

A Systematic Review of Research Evidence Reporting Educational Psychologists' Use of Contextual Observation in Practice

Dr Jennifer Leatherbarrow, Professor Kevin Woods, Dr George Thomas, and Dr Kathleen Tyldesley
School of Environment, Education and Development, Manchester Institute of Education

Considering the apparent centrality of contextual observation within educational psychologists' (EPs') daily practice, it is argued here that some level of rigour should be available through review of relevant research and self-evaluation, so that contextual observation meets the professional requirement for evidence-based practice (Health and Care Professions Council, 2016). This systematic literature review used an adapted critical evaluation checklist and a bespoke analytic framework for contextual observation, created by the authors, to systematically identify and critically evaluate the research evidence of EPs' use of contextual observation within practice. Across the studies, there were inconsistencies around which details of contextual observation were reported, which makes it difficult for the process of contextual observation to be reliably replicated across EP practice or to be communicated clearly to new entrants to the profession. The bespoke analytic framework was enhanced by incorporating features of the adapted evaluative checklist to produce an analytic framework for contextual observation which combines academic knowledge with practice-informed knowledge. This potentially provides a contribution to the evidence base for EP use of contextual observation, as well as laying a foundation for building an understanding of its process. Further research can be directed towards the development of guidelines for best practice of contextual observation in order to produce an evidence-based tool for use across the profession.

Keywords: educational psychologists, educational psychology practice, observation, contextual observation, assessment

Introduction

Contextual Observation in Educational Psychology Practice

Reconstructing Educational Psychology (Gillham, 1978) was published as a key text 41 years ago. This "landmark" work drew together contemporary perspectives from leading practitioners and academics. It marked a paradigm shift for educational psychology, set within a postmodern/social constructionist/phenomenological epistemology that challenged approaches used to assess "child difficulties", and the relevance of psychometric assessment in particular. It moved the focus onto methods that worked with the child's environment and school systems, using project work and research, to promote an interactionist perspective on child development within school and home "systems".

In the 2000s, evidence-based practice encouraged educational psychologists (EPs) to account for the effectiveness and cost-effectiveness (Frederickson, 2002) of their work and further supported the development of systemic models of professional practice and consultation, and across a wide range of role functions. Fallon et al. (2010, p. 4) describe the EPs' role in terms of "utilising psychological skills, knowledge and understanding through the functions of consultation, assessment, intervention, research and training, at or-

ganisational, group or individual level across educational, community and care settings, with a variety of role partners".

From earlier research carried out by the first author for her Doctoral thesis (Leatherbarrow, 2020), as well as the authors' many years of EP practice, it is known that observation is used to gather information across all of these core functions of the EP role as one of a "variety of tools, techniques and approaches", that a psychological assessment should involve (British Psychological Society, 2002, p. 26). It was found that observation may be considered one part in a wider information-gathering process set within a framework for practice, and places presented issues or concerns in the context in which they are being observed. For the purpose of this study, the authors used the term *contextual observation* and offer a definition which positions it as the rigorous tool for practice that this research sets the foundation for:

Contextual observation is used by EPs as an information-gathering method that occurs in vivo or by video capture and may be used across the range of role functions.

There is, however, a question around how this contextually sensitive work is consistently operationalised and utilised in the absence of any guidelines. Given the centrality of contextual observation within EPs' daily practice, it is

argued that some level of rigour around contextual observation should be available, developed through review of relevant research and self-evaluation, so that contextual observation meets the professional requirement for evidence-based practice (Health and Care Professions Council, 2016).

Professional Requirement for Evidence-Based Practice and Rationale of the Current Review

Considering the central position of contextual observation in EP practice, it is surprising that some important aspects of the process around it are not generally known, for example, when and why contextual observation is used, when and why it may be more or less structured, and how it is recorded, analysed and utilised. This knowledge gap leaves practitioners, and their professional training programmes, less able to effectively self-evaluate practice in this area, and potentially open to challenge that their practice falls below best practice standards.

The HCPC Standards of Proficiency (SoPs) require EPs to emerge from professional training able to demonstrate a logical and systematic approach to problem-solving (Health and Care Professions Council, 2016, SoP 14.12). Broad professional practice frameworks have become embedded within practice, such as the Mosen Problem-Solving Framework (1998), which prompt EPs to gather information using a variety of assessment tools. There is, however, very little research around EPs' use of specific assessment, or evaluation skills, including contextual observation, to support the teaching, practice and development of such skills in a systematic way. Some fifteen years ago, Woods and Farrell (2006) surveyed the range of EP assessment practices and provided a starting point for the development of knowledge around EP assessment. Its sustained relevance demonstrates the importance of continuing this work within the current context.

Crombie et al. (2014) identify literature relating to unconscious professional practice that may explain to some extent why there is very little peer-reviewed research about EPs' use of contextual observation. The paper set out to identify what was unconscious about the practice of professionals working in a school, in order to make the complexity of professional competence more explicit. This is directly transferable to the process of contextual observation. This is because without guidelines to capture the complexity of the process around contextual observation, the transference of skills and knowledge around it relies on processes learnt "experientially", which does not sit easily with the evidence-based practice requirement for "accountability".

