How can Educational Psychologists evaluate group supervision to ensure positive impact for supervisees’ development and skills, and for the children and families for whom they work?

Authors:

Clare Anderson, Senior Educational Psychologist, Dorset Council
clare.anderson@dorsetcouncil.gov.uk
Sturminster Newton Local Office, Bath Road, Sturminster Newton, Dorset DT10 1DR

Dr Ben Hayes, Associate Professor
b.hayes@ucl.ac.uk
University College London.
ORCID: 0000-0003-2347-9926

Educational Psychology Group, Department of Clinical, Educational and Health Psychology
University College London, 26 Bedford Way, London WC1H 0AP

5005 word count excluding Abstract, references and tables.
How can Educational Psychologists evaluate group supervision to ensure positive impact for supervisees’ development and skills, and for the children and families for whom they work?

Abstract
Aim: This review investigated how group supervision within education or social care settings has been evaluated, to identify tools for EP practice that ensure good outcomes for supervisees and the children and families they support.

Method: A systematic literature search was conducted using databases PsycINFO, Web of Science and ERIC. Twelve studies identified using the inclusion/exclusion criteria involved a range of research designs.

Findings: The selected studies suggest group supervision can benefit supervisees in terms of well-being, on direct practice with families, and on team relationships. The most innovative and revealing studies were small-scale and conducted in the context of social work. These highlighted the importance of having a ‘clinician’ involved in facilitation for practice change, and group supervision characterised as ‘practice-focused’ was found to improve parental engagement and goal agreement.

Limitations: There is a paucity and lack of breadth in methodological approaches in evaluating the impact of supervision, and challenges in data collection. Further research should incorporate triangulation of information (i.e., supervisee, supervisor, and client).

Conclusions: Evaluation tools relevant for EP use include making use of data sources such as supervision records, using a coding framework for self-evaluation and using questionnaires to assess how supervision has achieved the intended outcomes for supervisees. A subsequent follow-up of supervisee’s intended actions to identify practice change, should include gathering feedback directly from children and families.
Keywords: Supervision; evaluation; social work; educational psychology

Introduction
Multi-disciplinary supervision practices are increasing and offer ‘many opportunities for Educational Psychologists (EPs) to develop and enhance their supervisory skills’ (Dunsmuir & Leadbetter, 2010, p.12). The British Psychological Society (BPS) Division of Educational and Child Psychology (DECP) ‘Professional Supervision: Guidance for Educational Psychologists’ proposes a self-assessment ‘competencies framework’, which can be further supported by ‘feedback from supervisees, colleagues and others’ (Dunsmuir & Leadbetter, 2010). However, using this tool still leaves unanswered the question of impact on supervisee’s well-being, development and on their practice with clients. With increasing concerns regarding well-being, workload stress and capacity within public services, it is particularly important that EP supervisors consider the ‘ethical responsibility for all individuals to acquire competence in supervision’ (Dunsmuir & Leadbetter, 2010) to ensure that supervision improves outcomes for supervisees and for those they serve. Practitioner Psychologist Standards of Proficiency (Health and Care Professions Council; HCPC, 2015) require psychologists to ‘reflect on and review practice’ and ‘to be able to assure the quality of their practice’, and evaluation is one of six key competences in the BPS guidance (Dunsmuir & Leadbetter, 2010).

Characteristics of effective clinical supervision summarised in systematic reviews (Carpenter, Webb & Bostock, 2013; Rothwell et al., 2019) include positive outcomes for supervisees in terms of job satisfaction, self-efficacy and stress, and for organizations such as workload management, case analysis and retention. However, Carpenter, Webb and Bostock (2013) did not find any study evidencing impact for the supervisee’s clients. An overall appraisal of systematic reviews conducted between 1995 and 2019 raised significant concerns about weak existing research on supervision impact, models, methodological issues and evidence-based
supervision and highlighted the need for ‘better and broader evidence’ (Watson, 2020; p.205). There has been consistent criticism that the available research evidence is commonly based on correlational data which doesn’t evidence causality (Wheeler & Richards, 2007). However, interest and commitment to developing evidence-based practice in supervision is growing. An edition of Educational and Child Psychology journal was dedicated to the topic in 2015, noting the dearth of research (Leadbetter et al., 2015, p.7). Additionally, a special journal edition of the Australian Psychologist was dedicated to professional supervision in 2017, however there were no specific articles relating to the use of group supervision.

