, or handing the camera to your informants ...'.

Jimenez-Buedo (2021) discusses the distinction introduced above between benign and malignant reactivity in the context of experimentation. A similar logic may be an appropriate dimension of qualitative researchers' reflexivity in transformative studies. In addition, we have underlined that field observations with video recording may be data collection strategies within overall frameworks that are descriptive, interventional or combinations of these. However, the question needs to be posed whether reactivity in experimental or transformative studies is more consequential and less salient than similar processes in exploratory and descriptive research. The answer seems to be highly dependent on the type of research design. On one hand 'reflexivity' in strategies associated with 'action research' and 'designed based research' may involve less clearly articulated prescriptive and normative exigencies, but a reported negligence in clarifying pre-interventional conditions and poorly developed distanciation techniques. On the other hand the literature on applied behaviour analysis in classroom management includes reactivity as a key confounding factor when the

fidelity of teachers in the implementation of interventions is assessed (Coddling *et al.* 2008, Gresham *et al.* 2017).

A summarizing note

To repeat, we use ‘reactivity’ in a generic sense even though social constructivists may associate this term with the logic of naturalness contaminated by research. There is an issue about different ways of understanding ‘reflexivity’ that in our discussion mainly involves the contrast between on one hand the constitutive type (MacBeth 2001) grounded in participants reflection on their own discourse, and on the other hand a realist understanding of reflexivity as a methodological imperative to investigate not only the influence of research on the field, but also to what extent such ‘distortions’ are consequential for the inferences drawn from data. They may be benign in not generating biases of importance to the study. In that case, less concern about mitigating measures is needed.

The different voices in the debate on ‘reactivity’ seem to agree that studies rooted in conversation analysis / ethnomethodology provide valuable insights into the ‘co-production’ of the observed field when the research goal is mainly illustrative or explorative. However, a realist ontological platform should support a triangulating design that assists in documenting research-initiated changes and their epistemic value. The social constructivists’ reluctance to attribute any a priori categorization of reactivity such as demand characteristics etc., has a methodological parallel in Zahle’s (2019) recommendation of making ‘reactivity assumptions’ explicit and subject to empirical justification.

It may be objected to our review that we have deliberately inflated the meta-methodological differences in order to make the undertheorized character of these issues more visible. In practice, researchers committed to video observation cross these types, but their methodological training may be tailored to the respective key ideas on reactivity. Possibly PhD candidates are more likely to comply with these positions than senior researchers which seems to be a pattern we have observed in a review of doctoral dissertations. In the recent study, we will investigate how often reactivity is reported in the selected literature, how it is described and what evidence is provided for reported research-initiated influence. We also investigate which measures are taken to mitigate reactivity, and to what extent the authors’ discourse on these matters reflect meta-methodological positions.

Methodological steps and data sources

As pointed out above, our conceptual discussion and the literature search are intertwined and iterative processes in our design. This outline of conceptual distinctions provides a set of keywords for the literature search and introduces dimensions that will be included in the analysis of the data. In reviewing the literature on ‘reactivity,’ we had to follow somewhat different tracks delineated by four main type of data sources: textbooks, concept-focused methodological articles, empirical studies of ‘reactivity’ and empirical studies of video recording in classrooms. In a second round of literature search, both new categories from the methodological domain and video research of social interaction outside educational contexts were introduced in order to explore contrastive evidence (confer differential analysis of concepts). These steps and samples (e.g. SET 1, 2, 3, etc.) are illustrated in Table 1.

As can be seen from the table, we distinguish between seven datasets. (SET 1): An overview sample of various literature sources in the educational sciences addressing ‘reactivity’ were retrieved by combining the above keywords with ‘teach*,’ ‘learn*’ and ‘education*’ when searching in bibliographic databases such as Web of science (TOPIC data fields), Scopus (ALL, KEYWORDS data fields), ERIC, ORIA (University of Oslo library) and Google/Google Scholar. After a merge/purge of duplicates, a total of 34 citations satisfied our criteria, of which 23 were accessible. (SET 2 core sample): A subset of SET 1 was generated by narrowing the sample to studies with video observations of interactions

Table 1. Overview of literature sources, keywords, search procedures and results.

Literature sources	Main keywords	Search procedures	Results
SET 1 Studies educational research addressing 'reactivity'	'Reactivity' 'observer effect' 'Hawthorne effect' 'obtrusive' 'camera effect' 'video effect'	Databases, duplicate check	23 cases
SET 2 Video studies student/ teachers not reporting 'reactivity'		Databases	11 cases
SET 3 Video studies student/ teachers not reporting 'reactivity'		Databases	104 cases
SET 4 Video studies social interaction with 'reactivity'		Sweep in databases, manual search of citations and journals	12 entries reviewed
SET 5 Methodological textbooks		Databases and manual search	18 cases
SET 6 Methodological articles 1st round		Databases and searches in methodology journals	6 cases
SET 7 Methodological articles 2nd round	'reflexivity' in literature on 'intervention*' etc.	Sweep in databases, manual search of citations and journals	25 entries reviewed

between teachers and students – also including preschool and teacher education/professional development. Furthermore, 'reactivity' etc. had to be covered in depth in the studies and not just addressed as a general methodological issue. Only 11 entries satisfied our criteria. (SET 3): The same criteria and search procedure as for SET 2 but without reported 'reactivity' etc. This contrastive sample included 104 references – with roughly 40% having German affiliated authors. (SET 4): A sweeping replica of the main set, SET 2, was done, but the target for observation studies was widened to include video studies of 'social interaction' such as child–parents–professionals and training in health care. This step generated an additional 12 publications. (SET5): The textbook sample was selected from key publications on methodology in social science and educational research. Database overviews of textbooks combined with manual scanning of indexes with the keywords in Table 1 yielded 18 cases before the search was discontinued since the descriptions reached a saturation point. (SET 6): Through database and specific searches in high ranked journals of methodology in the social/educational sciences and psychology, a total of 6 concept-focused articles were identified. (SET 7): A sweep in methodological literature for the social/educational sciences and psychology using the above databases and a guiding keyword 'reflexivity' in a random sample of literature on 'transformative research', 'intervention* research', 'action research' etc.; 25 entries were reviewed.