Related to this, Ahrenkiel et al. (2013) refer to "unnoticed professional competence" (p. 4) referring to the individual aspects of professional practice which are not considered separately as individually important aspects of a whole process. Within contextual observation, this translates as all of the considerations around its process that individual EPs

give attention to but that are not necessarily recorded. This is because there is an assumption that the process is understood by all those who use it, prompting less urgency in practitioners' reporting of its details. In the absence of an overarching approach that captures all of these "unnoticed" elements of contextual observation, the effect may be that practice is based on custom and practice of individual psychologists being conveyed through professional conversations or shadowing experiences. As such, contextual observation as an information-gathering tool within a problem-solving approach is not currently embedded within an evidence base. This makes it difficult for the process to be defended as consistent and rigorous, or developed and improved, because there is not yet an explicitly agreed standard of best or effective practice from which to work.

The British Psychological Society (BPS, 2017) professional practice guidelines set out that "research provides the evidence base for the practice of psychology" (p. 9). They also make specific reference to the "application of systematic observation" (p. 9), providing further rationale for an urgent contribution to the evidence base in this area of practice.

Aims of the Current Review

Discussions between representatives of regional Educational Psychology Services (EPSs), which offer practice placements for trainee EPs, and the respective regional English university provider of EP training identified a valuable opportunity for research in this area to be commissioned. The aim of the current systematic literature review is to:

- systematically identify and critically evaluate research evidence of EPs' use of contextual observation;
- propose an initial analytic framework for contextual observation; and
- identify implications for further research needed for the development of the analytic framework.

Method of the Review

Literature Search Strategy and Review Process

Stage One

The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) framework (Moher et al., 2009) was used to identify and select appropriate papers for the review. Electronic searches of journal titles, abstracts and keywords were undertaken in December 2017. The following databases were systematically searched for relevant studies: Psych Info, Education Resources Information Centre, Google Scholar and the Manchester Online library. The following search terms: observation AND psychologist* AND education* AND practice, were used. This search, however,

produced only seven potentially relevant papers, all of which were cited in peer-reviewed journals. When the search was expanded to include full texts, the number of papers returned was too great to manage with limited time and resources, and so an alternative approach was taken.

Stage Two

It was decided that relevant research would be more reliably identified by a comprehensive hand-search of the most relevant journals. These were selected by consulting a reference group with academic and practitioner educational psychology roles across university and service settings. Reference group members were asked about their preferences for reading and publishing practice-based educational psychology research. This scoping exercise identified six journal titles which included: *Educational Psychology in Practice*, *British Journal of Educational Psychology*, *Educational and Child Psychology*, *Psychology in the Schools*, *School Psychology International*, *Pastoral Care in Education*. The searching period was a ten-year period covering 2007 to 2017 initially. This was updated with subsequent searches for new research in 2018 to identify later published research before the completion of this review. Reading titles and abstracts identified the seven papers produced through the database search in Stage One, as well as nine further papers. The sixteen papers identified through Stages One and Two searches were then read to ensure they met the following inclusion criteria:

1. published between 2007 and 2018 at the time of the search
2. written in English;
3. focusing on the empirical practice of EPs or school psychologists (SPs);
4. making specific reference to the use of observation carried out by an EP or SP; and
5. subjected to peer review in an academic journal.

No papers needed to be excluded, as they all met the inclusion criteria.

Data Classification

Firstly, the papers were subjected to review using methodological quality evaluation criteria which were adapted by the authors to fit the focus of the review, that being educational psychologists' use of observation ("the adapted quality evaluation checklist", available from the first author on request). As most of the identified papers were reporting other assessment methods, as well as observation, it was difficult to make a meaningful evaluation of the quality of the research as a whole, and so the checklist combined Weights of Evidence (WoE) B and C with WoE A to afford the authors greater precision relating to the focus.

The adapted quality evaluation checklist amalgamated qualitative and quantitative investigation research appraisal

criteria from the previously widely used University of Manchester review frameworks for qualitative and quantitative evaluation/investigation research 2017 and 2018 respectively (cf. Bond et al., 2013; Law & Woods, 2018; Sedgwick & Stothard, 2018) with distinctions made for quantitative- or qualitative-specific criteria, to reflect observation data collected through quantitative or qualitative approaches. Criteria applicable to both quantitative and qualitative observations include: appropriate utilisation of observation within research design; clear use of observation protocols; observation measures have demonstrable credibility/validity/trustworthiness/reliability; and well executed data collection using observations. Criteria specific to qualitative observations include: analysis close to the observation data; and negative case analysis from observation. Criteria specific to quantitative observations include: appropriate statistical analyses of observation data-descriptive or inferential; and multi-level, inter-group, individual and/or time trends identified from observation. This process crafted the single generic observation critical evaluation appraisal checklist with which the sixteen studies were examined. Each reporting of observation was able to achieve a score of 0 to 21 (some criteria carrying two points) regardless of whether the observation was qualitative or quantitative.

Each paper was read at least twice by the first author, and a 25 per cent sample (four papers) was read and coded by the co-authors, with moderation discussions of each to reach a consensus view on the quality of methodology and to ensure consistent interpretation and application of the quality criteria. This process indicated a high level of post-discussion agreement (average inter-coder agreement of 99 per cent).

To compare scores fairly, reporting of observation which scored 1 to 9 points were deemed to be low quality, those which scored 10 to 14 points were defined as being medium quality, whilst more than 14 points earned a high-quality judgement. Where the reporting of the observation was identified as low quality, these studies were not excluded from the analysis because the quality of reporting of observation is acknowledged to be an interesting finding in its own right.

Data Extraction and Synthesis

Sixteen papers were included in the final review. The authors read each paper at least twice and highlighted salient content to populate each category presented in the Appendix.