**Group supervision**

Group supervision is utilised within health (e.g., Rothwell et al., 2019), social work (e.g., Alschuler et al., 2015) and psychology (e.g., Enyedy et al., 2003; Ögren & Sundin, 2009). In facilitation of group supervision, EPs draw on a range of psychological models and processes including problem-solving frameworks and ‘therapeutic methods informed by psychodynamic, cognitive behavioural and humanistic approaches’ (Dunsmuir et al., 2015; p.14). Mastoras and Andrews (2011) note that ‘there is no clear definition, or single approach’ (p.108) however ‘the voices of supervisees were remarkably consistent in some of their perceptions and experiences of group supervision’ (p.108). Reported benefits are shown in Diagram 1 below. Group supervision can be an effective space to ‘develop, enhance, and sustain the reflective stance of practitioners’ therefore emulating attuned interactions central to children feeling safe, secure and emotionally contained by carers; experiencing this process themselves helps practitioners support families (Heffron et al., 2016). Group supervision requires supervisors to foster an emotionally safe space (Fleming et al., 2010) and to manage common group processes such as ‘between member problems’ (Enyedy et al., 2003), diversity/coherence or group think (Munroe, 2008). Barriers identified in research on group supervision are shown in Diagram 2 below.
Diagram 1: Supervisee reported benefits of group supervision.

Diagram 2: Perceived barriers to effective group supervision

Group supervision may involve a team where there are existing relationships amongst workers and a group culture or ethos (e.g., a residential setting), a group of practitioners within the same
organisation (e.g., a locality) or professionals with a common background or role but who do not work together e.g., Emotional Literacy Support Assistants (ELSAs). Contextual factors have been found to impact on how supervision is perceived by supervisees; for example, Callicott and Leadbetter (2013) found that there were issues arising related to the cultures and histories of the individuals and organisations involved in supervision leading to different expectations. Supervision experiences vary depending on whether a supervisor has line management responsibilities for supervisees; supervisees might value the ‘neutrality’ of a supervisor who is not in management role, however not belonging to the same organisation might reduce empathy for the supervisees (Callicott & Leadbetter, 2013; p.393). Further contextual issues include location, space, and timings, with some supervision now being facilitated virtually. Therefore, what might make group supervision ‘effective’ for one group may not be so for another and the way the EP supervisor introduces, structures, manages, and evaluates sessions and uses different frameworks will depend on the context, group dynamics and purpose of supervision (Mastoras & Andrews, 2011; p.109).

Both individual and group supervision can be equally effective in terms of impact on practice (Ray & Altekruse, 2000). One study, conducted over two years, compared the perceptions of trainee social workers of their experiences of either individual or group supervision and found no significant differences in perceptions of impact on interventions with clients, internalization of professional values, or general satisfaction with field instructor and field practice (p.342). However, supervisees had less favourable perceptions of the content of group supervision and of their relationships with facilitators compared with those who received individual supervision (Zeira & Schiff, 2010). Group supervision can also necessitate additional training that increases time and expense, which may be counter to assumptions made about the cost effectiveness of this approach (Zeira and Schiff, 2010). The format of supervision should therefore be chosen.
by supervisees rather than imposed upon them (Hawkins & Shohet, 2012), and both models should be offered as they provide complementary experiences (Ray & Altekruse, 2000). Drawing on the parallels with child development (see Heffron, 2016) Zeira and Schiff (2010) suggest that trainees might require one-to-one supervision when developing their skills and ‘group supervision may therefore be more suitable for the more advanced years of study or a later stage of professional development’ (p.432).

**Evaluation of supervision in practice**
It is challenging to evidence the impact of ‘indirect psychological practices’ such as supervision and consultation (Dunsmuir et al., 2015), and even more problematic evaluating outcomes for clients (e.g., child, family, students, colleagues). Alongside limited empirical research there also appears to be a lack of systematic use of evaluation of supervision in the context of EP practice, as well as in other professions including social work (Davys et al., 2017). In one study most workers felt that it was ‘best practice’ to evaluate the impact of supervision and cited a range of tools used to do this, however the most frequent focus of evaluation was ‘whether reflection is occurring in supervision’, thus exploring the content rather than the outcomes (Davys et al. 2017, p.114). In evaluating supervision Rothwell and colleagues (2019) suggest there should be a ‘feedback loop’ to supervisors which could involve multiple sources of information from supervisor, supervisee, a peer / external consultant, the clients with whom the supervisee practices (Vonk & Thyer, 1997) and the organisation for which the supervisees work (Wonnacott, 2012).