The initial segmentation of the material was done by reading the articles and chapter identified by the 'reactivity' keywords list and copying relevant passages into NVivo 12. This step provided a new set of concepts that triggered literature searches (SETS 6 and 7), but also established a conceptual framework for a structuring of the analysis and our presentation/discussion, for example 'social facilitation' in assessment contexts. It also generated categories and defined themes for a conceptual framing of the review, as presented in Figure 1.

The background factors in Figure 1 are meant to provide a profile of the studies to be reviewed. In our classification we will use the categorization of the respective authors (in vivo) except for the one on meta-methodological positions which is based on our reviews of methodological debates (see SET 6 and SET 7) about foundational issues related to philosophy of science addressing the conceptualization of 'reactivity.' The category 'reported reactivity' refers to descriptions grounded in the selected studies and includes an assessment of the evidence provided for such reporting, for example, interviews with teachers about reactive effects. As pointed out above, the effects of reactivity are often referred to as a biased observation of authentic behaviour, but also as an influence that may attenuate and fade away. A specific factor is their impact on the inferences drawn from the data (benign versus malignant effects). Conditional factors are defined as attributes of the observer, the observee and the video layout,

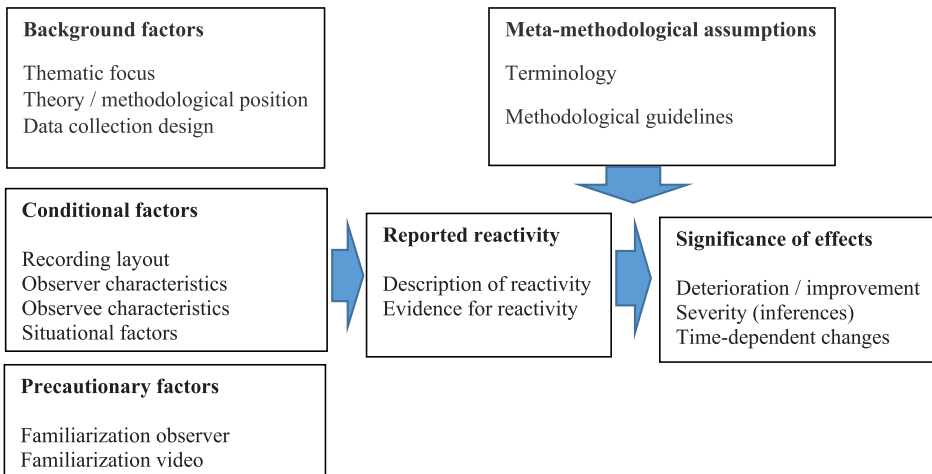


Figure 1. Factors from the literature review assumed to impact reported reactivity and effects of reactivity.

which also may interact with the research design, for example classroom studies or small group observation. In our framework, we will be looking for precautions taken by the authors to minimize this type of ‘contamination.’ The categories in [Figure 1](#) are mainly relevant when reviewing articles based on empirical studies and less so when referring to the methodological oriented contributions. They frame our main research questions: The occurrence of reactivity, and how is it researched? Do fields of research, theoretical and methodological/meta-methodological approaches have an impact on reported reactivity, and what are the effects of reactivity on participants’ behaviour? And how do they impact the interpretations and inferences made by the researcher(s)? What is the evidential basis for the reported reactivity?

Next, we will turn to literature that is apt to propose more definite answers to the issues such as: How often does reactivity occur? What are the effects on students and teachers? Do video recording or other factors make a difference? This latter question also interrogates the antecedents of reactivity in such settings and precautions taken to minimize the disturbing effects of observer and video presence. Are these measures assessed in terms of potential biasing effects and impact on the conclusions drawn from observational data?

Findings: reactivity in empirical literature on classroom video studies

In [Figure 1](#) we listed the factors and sub-items that have served as categories for our analysis of the 11 core articles on video observation in classroom settings reporting reactivity (SET 2). As pointed out above, these core entries were supplemented by texts ($n = 12$) addressing similar issues in parent–child interaction and health education (SET 4).

Background factors (profile of core articles)

Nearly all the core articles were focusing on teachers’ instruction and designed as small scale studies using multiple methods where video data provided the main empirical evidence. Since the main research approach adopted was interpretive, only two articles considered reported reactivity as a validity threat – these were also based on large scale quantitative studies. Five of the eleven entries made reference to research and methodology literature on reactivity and video recording. Three articles on teachers’ professional development combined video observations and stimulated recall (video-based) interviews. Comparing the core sample of 11 articles with classroom video studies that do not report reactivity (SET 3, $N = 104$), the latter were more often large scale quantitative

studies, authored by German scholars, and had a stronger thematic focus on ICT in science and math education.