To produce a synthesis of findings across the use of contextual observation, the studies were systematically analysed using the adapted quality evaluation checklist. Following this, the studies were further analysed using a bespoke overarching eleven-part framework (*The Bespoke Analytic Framework*, available from the first author on request). This was created by the authors, who include academics and practising EPs. This second phase in the analysis acknowledges that the studies exemplify the core functions of the EP role

and, as such, require psychological skills, knowledge and understanding to be explicitly communicated through the use of such a framework. The bespoke analytic framework makes explicit the skills, knowledge and understanding behind the process of observation, as identified by a group of academics and practising EPs.

Findings

The sixteen studies all exemplify EPs' use of contextual observation for the overarching purpose of research, which is an identified function of EP practice (Fallon et al., 2010). They have also been individually identified as having a focus that fits into the other four core functions of the role of the EP (assessment, consultation, intervention and training, Fallon et al., 2020). All studies were carried out in the UK.

Overview of Findings Using the Adapted Quality Evaluation Checklist

Presented first is an overview of the findings following an analysis of the studies that received a "high quality" rating for their reporting of observation (see the Appendix). The completed individual adapted quality evaluation checklists for the studies that received a "high" rating for the methodological quality of the reporting of observation were compared across categories to identify which categories were most commonly omitted. This provides an understanding of which details relating to contextual observation are not clearly described by the available literature. The common omissions in the high rated studies were: no, or little, information regarding negative case analysis for qualitative observations (Day, 2010; Regan & Howe, 2017); no, or little, information regarding researcher-participant negotiation of observation and process (Apter et al., 2010; Rait, 2012; Swinson & Knight, 2007); incomprehensiveness of observation documentation (K. Brown & Kennedy, 2011; Day, 2010; Swinson & Knight, 2007); and no, or little, evidence of attention to relevant ethical procedures and issues around observation (Day, 2010; Swinson & Knight, 2007)¹.

Overview of Findings Using the Bespoke Analytic Framework

Presented next is an overview of the findings following analysis of all of the studies in relation to the bespoke analytic framework. All of the studies used the term "observation(s)". Of the studies which used video as a medium through which to observe, one study referred to "video observation" (Rait, 2012). Two studies did not refer to "video" when describing the type of observation (K. Brown & Kennedy, 2011; Regan & Howe, 2017). Six studies used the term "schedule" (Apter et al., 2010; Burt & Stringer, 2018; Day, 2010; Hayes et al., 2007; Regan & Howe, 2017; Swinson, 2010). One referred to a "measure" (Rait, 2012)

and one study used the term "observation tool" (Vivash et al., 2018). Three studies used specific terms to describe the type of observation carried out: Sanders (2007) used "naturalistic observation", Crombie et al. (2014) and Fraser (2018) both used "narrative observation".

It was found that the *purposes* of the observations all fell within the range of practices that underpin the role of the EP, as identified by Fallon et al. (2010): *research* (Apter et al., 2010; Colville, 2013; Crombie et al., 2014; Day, 2010; Fraser, 2018; Sanders, 2007; Vivash et al., 2018), *training* (K. Brown & Kennedy, 2011; Burt & Stringer, 2018) *intervention* (Francis et al., 2017; Rait, 2012; Regan & Howe, 2017); and *consultation* (Hayes et al., 2007; Swinson, 2010). In relation to the function of *assessment*, it was found that observation was used for the purpose of identifying strengths and areas for development prior to training being delivered (Burt & Stringer, 2018), and to identify allocation to intervention groups (Francis et al., 2017), both within a research context.

It was also found across the studies that the purpose of observation was to provide contextual evidence about the topic or issue being addressed by the EPs' involvement. For example, Burt and Stringer (2018) supported a school to evaluate a maths programme and used observation to identify which aspects were working or needed development. K. Brown and Kennedy (2011, p. 382) described their videoed observations as "observations of interactive behaviour". Vivash et al. (2018) used observation to provide in-context information about how speech, language and communication needs (SLCN) are actually provided for in classroom environments and interactions, rather than relying on questionnaires alone. The Appendix provides further examples across all sixteen studies of observation being used for the purpose of generating contextualised information relating to the focus of the EP involvement. Fraser (2018, p. 650) refers specifically to the purpose of observation in the study, describing it as being "undertaken to triangulate the data through staff interviews and focus groups". Observation was a part contribution, with other data collection methods also contributing to the data set, in fourteen studies.

The studies discussed various elements of the process prior to an observation being carried out. In the bespoke analytic framework *contracting* describes the phase before consent for the observation has been obtained and incorporates explaining the purpose of the observation to the participants and ascertaining their feelings towards it. Hayes et al. (2007) gained an understanding of the attitude of participants to the research as a whole, which incorporated observation into its

¹Where a study scored 0 or ½ a point, it was considered that there was not sufficient information about that category in the reporting. Where a study scored 1 or 2 points, there was considered to be sufficient information to communicate the part of the process comprehensively.

methodology. The Appendix demonstrates that the purpose of the observations was clear across the studies. However, how this is communicated to participants and to what extent there is dialogue around how they feel about being observed was not clear in most of the studies (Burt & Stringer, 2018; Colville, 2013; Crombie et al., 2014; Day, 2010; Francis et al., 2017; Fraser, 2018; Rait, 2012; Sanders, 2007; Vivash et al., 2018).

In the category of *ethics*, the studies were considered in relation to whether they reported on obtaining consent, sensitivity and minimising harm when feeding back conclusions drawn when observation had been used. Half of the papers refer to obtaining consent for observation; three of these referred to “informed consent” (Apter et al., 2010; Francis et al., 2017; Regan & Howe, 2017). Examples of minimising harm and showing sensitivity to participants include occasions where the participants and the observer have a shared understanding that the process is collaborative (K. Brown & Kennedy, 2011; Fraser, 2018) and/or reflective. In other instances, feedback was also balanced, providing opportunity to discuss areas for change as well as what was working well, and this approach was well received by the participants (Burt & Stringer, 2018).