**The present review**
The aim of this systematic review is to identify evaluative tools that can be used to ensure that group supervision facilitated by EPs is effective in creating ‘positive change in the supervisee’s behaviours, level of knowledge and/or professional values (and)...also contribute to improvement in the supervisee’s clients’ functioning’ (Vonk & Thyer, 1997, p.105). The
review focuses on the following question: How do professionals evaluate the impact of group supervision for supervisees, and for their clients?

Method

Literature Search
The databases Web of Science, PsycINFO and Education Resources Information Center (ERIC) were systematically searched, and Google Scholar was also used to identify any additional relevant studies. The search process is illustrated in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram (Page et al., 2021) below. Key words used for the main searches in each database were as follows: “group supervision” OR “team supervision” OR “peer group supervision” AND “Evaluator*” OR “assess*” OR “review*”.

Search strategy
Studies were included if they involved evaluating group supervision in the context of practitioners working in social work or education, and identified the impact for supervisees or for the children and families with whom they work. Research involving student / trainee participants was excluded. The search scope was not restricted by research design (qualitative or quantitative) or by model of supervision, to avoid missing relevant evaluation tools. The literature search included studies evaluating group supervision where there was use of a supervisor / facilitator (peer supervision studies excluded). For full information about the inclusion / exclusion criteria, see Table 1 below.
PRISMA (Page et al., 2021) flow diagram for new systematic reviews which included searches of databases, registers and other sources

Identification of studies via databases and registers
- Records identified from:
  - PsycINFO (n = 63)
  - ERIC (n = 103)
  - Web of Science (n = 199)
- Records removed before screening:
  - Not peer reviewed (Reason 1)
  - PsycINFO (n = 14)
  - ERIC (n = 49)
  - Web of Science (n = 7)
- Records screened (n = 295)
- Records excluded** (n = 0)
- Reports sought for retrieval (n = 295)
- Reports assessed for eligibility (n = 295)
- Reports excluded:
  - Reason 2 (n = 10)
  - Reason 3 (n = 34)
  - Reason 4 (n = 28)
  - Reason 5 (n = 84)
  - Reason 6 (n = 2)
  - Reason 7 (n = 10)
- Studies included in review (n = 12)
- Reports of included studies (n = 12)

Identification of studies via other methods
- Records identified from:
  - Citation searching (n = 2)
- Reports sought for retrieval (n = 2)
- Reports not retrieved (n = 0)
- Reports assessed for eligibility (n = 2)
- Reports excluded (n = 0)
Data extraction and analysis
In order to minimise researcher bias, Gough’s (2007) Weight of Evidence framework was used as a transparent process to appraise and synthesise the twelve selected studies (Table 2). Analysis involved consideration of both the quality and relevance of the data to the specific question posed in this literature review. Furthermore, published coding protocols were used in making judgements of the quality of the research methodology (Weight of Evidence A). For three correlational studies the quality indicators suggested by Thompson, Diamond, McWilliam, Snyder and Snyder (2005) were used, providing standards regarding measurement, quantifying effects, avoiding common analysis errors, and using confidence intervals. The guide for appraising survey reports by Burns and Kho (2015) was utilised to appraise survey methods and use of a questionnaire. To assess qualitative research designs a coding protocol designed to appraise research in the field of special education was used (Brantlinger, Jimenez, Klingner, Pugach & Richardson, 2005).