Reported reactivity and effects

Not surprisingly, the core articles minimized the influence of the observer and camera recording. Except for one case of elicited self-defensiveness on the part of the observed pre-service teachers (Civitillo *et al.* 2019), no mention was made of the facilitating or contaminating effects of the researchers' presence. Four entries were concerned with the non-intrusiveness and habituation of wearing head mounted cameras (Blikstad-Balas and Sørvik 2015) and/or gaze tracking devices (Praetorius *et al.* 2017, Haataja *et al.* 2019, McIntyre *et al.* 2020). References in these cases to 'reactivity' were less related to observer effects than to novelty factors and the subjects getting used to the new recording instruments. In the same vein, three articles of the core and the extended sample argued for minimizing the researcher effect by 'handing the camera to your informants' (Flyvbjerg 2002). In some cases, this was achieved by having student teachers or younger children doing the recording with iPads and handing over the tapes or files to the researchers for analysis (Goh *et al.* 2019). Based on a research design where preschool children were assigned the role of recording their learning of simple electrical circuit building, Kucirkova and Falloon (2018) reported that there was no visual or verbal evidence of the children 'staging a performance,' and they quickly forgot the fact that the recorded material would be analysed.

Another study in the core sample involved remote observation with middle level and secondary education pre-service teachers using iPad recording and Wash *et al.* (2014, p. 61) reported that 81% of the eleven participating clinical students believe that their anxiety level decreased when using technology for remote observation. Several comments from the participants included that it allowed them to teach to the students and not to the observer, that it was convenient and easy. In a similar study using this observational strategy supported by iPad technology, the authors (MacMahon *et al.* 2019) remarked that the teacher students quickly forgot that they were being observed when no observer was present. Issues of remote observation and teachers, parents and students as research videographers will be addressed more in depth below.

The core articles provided evidence for reactivity or non-reactivity by interviewing teachers and students combined with direct observation and, in two cases, by noting changes in the observees over time (Blikstad-Balas and Sørvik 2015, Kucirkova and Falloon 2018). In summary, the instances of reported reactivity and associated elaborations are rather meager in the core literature (SET2) reviewed. It is noteworthy that such reporting is not initiated by unexpected validity threats to the video observation design, but rather by the authors' claim that new recording technologies attenuate any reactivity effect. An interesting exception to this pattern is provided by the extended sample (SET 4) by McCord and Matusovich (2019, p. 488) who investigated the metacognitive engagement of engineering students in self-directed learning environments. They concluded that '... the impact of the observer was minimized but never fully removed.' They observed an increase in crude language, unethical practices and less on task behaviour in a follow-up monitoring which narrowed down the specter of metacognitive strategies used by the student teams. The transient effects of a researcher as observer may thus have enabled the social production of a richer material that in the later stage of naturalness got lost. Except for a couple of articles on fidelity measures in the implementation of classroom interventions (SET 7), the issue of how reactivity occurrence may invalidate conclusions drawn from the observational data, was not addressed.

Conditional factors. In this review, we have asked to what extent the camera makes a difference, and how specific groups, younger and older children, pupils with special needs, etc. are affected. The core literature does not provide clear cut answers to the latter question. Miniaturization of recording instruments and remote control undoubtedly reduces the observer effect (Klette and Blikstad-Balas 2018). The fact that reactivity was more often reported in studies of students working in small groups with computers and iPads than research of whole class interaction indicates that the social

organization of the learning environment may be a determinant. This is also evident when the context is redesigned, which brings in the impact of novelty effects in addition to observer reactivity. Referring again to Wash *et al.* (2014), the authors argued for turning the five-year-olds into videographers by summarizing their experiences with preschool children being very mobile and the small group structure making researcher controlled recording a strenuous and unreliable task. However, if we want to make statements about age differences, a sweep into studies from health education and childhood research needs to be taken (e.g. SET 4).

Barriage and Darcey (2019) rehearse the point made by Heath *et al.* (2010) of a large variability in reported frequency, duration and ways that adult participants orient to the camera – typically at the start of the social interaction, when the observer is absent or during lulls in activity. Referring to studies of younger children, they maintained that these children displayed more attention to the camera than children over the age of 13 (Antal *et al.* 2015, Shoecraft and Fluckiger 2018) with talking to the recording devices as the most frequent effect. Glancing at the camera, waving and performing in front of the camera were more rare reactions (Given *et al.* 2016). However, in these studies, the overall frequency of children's attention to the camera recording was low (Barriage and Darcey 2019). There are surprisingly few systematic studies of such behaviour in K12 school contexts, and the same goes for the effects of video-based research observations when the target group is students with special needs (Dukuzumuremyi and Siklander 2018). In the core articles (SET 2), the focus was more often on the teachers'/student teachers' reactivity than on the attention and performance of the pupils.

Precautionary and mitigating factors. To the extent that the core articles and the added entries describe measures taken to minimize reactivity, they repeat the advice provided in the general methodological literature referred to above: (1) making the researcher and recording instruments/operations less visible, (2) sharing identity with those observed, for example by studying one's own class, (3) reducing apprehension by presenting the project etc. as non-consequential and (4) 'hanging around' as a familiarization strategy (Skårås 2018) for a period of time and gaining the trust of the participants. Additionally, triangulating behavioural observations with less reactive forms, such as nonverbal actions, physiological measures, human use of space etc., may be a way of monitoring changes (Goodwin and Velicer 2008) and establishing more non-obtrusive indicators. Departing from these common strategies for reducing reactivity, two types of recording device design for observational studies are advocated in the reviewed core literature: (1) partly covert recording, where participants are not actively informed about the target and timing of camera observation, and (2) vicarious recording where those observed are entrusted with this activity. The former is associated with a number of challenges in terms of informed consent, but also reactivity.