The category of *tools*, in the bespoke analytic framework, considers what the observer uses when observing, such as a specific schedule. The Appendix demonstrates that the use of a schedule was the most commonly used approach to observation, with more than half of the studies citing use of one. Nine studies reported the use of an observation tool (K. Brown & Kennedy, 2011; Burt & Stringer, 2018; Crombie et al., 2014; Day, 2010; Hayes et al., 2007; Rait, 2012; Swinson, 2010; Vivash et al., 2018). Three studies reported the use of a narrative style of observation (Crombie et al., 2014; Fraser, 2018; Sanders, 2007). Three studies reported the use of video observations (K. Brown & Kennedy, 2011; Rait, 2012; Regan & Howe, 2017) and two studies did not describe the type of observation used (Colville, 2013; Francis et al., 2017).

The studies were also analysed in relation to their reporting of what the EP(s) did *during* the time that they were in the immediate setting of the observation. This differs from exploring the “tools” used to observe, by considering aspects of the process such as where the EP physically positioned themselves, and whether there was any dialogue between the observer and anyone in the room where the observation took place. The reporting of this aspect of observation was particularly sparse, with almost all of the studies not providing relevant description beyond what tools were used. Exceptions to this include Swinson and Knight (2007) who explain that the observer sat at the back of the classroom, and began the observation when the class had settled, and Fraser (2018) who describes an uninvolved observer role.

The studies were analysed with regards to reporting the

contexts in which observations were carried out. This category explores whether observations were carried out in one context, such as one classroom, or across contexts, suggesting more than one place for the same purpose. It was found that in some studies observations were made in different classes, of different pupils, but for the same overall purpose of the use of observation such as observing behaviour or on- and off-task working. One study used video to record two observations of the same participant, both in the same classroom. In some studies it was not clear whether multiple sets of observation data were recorded on separate occasions in the same context or in different contexts; this was prevalent in the studies that used video recording and observation for the purpose of observing meetings as part of the research focus (Rait, 2012). There were studies where observations were reported to be carried out across contexts such as at different times of the school day (E. L. Brown et al., 2012). It was found that where observation was reported in an Early Years setting, the length and nature of the observation afforded the observer opportunity to observe the participants across contexts in the setting (Day, 2010). Where the observation was carried out for the purpose of observing with a narrow contextual focus such as a nurture group (Sanders, 2007) or for the purpose of assessing participants for a group or individual intervention-type (Francis et al., 2017) or for the purpose of training (Burt & Stringer, 2018), observation was in a single context relative to its purpose. Crombie et al. (2014) focused on individual children across a range of settings.

The studies were analysed in relation to their reporting of *reliability*. Joint observations were carried out in two studies (Apter et al., 2010; Swinson & Knight, 2007). Four studies reported inter-rater reliability (Apter et al., 2010; E. L. Brown et al., 2012; Rait, 2012; Swinson & Knight, 2007). Two papers referred directly to using other data in the study to triangulate with the observation data (E. L. Brown et al., 2012; Fraser, 2018). Five of the studies reported multiple observations of the same group over time although by the same observer (Crombie et al., 2014; Day, 2010; Fraser, 2018; Regan & Howe, 2017; Vivash et al., 2018). How reliability was achieved was not always clearly reported across the studies.

How the studies reported *validity* was also considered. Where studies reported using an observation schedule, the items on the schedule were vis-à-vis the purpose of the observation and linked to the overall aims of the research (Apter et al., 2010; Burt & Stringer, 2018; Hayes et al., 2007; Swinson & Knight, 2007; Swinson, 2010; Vivash et al., 2018). Rait (2012) designed a video coding scale specifically for coding the video-recorded observations; this was also provided as an appendix. E. L. Brown et al. (2012) did not use a schedule but agreed on the target behaviours which would be looked for during the observation, during a planning meeting. For narrative and naturalistic observations, the reported findings linked to the purpose of the observation (Fraser, 2018;

Sanders, 2007). Day (2010) reported that the data collection methods, which included observation, had been piloted.

In relation to actions *following an observation* analysis focused particularly upon content and process for feeding back. *Feedback* was found to be more evident in studies where there was a narrow focus of the observation, such as an individual participant, specific observable targets or as part of a directly reflective process to inform points for development for individual participants. For example, three of the studies which did refer to feedback were those using video to record the observation, those observed individual participants or parent–child interaction and individual teachers’ practice (K. Brown & Kennedy, 2011; Rait, 2012; Regan & Howe, 2017). One study used feedback as part of the approach to inform training (Burt & Stringer, 2018), and one study provided feedback to pupil participants which was directly related to identified target behaviours that were being looked for during the observation (E. L. Brown et al., 2012).

The studies were also analysed as to what *equipment* was reported to be used during the observations. A description of equipment was not provided by most of the studies. Regan and Howe (2017) identified the video medium as a tablet computer.

Discussion

Summary of Findings

Findings identified inconsistency in how much detail was included in the reporting of observation across the studies, illustrated through the inclusion of studies evaluated as having high, medium and low ratings for the quality of reporting of observation. Without comment from EPs, it cannot be assumed that the elements which were not reported were not practised. However, the absence of some areas of the bespoke analytic framework across the studies which were otherwise evaluated as high quality renders it potentially difficult for EPs to acquire comprehensive skills, knowledge and understanding about contextual observation.