Studies were compared by calculating the percentage of the quality indicators met for each study; percentage ranges were assigned nominal values to allow comparison across Weight of Evidence areas. Weight of Evidence B (WoE B) evaluated how appropriate the methodologies of the studies were for answering the review question (Gough, 2007); studies were rated most highly where the research design involved evaluation of outcomes for both supervisees and their clients. Weight of Evidence C (WoE C) examined the topical relevance of studies to the review question. The most relevant studies provided detailed information about evaluative tools to enable replication in EP practice. The studies were summarised and grouped according to context, research design, participants, measures and findings (see Table 2). The three Weight of Evidence scores calculated together provide an overall WoE D score. The ratings assigned to each study across all areas of Weight of Evidence can be found in Table 3.
<table>
<thead>
<tr>
<th>Criteria</th>
<th>Inclusion</th>
<th>Exclusion</th>
<th>Rationale</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Type of publication</td>
<td>Peer reviewed journal article.</td>
<td>Non-peer reviewed e.g., dissertation thesis, book chapter.</td>
<td>Peer reviewed studies have been subjected to scrutiny by relevant experts in the field to ensure quality and reliability.</td>
</tr>
<tr>
<td>2) Language</td>
<td>Published in English.</td>
<td>Not published in English.</td>
<td>Author not able to read in other languages and to obtain translation.</td>
</tr>
<tr>
<td>3) Research design</td>
<td>The research study must be original research seeking to evaluate the quality or effectiveness of group supervision for participants and their clients.</td>
<td>Studies that are exploring conceptual or theoretical models for supervision or processes within it.</td>
<td>This review aims to explore tools and measures used to evaluate the effectiveness of group/team supervision as experienced by supervisees and their clients.</td>
</tr>
<tr>
<td>4) Model of group supervision</td>
<td>Studies are included where the format of group supervision involves having an external facilitator / supervisor.</td>
<td>Supervisors are supervisees e.g., peer supervision with no external facilitator.</td>
<td>This study is focused identifying the methods that have been used to evaluate group supervision facilitated by a supervisor.</td>
</tr>
<tr>
<td>5) Context</td>
<td>Group supervision within social work or educational settings.</td>
<td>Context of supervision is with health professionals or another specific role that is not within social care (children and families) or education.</td>
<td>The review focus is on understanding how to facilitate group supervision effectively in role as EP supervisor.</td>
</tr>
<tr>
<td>6) Participants</td>
<td>Participants include supervisees (education or social care roles) or the children and families they support.</td>
<td>Participants solely are Supervisors.</td>
<td>The aim of the study is to inform Educational Psychologists who facilitate group supervision in education and social care settings.</td>
</tr>
<tr>
<td>7) Availability</td>
<td>Study accessible for review (available full text online)</td>
<td>Study not accessible online.</td>
<td>All research studies need to be available for review.</td>
</tr>
<tr>
<td>8) Focus of study</td>
<td>Original research focused on group supervision.</td>
<td>Study is not relevant, not original research or not about supervision.</td>
<td>Research limited to gathering research aimed at evaluating the impact of group supervision on supervisees and their clients.</td>
</tr>
<tr>
<td>Study</td>
<td>Aim</td>
<td>Context</td>
<td>Research design</td>
</tr>
<tr>
<td>-------</td>
<td>-----</td>
<td>---------</td>
<td>----------------</td>
</tr>
<tr>
<td>1</td>
<td>To examine the effects of structured group supervision (SGS) on counsellors’ self-efficacy, counselling competency, and job involvement in Singapore.</td>
<td>Counselling: 6 3-hour group supervision sessions following Structured Group Supervision approach SGS (1x 2 weeks) 12-week period, Singapore.</td>
<td>Correlational; A single-group, before-after design was adopted. Pre- and post-test questionnaires were administered.</td>
</tr>
<tr>
<td>2</td>
<td>Evaluation of a model of systemic group supervision or “systemic supervision” based on a wider evaluation of systemic social work practice in the UK.</td>
<td>Social work; Five English local authority children’s services departments who redesigned their child welfare provision in line with a systemic unit model (Reclaiming Social Work; RSW).</td>
<td>Correlational design; data on the relationship between supervision quality and direct practice quality to assess whether there is an association between the two practice forums.</td>
</tr>
<tr>
<td>3</td>
<td>To examine ELSAs’ views on group supervision: the extent to which supervision was meeting ELSAs’ needs, relationship with supervisor and group members, advantages and disadvantages of group supervision, and impact of supervision on practice (perception of).</td>
<td>Teaching Assistants trained as ELSAs in the UK being supervised in group by EPs.</td>
<td>Qualitative; Questionnaire study.</td>
</tr>
<tr>
<td>4</td>
<td>To examine the views of teaming mentors on the outcomes of group supervision; enablers and barriers to attending and key characteristics of groups who might benefit from group supervision.</td>
<td>Learning mentors experience of group supervision. The group has been running for approximately three years. Based in one school in the UK.</td>
<td>Qualitative; Single case study methodology, using the following methods: a focus group, review of written records of attendance and content, and questionnaires evaluating the sessions after a year.</td>
</tr>
<tr>
<td>5</td>
<td>To explore the use of group supervision with staff supporting young people with social, emotional and</td>
<td>Specialist educational setting: Educational Psychologists led fortnightly group supervision sessions to promote collaborative problem solving</td>
<td>Qualitative; focus group</td>
</tr>
<tr>
<td></td>
<td>Authors and Year</td>
<td>Title</td>
<td>Methodology</td>
</tr>
<tr>
<td>---</td>
<td>------------------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td>10</td>
<td>Soni, A. (2013)</td>
<td>Group Supervision: Supporting Practitioners in Their Work with Children and Families in Children’s Centres. Early Years: An International Journal of Research and Development, 33, 2, 146 – 160.</td>
<td>To explore how group supervision can be used to support the Continuing Professional Development (CPD) of those working with children and families in early years provision in England. Early years education staff: Group supervision every 6-8 weeks, 90 mins. Cooperative approach to group supervision – facilitator and group co-supervise the sessions. UK. Qualitative; Realistic evaluation. Semi-structured interviews. Three Children’s Centre managers and the 12 Family Support Workers. Realistic, semi-structured interviews Importance of mechanisms including the professional contract with managers and group working agreement.</td>
</tr>
<tr>
<td>11</td>
<td>Bostock, L., Patrizo, L., Godfrey, T., Munroe, E. and Forrester, D. (2019)</td>
<td>How do we assess the quality of group supervision? Developing a coding framework. Children and Youth Services Review, 100, 515 – 524.</td>
<td>To develop a framework for evaluating quality of group supervision. Social work: Consultant social workers were supervisors (where available supported by clinician trained in systemic family therapy) and social workers supervisees. UK. Qualitative study – observations, focused interviews. 29 observations of ‘live’ supervision sessions for social workers. 5 LA in England developing a systemic social work practice model (Reclaiming Social Work RSW). 185 practitioner participants over 29 observations (4-8 in a group).</td>
</tr>
<tr>
<td>12</td>
<td>Willis, J. and Baines, E. (2018)</td>
<td>The Perceived Benefits and Difficulties in Introducing and Maintaining Supervision Groups in a SEMH School. Educational Review, 70, 259 – 279</td>
<td>To examine the perceived benefits and difficulties of introducing and maintaining effective group supervision. One school for SEMH in the UK. Case study – one SEMH school context. 6 teachers, 10 TAs and 1 office manager 3 supervision groups (5/6 in a group). Supervisor is Psychotherapist. Use of semi-structured one-to-one interviews by 2 independent research assistants. The supervisees provide evidence of how group supervision has had a notable, positive impact in improving their interactions with colleagues and on the management of their stress.</td>
</tr>
</tbody>
</table>
Results
Twelve studies were selected for review and appraisal; three correlational, one survey, and eight qualitative studies. Two studies achieved the highest overall (WoE D) appraisal scores; one (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019) scored highest for methodological quality (WoE A) and the other (Bostock, Patrizo, Godfrey & Forrester, 2019) received the highest rating for topic relevance (WoE C); this was the only study that attempted to explore the links between quality of systemic group supervision and quality of social work practice. Only three studies attempted to assess the impact of supervision on supervisee’s practice with their clients (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019; Bostock, Patrizo, Godfrey & Forrester, 2019; Wilkins et al., 2018) and as such, these studies have been particularly informative in answering the review question. The study reported by Tan and Chou (2018) was the only other study that received the highest rating for WoE A regarding methodological quality; a single-group, before-after design was utilised, allowing comparison of scores on three published self-report scales measuring counselling competence, job involvement and counselling self-efficacy. The authors made use of an external supervisor to try to minimise study bias or conflict of interest and were explicit about limitations of the study and findings.
Table 3: Weight of Evidence appraisal summary