Vicarious video recording in classrooms and educational settings is likely to take many forms – of which the use of permanent video recording ('surveillance') is one variant. As noted above, the ubiquity of video recording devices such as iPads in the hands of teachers and students may encourage strategies that eliminate or reduce the presence of the researcher as observer. Teachers, parents and sometimes children and students take on a vicarious role by doing the technical filming in line with the researchers' instructions (MacMahon *et al.* 2019). A more participative approach would be to have subjects define the script, which is believed to provide a more 'emic' research design (Given *et al.* 2016) resulting in 'very rich and detailed data about children's use of technology ...' (p. 1). Again, this approach is more common in health research and children studies (Burn and Richards 2014, Aarsand 2016) than educational research, but has a growing support in the latter field (Hanks 2019). No doubt the engagement of students as co-researchers in this respect may provide new insights in the field. However, reactivity is certainly not eliminated or reduced, since the 'handing of the camera to informants' represents a transformation of the research design into an interventional format where participants' engagement in recording activities complicates the sorting out of 'reactivity'. Our scan of literature on action research and educational design-based research confirms that 'reactivity' or the 'Hawthorne effect' are rarely addressed and that more general appeals to 'reflexivity' prevail (Robertson 2000, Lyngsnes 2016).

In review. Our comparison of our core sample, SET 2, with video-based classroom studies that do not report reactivity or observer effect, SET 3, indicates that research design features associated with a quantitative tradition have to a larger degree standardized data collection layout and procedures. From our reading of these texts, it is not possible to sort out whether these differences are generated by a lack of methodological attentiveness or norms for research reporting. Several articles in SET 3 refer to a standard setup for large scale video studies (Jacobs *et al.* 2007, Klette 2009, Seidel and Thiel 2017), which are in accordance with the technical aspects of international video studies in education (Seidel *et al.* 2005, OECD 2020). This contrasts with our core sample of mainly smaller scale studies with a research focus on the learning potential of new digital learning environments, sometimes combined with participative and vicarious collection strategies using unfamiliar recording devices such as eye trackers. These designs should provoke reflections on how to disentangle the causal patterns at play (Morrison 2009), but, again, we need to stress that the number of video-based classroom studies reporting reactivity is rather low, and that the description provided was in most cases rather general. Recommendations for minimization and acclimatization dominate. Our original bibliographic procedure did not generate a sufficient data material to fully explore this issue and to identify patterns between the categories aligned with the classification in Table 1. When combined with the others literature sets including the methodological, the core material indicates that there is little reported concern for 'reactivity' in classroom video studies. This does not align with focused methodological debates that advice researchers to pay attention to and be explicit about such matters. There is a lack of studies addressing reactivity effects on different dimensions of teaching / learning and the determinants of such influences.

Discussion and concluding comments

Unlike many attempts at using conceptual analysis as a platform for clarifying ambiguities and vagueness, our intention with this review is rather to go 'behind' the terminological differences by outlining their dependency on meta-methodological positions. In this effort we ran into various skirmishes between epistemic paradigms. The research strand founded in conversation analysis and ethnomethodology has generated a prolific terminology about the co-construction between researchers and participants in terms of the latter's 'staging of performances'. Our review has indicated that this approach may generate rich casuistic evidence (Silverman 2013) of strategies adopted by children of varying age, in private / public settings and confronted with new recording technology and research designs. However, to the extent that social constructivism refrains from making generalizations contenting with moment-by-moment demonstrations, one could envisage a division of work between explorative studies of video effects and research aiming at causal explanations of underlying mechanisms – to be integrated as elements of a programmatic research strategy (Klette *et al.* 2021). We need to add that the majority of video-based studies in our sample that reported reactivity, were concerned with mitigating aspects such as the reduced salience with small, stabilized cameras and fading effects over time. No reference was made to the debate about reactivity as 'staged performance' versus reactivity as contaminating effect.

The low percentage of articles addressing reactivity in concert with the brief mentioning of such effects and a lack of provisions for confirming their occurrence and severity, suggest that reactivity in video-based classroom studies may not be regarded as a serious methodological issue. Our excursions into research on social interaction within neighbouring fields indicate that such a diagnostic is more widespread. In both the former and the latter literature reactivity was only rarely an object for specific investigations guided by assumptions (Zahle 2019) about its potential effects.

Despite the low occurrence of reported reactivity in our sampled literature, the study revealed a number of factors that should be attended to for systematic validation: factors linked to recording devices and their layout including provisions for distant and vicarious data collection, the social organization of learning environments in classes, small group works, projects, age differences of

subjects and teacher/researcher characteristics, and degrees of intrusiveness in fields considered as public versus private. Also, the need to clarify the impact of different research purposes, notably the distinction between descriptive and interventional designs, is highlighted in our article. However, in the literature with reported reactivity there were few studies that provided solid convergent evidence for these effects and the success of mitigating measures. Thus we tend to agree with Praetorius *et al.* (2017, p. 68) that future video-based educational research should be investigating potential determinants and the multidimensionality of reactivity. They refer to own and other quantitative studies indicating that emotional variables may be more affected than behavioural in video research with teachers and students. This may be especially true when participants feel they are surveyed and the situation is evaluative (Kazdin 1982). However, nearly all the studies reviewed in this review focused on behavioural changes during video recording, which suggests that more attention should be devoted to cognitive and emotional mechanisms in combinations with material factors.

The discussion on naturalness and contamination takes a one key premise that our data should be descriptively adequate, and that such a quality is important for the inferences we draw from our empirical material. The OECD Talis Video Study (2020, p. 43) conclude with reference to Praetorius *et al.* (2017);

While there is evidence from both teachers and students that the videotaped lessons may have differed in some ways from typical lessons in the unit, previous research on this issue indicates this may not significantly influence conclusions drawn from videotaped lessons...