Six studies received a “high” rating for their reporting of observation following application of the adapted quality evaluation checklist. These six studies all contained elements relating to eight categories of the bespoke analytic framework encompassing considerations of: terminology (referring to “observation”), the purpose of the observation, what happens during the observation, reliability and validity, observation across contexts and what occurs following an observation. Omitted considerations in relation to the bespoke analytic framework, which included: contracting, tools used, ethics and the identification of equipment used, occurred once.

Findings also identified a continuum of contribution to observation made to data-gathering in the studies, illustrated through the inclusion of studies where observation made a whole, part or majority contribution (see the Appendix). The

prevalence of observation making a part contribution reflects triangulation where an EP may draw on other sources of information in addition to observation to contribute to their psychological formulation (Cohen et al., 2007). It also reflects what was found in the preliminary study (Leatherbarrow, 2020) and what was discussed between the authors, that the use of observation is often one part of a wider problem-solving process. There was, however, no reporting of observation being used as part of the individual assessment of a child or young person, as might be carried out by an EP using observation as part of a consultation process in everyday practice. This also contributes to the potential difficulty for EPs, particularly new entrants to the profession, to acquire comprehensive skills, knowledge and understanding about contextual observation in practice.

Implications for Practice and Future Research

Since the search period for the literature reviewed in the current paper, there has been recent interest in EP use of observation in a professional context. Speed (2019) reported that there is very little literature relating specifically to the process of observation in EP practice. Speed (2019) also identified some of the same aspects of observation worthy of consideration that the current review also found to be unclear in the literature, (broadly and not exhaustive: defining observation within EP practice; ethics; validity and reliability; structured and unstructured tools; and individual differences between EP approaches to observation). Speed (2019, p. 1) also sets out the importance of knowing more about these things, considering that it is “frequently used”. Furthermore, equality issues were not identified within the reviewed research, and it is suggested that salient equality issues (e.g., gender, ethnicity, race) should also be considered in future research in this area as they have been in other areas of psychological study (e.g., William Labov). The current review provides a timely next step by providing a means of addressing these unknowns by bringing them to the fore through the use and development of the adapted quality evaluation checklist and the bespoke analytic framework. This review has offered further development of knowledge in the area of EP use of contextual observation.

A rigorous two-phase process was used to review and analyse a set of research papers, each of which reported EP use of contextual observation. The two approaches used were an adapted quality evaluation checklist and a bespoke analytic framework. Each afforded its own value in its contribution: the adapted quality evaluation checklist captures the authors’ systematic evaluation of how clearly elements of contextual observation are reported across research evidence reporting EPs’ use of contextual observation in practice, and the bespoke analytic framework reflects a practitioner EP view on what should be considered in relation to contextual observation. The authors identified affordance in combin-

ing the two approaches to contribute to an analytic framework (*The Analytic Framework*, available from the first author on request), that incorporates academic knowledge with practice-informed knowledge. In order to achieve this, additional prompts which reflect the evaluative checklist were incorporated into the analytic framework.

The analytic framework potentially provides a contribution to the evidence base for EP use of contextual observation across the five core functions: research, assessment, consultation, intervention and training, as it identifies key considerations for an EP carrying out contextual observation and lays a foundation for building an understanding of its process.

To develop this into a useful tool for use across the profession, there are opportunities for further research. It is envisaged that this may take the form of an empirical study that uses the analytic framework in its current form as an interview schedule to facilitate individual EPs to evaluate their day-to-day practice of contextual observation. This would provide practice-informed evidence which could be used to develop the analytic framework into initial best practice guidelines with application that is transferable across the five core functions.

It is hoped that the findings from this systematic literature review, and the further research it has set in motion, will promote consistency in EP practice and reduce ambiguity and opacity in communicating the skills of contextual observation. This will allow EPs to (a) emerge from professional training able to apply a logical and systematic approach to contextual observation as part of a wider problem-solving approach and (b) develop contextual observation as part of a wider evidence-based community of practice through further research, peer review and refinement.