<table>
<thead>
<tr>
<th>Participant</th>
<th>WoE A: Methodological Quality</th>
<th>WoB B: Methodological Appropriateness</th>
<th>WoE C: Topic relevance</th>
<th>WoE D Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bartle &amp; Trevis (2015)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Bostock, Patrizo, Godfrey, Munroe &amp; Forrester (2019)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>Bostock, Patrizo, Godfrey &amp; Forrester (2019)</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>France &amp; Billington (2020)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Iwasaki, Watanabe &amp; Tamura (2017)</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Osborne &amp; Burton (2014)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Soni (2013)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Soni (2015)</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Tan (2019)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Tan &amp; Chou, (2018)</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>Willis &amp; Baines (2018)</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Wilkins, Lynch &amp; Antonopoulou (2018)</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Participant characteristics
This review focuses on research conducted with participants practising in either education or social care settings. Participants were Emotional Literacy Support Assistants (ELSAs; France & Billington, 2020; Osborne & Burton, 2014), learning mentors (Soni, 2015), teachers (Bartle & Trevis, 2015; Willis & Baines, 2018), social workers (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019; Bostock, Patrizo, Godfrey & Forrester, 2019; Wilkins et al., 2018), school
counsellors (Tan & Chou, 2018; Tan, 2019), family support workers in early years settings (Soni, 2013) and ‘Yogo teachers’ similar to school nurses (Iwasaki et al., 2017). Most studies involve small samples (between 5 and 270), therefore limiting generalisability. The participant sample (n=5) reported by France and Billington (2020) is particularly small, however qualitative studies such as these are valuable because they provide insight into the lived experiences of participants. Supervision is a process that encourages self-reflection and practice development by supervisees; such outcomes are challenging to measure using quantitative methods.