But the research referred to in this quotation and the cases reviewed in our article have mainly concentrated on departures from naturalness or the occurrence of reactivity, and not on how such effects invalidate inferences from the field. In this respect, Jimenez-Buedo's (2021) underline on malignancy effects may steer our 'reactivity assumptions' towards going beyond contamination of the typical and to investigate whether it matters for our conclusions. This strategy will strengthen not only the validation of our research. It should provide better evidence for targeted mitigating measures that are not merely based on a priori expectations about precautionary effects. More transparency about the empirical basis for field researchers' judgement concerning reactivity and validity claims should strengthen the methodological training of PhD-students on such topics.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Leif Christian Lahn  <http://orcid.org/0000-0001-9032-1358>

References

- Entries marked with * are part of core SET 2 (literature sources) video studies in education addressing "reactivity".
- Aarsand, P., 2016. Children's media practices: challenges and dilemmas for the qualitative researcher. *Journal of children and media*, 10 (1), 90–97. doi:10.1080/17482798.2015.1121894.
- Aarsand, P., and Forsberg, L., 2010. Producing children's corporeal privacy: ethnographic video recording as material-discursive practice. *Qualitative research*, 10 (2), 249–269. doi:10.1177/2F1468794109356744.
- Adler, P.A., and Adler, P., 1994. Observational techniques. In: N.K. Denzin and Y.S. Lincoln, eds. *Handbook of qualitative research*. Thousand Oaks, CA: Sage, 377–392.
- Anderson, T., and Shattuck, J., 2012. Design-based research: a decade of progress in education research? *Educational researcher*, 41 (1), 16–25. doi:10.3102/0013189X11428813.
- Andersson, E., and Sørvik, G.O., 2013. Reality lost? Re-use of qualitative data in classroom video studies. *Forum Qualitative Sozialforschung / Forum: qualitative social research*, 14 (3), doi:10.17169/fqs-14.3.1941.
- Antal, H., *et al.*, 2015. Audio-video recording of health care encounters for pediatric chronic conditions: observational reactivity and its correlates. *Journal of pediatric psychology*, 40 (1), 144–153. doi:10.1093/jpepsy/jsu046.

- Baldassarri, D., and Abascal, M., 2017. Field experiments across the social sciences. *Annual review of sociology*, 43, 41–73. doi:10.1146/annurev-soc-073014-112445.
- Barriage, S., and Darcey, K., 2019. Okay, now the video is on: an analysis of young children's orientations to the video camera in recordings of family interactions. *Journal of childhood studies*, 44 (3), 18–41. doi:10.18357/jcs00019172.
- Baum, C.G., Forehand, R., and Zegjob, L.E., 1979. A review of observer reactivity in adult-child interactions. *Journal of behavioral assessment*, 1 (2), 167–178.
- Bergold, J., and Thomas, S., 2012. Participatory research methods: a methodological approach in motion. *Forum qualitative social research*, 13 (1), Art. 30. <http://nbn-resolving.de/urn:nbn:de:0114-fqs1201302>.
- Bhatt, I., 2017. Classroom digital literacies as interactional accomplishments. In: M. Knobel and C. Lankshear, eds. *Researching new literacies*. New York: Peter Lang, 127–149.
- Blease, D., 1983. Observer effects on teachers and pupils in classroom research. *Educational review*, 35 (3), 213–217. doi:10.1080/0013191830350302.
- Blikstad-Balas, M., 2017. Key challenges of using video when investigating social practices in education: contextualization, magnification, and representation. *International journal of research and method in education*, 40 (5), 511–523. doi:10.1080/1743727X.2016.1181162.
- *Blikstad-Balas, M., and Sørvik, G.O., 2015. Researching literacy in context: using video analysis to explore school literacies. *Literacy*, 49 (3), 140–148. doi:10.1111/lit.12037.
- Bracht, G.H., and Glass, G.V., 1968. The external validity of experiments. *American educational research journal*, 5, 437–474.
- Burn, A., and Richards, C., eds. 2014. *Children's games in the new media age: childlore, media and the playground*. London: Routledge.
- *Civitillo, S., et al., 2019. The interplay between culturally responsive teaching, cultural diversity beliefs, and self-reflection: a multiple case study. *Teaching and teacher education*, 77, 341–351. doi:10.1016/j.tate.2018.11.002.
- Codding, R.S., et al., 2008. Using performance feedback to improve treatment integrity of classwide behavior plans: an investigation of observer reactivity. *Journal of applied behavior analysis*, 41 (3), 417–422.
- Collier, A., and Wyer, M., 2016. Researching reflexively with patients and families: two studies using video-reflexive ethnography to collaborate with patients and families in patient safety research. *Qualitative health research*, 26 (7), 979–993. doi:10.1177/1049732315618937.
- Derry, S.J., et al., 2010. Conducting video research in the learning sciences. *Journal of the learning sciences*, 19 (1), 3–53. doi:10.1080/10508400903452884.
- Dukuzumuremyi, S., and Siklander, P., 2018. Interactions between pupils and their teacher in collaborative and technology-enhanced learning settings in the inclusive classroom. *Teaching and teacher education*, 76, 165–174. doi:10.1016/j.tate.2018.08.010.
- Erickson, F., 2006. Definition and analysis of data from videotape: Some research procedures and their rationales. In: J. L. Green, G. Camilli, and P. B. Elmore, eds. *Handbook of complementary methods in education research*. Mahwah, NJ: Lawrence Erlbaum Associates, 177–192.
- Flyvbjerg, B., 2002. *Making social science matter*. Cambridge: Cambridge University Press.
- Gilbert, L.S., 2002. Going the distance: 'closeness' in qualitative data analysis software. *International journal of social research methodology*, 5 (3), 215–228. doi:10.1080/13645570210146276.
- Given, L.M., et al., 2016. Parents as co-researchers at home: using an observational method to document young children's use of technology. *International journal of qualitative methods*, doi:10.1177/1609406915621403.
- Goh, F., A. Carroll and R.M. Gillies, 2019. A review of the use of portable technologies as observational aids in the classroom. *Information and learning science*, 120, doi:10.1108/ILS-08-2018-0080.
- Goodwin, M.S., and Velicer, W.F., 2008. Telemetric monitoring in the behavior sciences. *Behavior research methods*, 40 (1), 328–341. doi:10.3758/BRM.40.1.328.
- Gordon, C., 2010. Beyond the observer's paradox: the audio-recorder as a resource for the display of identity. *Qualitative research*, 13 (3), 299–317. doi:10.1177/1468794112442771.
- Gresham, F.M., Dart, E.H., and Collins, T.A., 2017. Generalizability of multiple measures of treatment integrity: comparisons among direct observation, permanent products, and self-report. *School psychology review*, 46 (1), 108–121. doi:10.1080/02796015.2017.12087606.
- *Haataja, E., et al., 2019. Teacher's visual attention when scaffolding collaborative mathematical problem solving. *Teaching and teacher education*, 86, 2–14. doi:10.1016/j.tate.2019.102877.
- Hammersley, M., 2003. Analytics are no substitute for methodology: a response to Speer and Hutchby. *Sociology*, 37 (2), 339–352. doi:10.1177/0038038503037002007.
- Hammersley, M., 2008. Troubles with triangulation. In: M.M. Bergman, ed. *Sage research methods in innovation and social research methods*. Thousand Oaks, CA: Sage, 22–36.
- Hammersley, M., 2011. Objectivity: a reconceptualization. In: M. Williams, and W.P. Vogt, eds. *The Sage handbook of methodological innovation*. London: Sage, 8–21. doi:10.4135/9781446268261.n3.
- Hammersley, M., 2018. What is ethnography? Can it survive? Should it? *Ethnography and education*, 13 (1): 1–17. doi:10.1080/17457823.2017.1298458.