References

- Abidin, R. R. (1990). *Parenting Stress Index* (3rd ed.). Pediatric Psychology Press.
- Ahrenkiel, A., Schmidt, C., Nielsen, B., Sommer, F., & Warring, N. (2013). Unnoticed professional competence in day care work. *Nordic Journal of Working Life Studies*, 3(2), 79–96. <https://doi.org/10.19154/njwls.v3i2.2551>
- Apter, B. (2004). *TaMBiS — Teaching and managing behaviour in schools* [Unpublished INSET course materials], Wolverhampton City Council/University of Wolverhampton.
- Apter, B., Arnold, C., & Swinson, J. (2010). A mass observation study of student and teacher behaviour in British primary classrooms. *Educational Psychology in Practice*, 26(2), 151–171. <https://doi.org/10.1080/02667361003768518>
- Bond, C., Woods, K., Humphrey, N., Symes, W., & Green, L. (2013). Practitioner review: The effectiveness of solution focused brief therapy with children and families: A systematic and critical evaluation of the literature from 1990–2010. *Journal of Child Psychology and Psychiatry*, 54(7), 707–723. <https://doi.org/10.1111/jcpp.12058>
- British Psychological Society. (2002). *Professional practice guidelines: A framework for psychological assessment and intervention*.
- British Psychological Society. (2017). *Practice guidelines* (3rd ed.).
- Brown, E. L., Powell, E., & Clark, A. (2012). Working on what works: Working with teachers to improve classroom behaviour and relationships. *Educational Psychology in Practice*, 28(1), 19–30. <https://doi.org/10.1080/02667363.2011.639347>
- Brown, K., & Kennedy, H. (2011). Learning through conversation: Exploring and extending teacher and children's involvement in classroom talk. *School Psychology International*, 32(4), 377–396. <https://doi.org/10.1177/0143034311406813>
- Burt, E., & Stringer, P. (2018). How does the “PACE Maths” approach impact on the practice of school staff? *Educational Psychology in Practice*, 34(3), 245–261. <https://doi.org/10.1080/02667363.2018.1431767>
- Clark, A., & Moss, P. (2001). *Listening to young children, the MOSAIC approach*. National Children's Bureau for the Joseph Rowntree Foundation.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in education* (6th ed.). Routledge. <https://doi.org/10.4324/9780203029053>
- Colville, T. (2013). Strengths-based approaches in multi-agency meetings: The development of theory and practice. *Educational and Child Psychology*, 30(4), 100–123.
- Crombie, R., Sullivan, L., Walker, K., & Warnock, R. (2014). Unconscious and unnoticed professional practice within an outstanding school for children and young people with complex learning difficulties and disabilities. *Support for Learning*, 29(1), 7–23. <https://doi.org/10.1111/1467-9604.12042>
- Day, S. (2010). Listening to young children: An investigation of children's day care experiences in children's centres. *Educational and Child Psychology*, 27(4), 45–55.
- Dockrell, J. E., Bakopoulou, I., Law, J., Spencer, S., & Lindsay, G. (2012). *Communication Supporting Classroom Observation Tool*. https://web.archive.org/web/20210427054722id_/http://www.thecommunicationtrust.org.uk/media/93866/tct_bcrp_csc_final.pdf

- Dunsmuir, S., Brown, E., Iyadurai, S., & Monsen, J. (2009). Evidence-based practice and evaluation: From insight to impact. *Educational Psychology in Practice*, 25(1), 53–70. <https://doi.org/10.1080/02667360802697605>
- Eyberg, S., & Pincus, D. (1999). *Eyberg Child Behaviour Inventory and Sutter-Eyberg Student Behaviour Inventory-Revised: Professional manual*. Psychological Assessment Resources.
- Fallon, K., Woods, K., & Rooney, S. (2010). A discussion of the developing role of educational psychologists within children's services. *Educational Psychology in Practice*, 26(1), 1–23. <https://doi.org/10.1080/02667360903522744>
- Francis, Y. J., Bennion, K., & Humrich, S. (2017). Evaluating the outcomes of a school based Theraplay® project for looked after children. *Educational Psychology in Practice*, 33(3), 308–322. <https://doi.org/10.1080/02667363.2017.1324405>
- Fraser, D. M. (2018). An exploration of the application and implementation of growth mindset principles within a primary school. *British Journal of Educational Psychology*, 88(4), 645–658. <https://doi.org/10.1111/bjep.12208>
- Frederickson, N. (2002). Evidence-based practice and educational psychology. *Educational and Child Psychology*, 19(3), 96–111.
- Gillham, B. (Ed.). (1978). *Reconstructing educational psychology*. Croom Helm Ltd.
- Gray, J., & Sime, N. (1989). Findings from the national survey of teachers in England and Wales. *Discipline in schools: Report of the committee of enquiry chaired by Lord Elton* (pp. 222–250). Department of Education and Science and the Welsh Office. <http://www.educationengland.org.uk/documents/elton/elton1989.html>
- Hayes, B., Hindle, S., & Withington, P. (2007). Strategies for developing positive behaviour management: teacher behaviour outcomes and attitudes to the change process. *Educational Psychology in Practice*, 23(2), 161–175. <https://doi.org/10.1080/02667360701320861>
- Health and Care Professions Council. (2016). *Standards of conduct, performance and ethics*. <https://www.hcpc-uk.org/globalassets/resources/standards/standards-of-conduct-performance-and-ethics.pdf>
- Jolly, M., & McNamara, E. (1992). *Pupil behaviour schedule towards better behaviour*. Towards Better Behaviour.
- Law, C. E., & Woods, K. (2018). The representation of the management of behavioural difficulties in EP practice. *Educational Psychology in Practice*, 34(4), 352–369. <https://doi.org/10.1080/02667363.2018.1466269>
- Leatherbarrow, J. (2020). *An exploration of educational psychologists' use of contextual observation in practice* [Doctoral thesis, University of Manchester]. University of Manchester Research Archive. https://www.research.manchester.ac.uk/portal/files/177286254/FULL_TEXT.PDF
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Rait, S. (2012). The Holding Hands Project: Effectiveness in promoting positive parent–child interactions. *Educational Psychology in Practice*, 28(4), 353–371. <https://doi.org/10.1080/02667363.2012.712916>
- Regan, H., & Howe, J. (2017). Video self-modelling: An intervention for children with behavioural difficulties. *Educational Psychology in Practice*, 33(1), 93–102. <https://doi.org/10.1080/02667363.2016.1233862>
- Roggman, L. A., Cook, G. A., Innocenti, M. S., & Jump Norman, V. (2013). *Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO) user's guide*. Brookes Publishing.
- Sanders, T. (2007). Helping children thrive at school: The effectiveness of nurture groups. *Educational Psychology in Practice*, 23(1), 45–61. <https://doi.org/10.1080/02667360601154600>
- Sedgwick, A., & Stothard, J. (2018). A systematic review of school-based, mainstream, oral language interventions for key stage 1 children. *Support for Learning*, 33(4), 360–387. <https://doi.org/10.1111/1467-9604.12225>
- Speed, E. (2019). The process of psychological assessment: A critique of non-participatory observations within educational psychology practice and the process of psychological assessment. *Educational Psychology Research and Practice*, 5(2), 1–8. <https://doi.org/10.15123/uel.88739>
- Swinson, J., & Knight, R. (2007). Teacher verbal feedback directed towards secondary pupils with challenging behaviour and its relationship to their behaviour. *Educational Psychology in Practice*, 23(3), 241–255. <https://doi.org/10.1080/02667360701507327>
- Swinson, J. (2010). Working with a secondary school to improve social relationships, pupil behaviour, motivation and learning. *Pastoral Care in Education*, 28(3), 181–194. <https://doi.org/10.1080/02643944.2010.504221>