**Study design**
Most of the selected studies involved qualitative methodology to explore supervisee’s perceptions (Bartle & Trevis, 2015; France & Billington, 2020; Iwasaki et al., 2017; Osborne & Burton, 2014; Soni, 2013; Tan, 2019; Tan & Chou, 2018; Willis & Baines, 2018). Correlational designs exploring the relationships between supervision and practice outcomes (Bostock, Patrizo, Godfrey & Forrester, 2019; Tan & Chou, 2018; Wilkins et al., 2018) were the most methodologically and topically relevant to the question explored in this review. Although the causal nature of the associations cannot be inferred, this research design offers a more practical solution for EP practice because setting up the conditions for an experimental design present greater ethical and logistical challenges. Realistic Evaluation methodology (Pawson & Tilley, 1997) was used by Soni (2013), to explore not only the outcomes of group supervision, but the mechanisms involved in achieving these, which is very relevant for research on group supervision because as discussed previously, every group, context and model will be different. However, this study did not involve evaluating the outcomes of supervision, therefore achieving lower scores for Weight of Evidence B.

**Supervision outcome measures**
The selected studies evaluated the impact of group supervision on supervisees using surveys (Osborne & Burton, 2014), focus groups using semi-structured questions (Bartle & Trevis,
2015; Iwasaki et al, 2017; Soni, 2013; Soni, 2015; Tan, 2019), document analysis (Soni, 2015), observations (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019; Bostock, Patrizo, Godfrey & Forrester, 2019; Wilkins et al, 2018) and standardised scales and questionnaires (Tan & Chou, 2018). Bostock, Patrizo, Godfrey and Forrester (2019) used a highly rigorous methodological approach to minimize bias and established a system for assessing quality of supervision and quality of practice. A coding framework involving a three-point ordinal scale was used in three studies to assess supervision (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019; Bostock, Patrizo, Godfrey & Forrester, 2019; Wilkins et al., 2018). However, the data collection process involved in these studies, using live observations of group supervision sessions and of social workers in practice with families during home visits would be unrealistic in EP practice due to the time and costs in terms of observers required.

Tan and Chou (2018) made use of standardised questionnaires to evaluate the impact of group supervision on aspects of supervisee’s emotional well-being and skills for example ‘counselling self-efficacy’, ‘job involvement’ and ‘counselling competency’. These measures were specifically for use with counsellors however, Educational Psychologists could make use of other self-report measures in the process of evaluating the impact of group supervision on their supervisees. Using standardised tools allows more confidence in the perceived competencies being measured however, it is not possible to be certain as to whether supervision, as opposed to a different variable, has led to the perceived changes. There are also inherent limitations with using self-report measures, for example social desirability, selective recall, or other biases (Althubaiti, 2016). Tan and Chou (2018) measured changes over a period of only 12 weeks, therefore not capturing longer-term impact. It is difficult to consider how such tools could be used to assess the impact of supervision regularly as the constructs may be too broad to detect small shifts. Furthermore, using published scales and questionnaires alone
does not provide any information about what aspects of the supervision were helpful (or not helpful) therefore not providing supervisors with information to support their professional development. Osborne and Burton (2014) designed a questionnaire specifically for ELSAs. Limitations of this study are again those inherent to using self-report measures, and this questionnaire was designed to evaluate supervision more generally, as opposed to using it to evaluate individual sessions or the impact of these over time.