- Hanks, J., 2019. From research-as-practice to exploratory practice-as-research in language teaching and beyond. *Language teaching*, 52 (2), 143–187. doi:10.1017/S0261444819000016.
- Harris, F., and Lahey, B., 1982. Subject reactivity in direct observational assessment: a review and critical analysis. *Clinical psychology review*, 2, 523–538.
- Haw, K., and Hadfield, M., 2011. *Video in social science research. Functions and forms*. London: Routledge.
- Hazel, S., 2016. The paradox from within: research participants doing-being-observed. *Qualitative research*, 16 (4), 446–467. doi:10.1177/1468794115596216.
- Heath, C., Hindmarsh, J., and Luff, P., 2010. *Video in qualitative research: analysing social interaction in everyday life*. Thousand Oaks, CA: SAGE.
- Ilin, G., Kutlu, O., and Kutluay, A., 2013. An action research: using videos for teaching grammar in an ESP class. *Procedia – social and behavioral sciences*, 70, 272–328. doi:10.1016/j.sbspro.2013.01.065.
- Jacobs, J.K., Hollingsworth, H., and Givvin, K.B., 2007. Learned from the TIMSS video studies video-based research made “easy”: methodological lessons. *Field methods*, 19, 284–299. doi:10.1177/1525822X07302106.
- Jewitt, C., 2012. An introduction to using video for research. *NCRM Working Paper 03/12*. http://eprints.ncrm.ac.uk/2259/4/NCRM_workingpaper_0312.pdf
- Jimenez-Buedo, M., 2021. Reactivity in social scientific experiments: what is it and how is it different (and worse) than a placebo effect? *European journal for philosophy of science*, 11 (42), 1–22. doi:10.1007/s13194-021-00350-z.
- Jordan, B., and Henderson, A., 1995. Interaction analysis: foundations and practice. *The journal of the learning sciences*, 4 (1), 39–103.
- Kazdin, A.E., 1982. Observer effects: reactivity of direct observation. *New directions for methodology of social and behavioral science*, 14, 5–19.
- Kleckner, I.R., et al., 2020. Framework for selecting and benchmarking mobile devices in psychophysiological research. *Behavior research methods*, doi:10.3758/s13428-020-01438-9.
- Klette, K., 2009. Challenges in strategies for complexity reduction in video studies. Experiences from the PISA+ study: a video study of teaching and learning in Norway. In: T. Janik and T. Seidel, eds. *The power of video studies in investigating teaching and learning in the classroom*. New York: Waxmann, 61–82.
- Klette, K., and Blikstad-Balas, M., 2018. Observation manuals as lenses to classroom teaching: pitfalls and possibilities. *European educational research journal*, 17 (1), 129–146.
- Klette, K., Blikstad-Balas, M., and Tengberg, M., eds., 2021. *Ways of analyzing teaching quality. Potentials and pitfalls*. Oslo: Universitetsforlaget. doi:10.18261/9788215045054-2021-06.
- Knoblauch, H., and Tuma, R., 2011. Videography: an interpretative approach to video-recorded micro-social interaction. In: E. Margolis, and L. Pauwels, eds. *SAGE handbook of visual research methods*. London: Sage, 414–430.
- Knoblauch, H., Schnettler, B., and Raab, J., 2009. Video-analysis. Methodological aspects of interpretive audio-visual analysis in social research. In: H. Knoblauch, B. Schnettler, J. Raab, and G. Soeffner, eds. *Video-analysis: methodology and Methods*. Frankfurt am Main: Peter Lang, 9–29.
- Knoblauch, H., Schnettler, B., and Tuma, R., 2018. Videography. In: U. Flick, ed. *The Sage handbook of qualitative data collection*. London: Sage. doi:10.4135/9781526416070.
- *Kucirkova, N., and Falloon, G., 2018. iPad apps and visual methodologies: empirical and ethical issues in achieving authentic data. *Research in learning technology*, 26, 20–29. doi:10.25304/(1).
- Labov, W., 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania.
- Lee, R.M., 2000. *Unobtrusive methods in social research*. Buckingham: Open University Press.
- Levering, B., 2002. Concept analysis as empirical method. *International journal of qualitative methods*, 1 (1), 35–48.
- Lyngsnes, K., 2016. A reflexive eye on a collaborative action research project in school. *The qualitative report*, 21 (2), 196–211. <https://nsuworks.nova.edu/tqr/vol21/iss2/3>.
- MacBeth, D., 2001. On “reflexivity” in qualitative research: two readings, and a third. *Qualitative inquiry*, 7 (1), 35–68.
- *MacMahon, B., Grádaigh, S.O., and Ghuidhir, S.N., 2019. Super vision: the role of remote observation in the professional learning of student teachers and novice placement tutors. *TechTrends*, 63, 703–710. doi:10.1007/s11528-019-00432-z.
- *Magnussen, R., et al., 2017. Mobile eye tracking methodology in informal e-learning in social groups in technology enhanced. *Journal of e-learning*, 15, 46–58.
- Maxwell, J.A., 2012. *A realist approach for qualitative research*. London: Sage.
- McCambridge, J., Kypri, K., and Elbourne, D., 2014a. Research participation effects: a skeleton in the methodological cupboard. *Journal of clinical epidemiology*, 67 (8), 845–849. doi:10.1016/j.jclinepi.2014.03.002.
- McCambridge, J., Witton, J., and Elbourne, D.R., 2014b. Systematic review of the Hawthorne effect. New concepts are needed to study research participation effects. *Journal of clinical epidemiology*, 67 (2), 267–277. doi:10.1016/j.jclinepi.2013.08.015.
- McCord, R.E., and Matusovich, H.M., 2019. Naturalistic observations of metacognition in engineering: using observational methods to study metacognitive engagement in engineering. *Journal of engineer education*, 108, 481–502. doi:10.1002/jee.20291.
- *McIntyre, N.A., Mulder, K.T., and Mainhard, M.T., 2020. Looking to relate: teacher gaze and culture in student-rated teacher interpersonal behaviour. *Social psychology of education*, 23, 411–431. doi:10.1007/s11218-019-09541-2.