- Vivash, J., Dockrell, J., & Lee, F. (2018). The re-alignment of educational psychologists in supporting primary schools to enhance provision for children with speech, language and communication needs. *Educational and Child Psychology, 24*(2), 77–88.
- Wheldall, K., & Merrett, F. (1984). *Positive teaching: The behavioural approach*. George Allen & Unwin Ltd.
- Woods, K., & Farrell, P. (2006). Approaches to psychological assessment by educational psychologists in England and Wales. *School Psychology International, 27*(4), 387–404. <https://doi.org/10.1177/0143034306070425>

**Appendix
Data Extraction**

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Vivash et al., 2018. Northamptonshire, UK. Six mainstream primary school KS1 classrooms; Professionals involved with SLCN.	To identify gaps between professional perspectives and practice around how SLCN is supported in schools.	Observation. Questionnaires.	KS1 classroom, staff and children.	Classroom environments and interactions. How is SLCN provided for in practice?	Communicating and Supporting Classroom Observation Tool (Dockrell et al., 2012).	Part Contribution. Provided a sample of how SLCN provision is realised in practice, following professionals' views around how it should look.	Medium
Burt and Stringer, 2018. Hampshire, UK. Year 3 and Year 5 pupils in one mainstream school; PACE maths approach.	To evaluate the effect the PACE maths programme has on staff practice and children's participation in learning; to improve effectiveness.	Observation. Focus Group.	Two teachers. Two learning support assistants.	PACE maths sessions. Update and improve the PACE maths programme.	Direct observation using a schedule designed by first author based on key elements of the programme.	Part contribution. Observation identified which aspects of the programme were being used and how effectively to allow the programme to be adapted for the future.	Low

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
<p>Regan and Howe, 2017. Telford, UK University of Birmingham. Six-year-old boy in a mainstream primary school; Video Self Modelling (VSM) intervention.</p>	<p>To measure the impact of VSM in modifying one pupil's challenging behaviour.</p>	<p>Videoed observations.</p>	<p>One six-year-old boy.</p>	<p>Identified target behaviours to provide baseline data. Provide material for the VSM intervention and directly observe measurable changes post-intervention.</p>	<p>Event-based time sampling; Momentary time sampling.</p>	<p>Whole contribution. Observation alone provided all data for the findings of the study.</p>	<p>High</p>
<p>Fraser, 2018. Ellon, UK. A mainstream primary school with embedded growth mindset approach.</p>	<p>To explore growth mindset principles in a primary school.</p>	<p>Semi-structured interviews. Focus groups. Observations in classrooms.</p>	<p>Three class groups of children in morning and afternoon lessons.</p>	<p>Classroom environments, behaviour and approaches to learning. Triangulate data obtained through staff interviews and pupil focus groups.</p>	<p>Narrative style.</p>	<p>Part contribution. Observation contributed to the dataset, which was thematically analysed to provide themes pertinent to the research aim.</p>	<p>Low</p>

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Francis et al., 2017. Leicester, UK. Looked After Children (LAC) from 9 schools, over 8 months.	To evaluate Theraplay as an intervention for LAC.	Consultation with staff. Observation. Pre- and Post-intervention Strengths and Difficulties Questionnaire. Interviews	20 Looked After Children.	Participants in the classroom. Ascertain level of need to inform allocation to research group.	Type of observation not discussed.	Part contribution. Classroom observations described as forming one method of initial assessment. No further information provided thereafter.	Low
Crombie et al., 2014. Blackburn, UK. 3-year research project at a special school.	To generate information about the nature of professional practice in the school.	Observation. Vignettes. Consultations with parents. Staff consultations.	15 children and young people with severe and complex learning difficulties/ disabilities and their teaching/ support staff.	Individual children and staff in classrooms, around school, visits to local area. Behaviour of participants across contexts.	Narrative Observation. Engagement Profiles.	Part contribution. Identified examples of good practice and informed the contexts of the vignettes as staff identified examples of this good practice in day-to-day work.	Low

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Colville, 2013. Scotland, UK. Development and evaluation of strengths-based approaches to multi-agency meetings.	To gain the views of stakeholders of Strengths-Based Multi-Agency Meetings (SBMAs), and the impact of the approach in the local authority.	Observation of meetings. Interviews Focus Groups Analysis of documentary data.	Stakeholders of Strengths-based Multi-Agency Meetings (SBMAs) in the local authority.	SBMAMs Inform the researcher in gaining the views of stakeholders.	Type of observation not discussed: “qualitative data collected” (Colville, 2013).	Part contribution Whole data set relating to views of stakeholders and impact of SBMAs, including observation data analysed using template analysis.	Low
Rait, 2012. Buckinghamshire, UK. Parenting programme: Holding Hands Project.	To evaluate the effectiveness of the Holding Hands Project.	The Eyeberg Child Behaviour Inventory (Eyberg & Pincus, 1999). The Parenting Stress Index, Short Form (Abidin, 1990). A video observational coding scale. Parental Confidence Rating Scale. End of Pilot Questionnaire.	12 parent/child sets.	Parent–child interactions. Provide feedback based on a checklist of 8 observable behaviours.	Observation Schedule (VOCS), adapted from PICCOLO (Roggman et al., 2013).	Part contribution. Data from coding of videos found an increase in positive parent–child interactions between identified time periods.	High