The study reported by Soni (2015) made use of three methods to elicit qualitative data from supervisees; focus groups, questionnaires, and analysis of the recorded notes from group supervision. To analyse the notes, Soni (2015) used a coding system to explore the main functions (based on Hawkins & Shohet, 2006 and Proctor, 2012); this method provides a means of exploring the content, process, or quality of supervision.

**Discussion**

This literature review provides further evidence of the value of group supervision in education and social work settings. In terms of benefits for individual supervisees, group supervision has been found to bring the following benefits: ‘authority-based’ and ‘relationship-building’ skills (Bostock, Patrizo, Godfrey & Forrester, 2019), counselling self-efficacy and competency (Tan & Chou, 2018), new ideas (France & Billington, 2020; Osborne & Burton, 2014), ‘sharing good practice’ (Soni, 2015), ‘shared emotional experiences’ and ‘therapeutic effects’ (Willis & Baines, 2018) and improving ‘personal and professional development’ and increased confidence (Osborne & Burton, 2014). There also appears to be positive impact for the group or team itself including improving team communication and coherence (Bartle & Trevis, 2015),
team relationships and peer support (Soni, 2013), helping supervisees feel more connected and experience more support and help (Tan, 2019).

Factors that facilitate effective group supervision include support from management to attend (Soni, 2015) and protected time (Bartle & Trevis, 2015), having a ‘relaxed informal approach’ (Soni, 2013) and ‘sharing of ideas, experiences and resources’ (Osborne & Burton, 2014). Furthermore, there is evidence that the inclusion of ‘clinicians’ (trained in family therapy) within systemic group supervision (Bostock, Patrizo, Godfrey & Forrester, 2019) and having an effective ‘external independent supervisor’ (Willis & Baines, 2018) might increase the effectiveness in some contexts; these factors are particularly salient in the context of Educational Psychologists as supervisors to other professional groups. The study by Bostock, Patrizo, Godfrey and Forrester (2019) found a significant relationship between high quality supervision and overall quality of direct practice of social workers. As this is the only study of its kind attempting to evaluate both quality of supervision and practice, there is a need for this to be replicated and extended.

The findings will now be considered in relation to the review question:

*How can EPs evaluate the impact of group supervision for supervisees, and for the children and families for whom they work?*

The three studies that received the highest quality appraisal ratings used coding procedures to assess both the quality of supervision and practice with families, allowing exploration of associations between these (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019; Bostock, Patrizo, Godfrey & Forrester, 2019; Wilkins et al., 2018). Coding frameworks for supervision practice described by Bostock, Patrizo, Godfrey and Forrester (2019) and Wilkins et al. (2018)
were formulated by practitioner ‘experts’, and then used in live observations of group supervision and subsequent quality ratings applied. Alongside these, coding protocols based on published standards in social work practice were used to evaluate practice, to explore associations between the quality of supervision and subsequent work with families.

Bostock, Patrizo, Godfrey and Forrester (2019) describe a model of systemic supervision that allows supervisees to plan, rehearse and practice having such conversations before they meet again with families, and found evidence that this model of group supervision improves ‘practice-making’ (Bostock, Patrizo, Godfrey & Forrester, 2019). Supervision therefore might usefully incorporate elements of coaching, focusing on strengths observed in the interaction performed in a safe space, and supporting the supervisee to develop confidence and skills. Wilkins et al. (2018) explored the difference between ‘practice-focused supervision’ involving a structured discussion focusing on client goals and action planning and ‘other-focused supervision’ which ‘tended to be characterized by a lack of structure, a focus on the past and a superficial discussion of problems and issues’ (p. 7). This study identified a positive association between practice-focused supervision and improved social work skills, ‘particularly the use of good authority (purposefulness, clarity about risk, and child focus)’ (Wilkins et al., 2018; p.500). The model of supervision described as ‘practice-focused’ involves exploring explicitly what (goal), why (hypotheses) and how (practice) social workers are going to work with families following supervision. Feedback regarding the impact of the supervision on practice was gathered by using parent questionnaires; this offers a potential way forward for evaluating group supervision for social workers, and a similar process could be used with different service users for education supervisees (e.g., other teaching staff, parents) or psychologists (e.g., parents, school staff, colleagues).
Implications for EP practice
Interestingly, the selected studies highlight the value of having an external facilitator compared with peer supervision groups where all group members co-facilitate, providing support for EP facilitated group supervision. Bostock, Patrizo, Godfrey and Forrester (2019, p.7) comment that ‘even within a group-based format, it appears important that there is a leader who is helping the group with the task of turning hypotheses into actions and rehearsing conversations’. Another important role for the supervisor is to ensure effective use of the time; Soni (2015) found that the educative function dominated the discussions between learning mentors, suggesting that ‘it is also important for the facilitator to encourage the group to consider the supportive’ (Soni, 2015; p.74). Further research exploring other facilitators to effective group supervision such as the supervisor’s specific skills and approaches valued most by supervisees is needed.