- Mercators, M., and Craighead, E., 1974. Effects of nonparticipant observation on teacher and pupil classroom behavior. *Journal of educational psychology*, 66 (4), 512–519.
- Monahan, T., and Fisher, J., 2010. Benefits of 'observer effects': lessons from the field. *Qualitative research*, 10 (3), 357–376. doi:10.1177/1468794110362874.
- Mondada, L., 2006. Video recording as the reflexive preservation and configuration of phenomenal features for analysis. In: H. Knoblauch, J. Raab, H.-G. Soeffner, and B. Schnettler, eds. *Video analysis*. Bern: Peter Lang, 51–68.
- Mondada, L., 2014. Ethics in action: anonymization as a participant's concern and a participant's practice. *Human studies*, 37, 179–209. doi:10.1007/s10746-013-9286-9.
- Morrison, K., 2009. *Causation in educational research*. Oxon: Routledge.
- Muench, J., Sanchez, D., and Garvin, R., 2013. A review of video review: New processes for the 21st century. *International journal of psychiatry in medicine*, 45 (4), 413–422. doi:10.2190/PM.45.4.k.
- Nassauer, A., and Legewie, N.M., 2019. Analyzing 21st century video data on situational dynamics. Issues and challenges in video data analysis. *Sociology, social sciences*, 8, doi:10.3390/socsci8030100.
- Newman, M., 2010. Beyond structural realism: pluralist criteria for theory evaluation. *Synthese*, 174, 413–443. doi:10.1007/s11229-009-9463-6.
- Newman, M., and Gough, D., 2020. Systematic reviews in educational research: methodology, perspectives and application. In: O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, and K. Buntins, eds. *Systematic reviews in educational research*. doi:10.1007/978-3-658-27602-7_1M.
- Niiniluoto, I., 2003. *Critical scientific realism*. Oxford: Oxford Scholarship Online. doi:10.1093/0199251614.001.0001
- Nolan, A., Paatsch, L., and Scull, J., 2018. Video-based methodologies: the affordances of different viewpoints in understanding teachers' tacit knowledge of practice that supports young children's oral language. *International journal of research and method in education*, 41 (5), 536–547.
- OECD, 2020. *Global teaching insights. A video study of teaching*. Paris: OECD Publishing. doi:10.1787/20d6f36b-en
- Ogden, T., and Fixsen, D.L., 2014. Implementation science. A brief overview and a look ahead. *Zeitschrift für Psychologie*, 222 (1), 4–11. doi:10.1027/2151-2604/a000160.
- Orne, M.T., 1962. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *American psychologist*, 17, 776–783.
- Patton, M.Q., 2015. *Qualitative research and evaluation methods*. 4th ed. Thousand Oaks, CA: Sage.
- Praetorius, A.-K., and Charalambous, C.Y., 2018. Classroom observation frameworks for studying instructional quality: looking back and looking forward. *ZDM*, 50, 535–553. doi:10.1007/s11858-018-0946-0.
- Praetorius, A.-K., McIntyre, N.A., and Klassen, R.M., 2017. Reactivity effects in video-based classroom research: an investigation using teacher and student questionnaires as well as teacher eye-tracking. *Zeitung Erziehungswissenschaft*, 20 (Suppl 1), 49–74. doi:10.1007/s11618-017-0729-3.
- Prieto, L.P., et al., 2018. Orchestration load indicators and patterns: in-the-wild studies using mobile eye-tracking. *IEEE transactions on learning technologies*, doi:10.1109/TLT.2017.2690687.
- Psillos, S., 2011. Choosing the realist framework. *Synthese*, 180 (2), 301–316.
- Redman, S., et al., 1989. The assessment of reactivity in direct observation studies of doctor-patient. *International journal of psychology and health*, 3 (1), 17–28. doi:10.1080/08870448908400362.
- Robertson, J., 2000. The three Rs of action research methodology: reciprocity, reflexivity and reflection-on-reality. *Educational action research*, 8 (2), 307–326. doi:10.1080/0965079000200124.
- Rosenthal, R., 1966. *Experimenter effects in behavioral research*. New York: Appleton.
- Sahlström, F., Tanner, M., and Valasmo, V., 2019. Connected youth, connected classrooms. Smartphone use and student and teacher participation during plenary teaching. *Learning, culture and social interaction*, 21, doi:10.1016/j.lcsi.2019.03.008.
- Sechrest, L., ed., 1982. *Unobtrusive measurement today: new directions for methodology of behavioral science*. San Francisco, CA: Jossey-Bass.
- Seidel, T., and Thiel, F., 2017. Video based teaching research. Analysis of teaching quality, development of learnability and measurement [Videobasierte unterrichtsforschung. Analysen von unterrichtsqualität, gestaltung von lerngelegenheiten und messung]. *Zeitschrift für Erziehungswissenschaft*, 32, 1–21.
- Seidel, T., Prenzel, M., and Kobarg, M., eds., 2005. *How to run a video study: technical report of the IPN video study*. Münster: Waxmann.
- Shapero, B.G., et al., 2016. Emotional reactivity. In: R. Levesque, ed. *Encyclopedia of adolescence* (pp. 1–9). Cham: Springer.
- Shoecraft, K., and Fluckiger, B., 2018. Conducting qualitative video research with young children. *Qualitative research journal*, 18 (3), 238–247. doi:10.1108/QRJ-D-17-00031.
- Silverman, D., 2013. What counts as qualitative research? Some cautionary comments. *Qualitative sociology review*, 9 (2), 48–55. doi:10.31857/S013216250006160-9.
- *Skårås, M., 2018. Focused ethnographic research on teaching and learning in conflict zones: history education in South Sudan. *Forum for development studies*, 45 (2), 217–238. doi:10.1080/08039410.2016.1202316.
- Soltis, J.F., 1978. *An introduction to the analysis of educational concepts*. Boston, MA: Addison Wesley.
- Spano, R., 2006. Observer behavior as a potential source of reactivity: describing and quantifying observer effects in a large-scale observational study of police. *Sociological methods and research*, 34 (4), 521–553. doi:10.1177/0049124106286331.