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
E. L. Brown et al., 2012. Aberdeen, UK. Working on What Works (WOWW) approach.	To improve the behaviour and relationships through collaborative working with school staff and pupils. Target Monitoring and Evaluation form (“TME”, adapted from Dunsmuir et al., 2009). Observation of participant children. Consultation with staff and pupils.	Class group of 25 children (aged 5–6 years).	Participants in the classroom. Stage 1: positive things to feedback to pupils. Stage 2: Goals set by children themselves actively looked for during observation period. To improve classroom behaviour and relationships.	“An observation schedule”: child’s name noted alongside space for the observation, and space to indicate which target was being met at the time of the observation.	Part contribution. Data from observation fed back to pupils and class teacher, and used to set and scale new targets for each period of observations.	Medium	
K. Brown and Kennedy, 2011. Surrey, UK. Six teachers and their classes at a residential special school using Video Interactive Guidance (VIG).	To support teachers and pupils develop their use of dialogue to involve children in the learning process.	Discussion groups. Video-recorded interactive sequences in classrooms.	Six teachers and their class group children.	Teachers’ whole-class teaching. Professional development; Identify successes and areas for development in recorded sequences.	Videos coded for “talk types” as identified through discussion groups with staff.	Majority contribution. Videoed classroom interactions formed main data set, but analysis was facilitated by discussion groups.	High

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Apter et al., 2010. Wolverhampton UK. 141 primary classrooms across the UK.	To investigate the nature of verbal behaviours of teachers and its association with how children work.	Observation.	Three randomly chosen pupils per class.	Classroom environments and interactions. “On task” or “off task” behaviour of pupils, and teacher verbal behaviour.	Observation Schedule (TaMBiS, Apter, 2004).	Whole Contribution. Observation alone provided all data for the findings of the study.	High
Swinson, 2010. Area not specified. UK. One 11–18 comprehensive school.	To help revise and improve school’s behaviour policy and practice.	Staff and class meetings (consultations) Questionnaires. Classroom Observations.	Class group children (Year 9, aged 13–14 yrs) before and after policy improvements in the research school.	Pupil behaviour and teacher interactions. Record incidents of disruptive behaviour; positive and negative statements made by teachers; on-task rates.	Observation Schedule devised for the study based on behaviours used by Gray and Sime, 1989.	Part Contribution. Observation provided a comparison of various classroom behaviours before and after new behaviour policy.	Medium

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Day, 2010. Essex, UK. Six children's experience and enjoyment of day care in Children's Centres.	To investigate how a sample of children experience day care in Children's Centres, what they enjoy, and how experiences can be improved for other children.	Age-appropriate pupil voice gathering tools. Interviews. Observation.	Three boys and three girls with age range spanning 20 months to 4 years.	Behaviour and interactions in the nursery environment. Describe what the children were doing and saying; noises, facial expression and body language.	Observation schedule developed for the study.	Part contribution. Observations were coded and contributed to the Mosaic Approach of data collection (Clark & Moss, 2001).	High
Hayes et al., 2007. Kent, UK. 68 classroom observations in one secondary.	To challenge and enable teachers to increase number of positive feedback statements to pupils in classrooms.	Observation. Self-Report Questionnaire. Focus Group.	Teachers and Pupils.	Pupil and teacher behaviour and interactions in the classroom. Types of feedback statements made by teachers and pupils' "on-task" and "off-task" behaviour, before and after staff training for behaviour management.	Observation Schedule (OPTIC observation schedule; Wheldall and Merrett, 1984).	Part Contribution. Identified that the staff training on using a problem-solving framework approach to behaviour management had been successful.	Medium

Authors and research context	Aim of Study	Data Collection	Participants in observation	What was the focus and why?	How was the observation carried out?	Contribution of the observation to the research aim	WoE A
Sanders, 2007. Hampshire, UK. A nurture group pilot project. 3 schools with nurture groups for KS1 and one comparison school without a nurture group.	To investigate whether children in nurture groups make significant social and emotional gains after attending a group.	Staff and pupil interviews. Staff questionnaire. Observation.	Pupils attending KS1 nurture group children (aged 4–7yrs) across 3 schools.	Nurture group pupils during play and academic tasks. Concentration, engagement and interactions.	Naturalistic observations.	Part contribution. Naturalistic observations suggested concentration, engagement and interactions improved.	Medium
Swinson and Knight, 2007. Manchester, UK. Large comprehensive school, 24 Year 8 pupils.	To investigate patterns of teacher feedback towards pupils designated as having behaviour problems and the rest of the class.	Observation.	24 pupils designated as regularly displaying challenging behaviour (18 boys and six girls in Year 8).	Pupil on-task and off-task behaviour; Teacher feedback. Is there a correlation between feedback directed to individuals and individual pupil behaviour?	Observation Schedule (The Pupil Behaviour Schedule, Jolly and McNamara, 1992).	Whole contribution.	High