- Speer, S.A., and Hutchby, I., 2003a. From ethics to analytics: aspects of participants' orientations to the presence and relevance of recording devices. *Sociology*, 37 (2), 315–337. doi:10.1177/0038038503037002006.
- Speer, S.A., and Hutchby, I., 2003b. Methodology needs analytics: a rejoinder to Martyn Hammersley. *Sociology*, 37 (2), 353–359. doi:10.1177/0038038503037002008.
- Steinmetz, J., and Pfattheicher, S., 2017. Beyond social facilitation: a review of the far reaching effects of social attention. *Social cognition*, 35 (5), 585–599. doi:10.1521/soco.2017.35.5.585.
- Torraco, R.J., 2016. Writing integrative literature reviews: using the past and present to explore the future. *Human resource development review*, 15 (4), 404–428. doi:10.1177/1534484316671606.
- Vogelsang, K., 2012. Conceptual history: a short introduction. *Oriens extremus*, 51, 9–24.
- *Wash, P.D., Bradley, G., and Beck, J., 2014. Remote classroom observations with preservice teachers. *SRATE journal*, 24 (1), 58–65.
- Webb, E., et al., eds., 2000. *Unobtrusive measures*. Thousand Oaks, CA: Sage.
- Weinrott, M.R., Garrett, B., and Todd, N., 1978. The influence of observer presence on classroom behavior. *Behavior therapy*, 9, 900–911.
- Wolfswinkel, J.F., Furtmueller, E., and Wilderom, C.P., 2013. Using grounded theory as a method for rigorously reviewing literature. *European journal of information systems*, 22, 45–55. doi:10.1057/ejis.2011.51.
- Xu, L., and Clarke, D., 2019. Validity and comparability in cross-cultural video studies of classrooms. In: L. Xu, G. Aranda, W. Widjaja, and D. Clarke, eds. *Video-based research in education: cross-disciplinary perspectives*. Abingdon: Routledge, 19–33.
- Zahle, J., 2019. Data, epistemic values, and multiple methods in case study research. *Studies in history and philosophy of science*, 78, 32–39. doi:10.1016/j.shpsa.2018.11.005.