EP facilitators can make use of several possible methods to evaluate the outcomes of group supervision. To assess whether supervision is meeting supervisees needs (e.g. formative, normative and restorative; Proctor, 2012), the initial group agreement should always involve discussion with group members of the desired overall purpose of the supervision so that as part of the group contracting, a set of functions and intended outcomes for supervisees could be agreed and subsequently monitored each session and longer term by both the supervisor and supervisees using self-report measures. Using standardised measures to track impact on supervisee’s self-efficacy, well-being, or motivation for their role (see Tan & Chou, 2018) is a time-efficient approach that provides a quantitative assessment of impact, and facilitates exploration of correlations between variables such as frequency of supervision (attendance), model of practice or quality of supervision with a range of outcomes for supervisees.
To evaluate the quality of supervision, questions such as used in the questionnaire to gather ELSA views on the role of the supervisor (Osborne & Burton, 2014) could be adapted for regular evaluation of group supervision as part of EP practice. These involved rating scales to assess how much supervisees agreed with the statements e.g. ‘To what extent does your supervisor… help you to clarify objectives in working with children involved in ELSA’.

Qualitative feedback could also be gathered from supervisees, so that supervisors gain feedback regarding their practice and ways that they could improve. In the absence of external observers (Bostock, Patrizo, Godfrey and Forrester, 2019; Wilkins et al., 2018) with consent from supervisees supervision sessions could be recorded using video feedback, for regular review by the supervisor perhaps in the context of their own individual supervision as a coaching tool to notice strengths as well as areas for development.

Recording supervision sessions is an important part of the supervisor role (Dunsmuir & Leadbetter, 2010) and using the ‘domains of systemic group supervision’ for example (Bostock, Patrizo, Godfrey, Munroe & Forrester, 2019) as a format to structure notes for each case discussion, might simultaneously prove helpful as a means of evaluating the quality of discussions. Evaluating documents such as meeting notes / records, to assess the content of supervision (Soni, 2015), could be conducted by EPs in practice to explore the themes or functions emerging, and monitoring over time whether these remain static or change. Analysing data such as this might indicate for example, when the supervision model should be changed to encourage more diversity in the issues brought to supervision by highlighting when it being used mostly for one purpose so that the balance of all intended functions or outcomes is addressed.
Evaluation of the impact of supervision on practice could be achieved by ensuring goal-focused discussions and rehearsal takes place during supervision (Bostock?????), and subsequently strategies or actions planned by supervisees could be followed-up using questionnaires designed to gather information from parents and children about whether that particular action was helpful to them. Data could be gathered from multiple sources regarding supervisee’s use of strategies discussed in supervision (Bartle & Trevis, 2015).

Finally, having the skills and confidence to be able to tailor supervision to the needs and context of different group members or contexts has implications for training; ensuring that EPs have access to appropriate professional development opportunities should be an important consideration for EP services.

**Limitations of the review**
The literature search was conducted systematically, however studies not written in English were excluded therefore the search is not representative of international practice and important research may have been missed. Limiting the search to peer-reviewed articles means that gray literature including doctoral theses were excluded, which means further relevant research may not have been included and raises the possibility of publication bias where counter evidence may not be included. The main author is from and working in the UK, and most studies in the review are based in UK contexts. While this is a strength in terms of implications for educational practice in the UK, it means that construction of the inclusion and exclusion criteria may have been influenced by use of terms that are in common use in the UK and may not have used terms for group supervision that describe this process in other countries. It would therefore be useful to search the literature and evaluate studies published in other languages. Finally, this review only considered group supervision practice in the context of educational and social work.
practice. There might be some innovative ideas and relevant evaluation tools in the context of other professions that could be of use to Educational Psychologists.

**Conclusion and future directions**

Very few studies have explored the relationship between group supervision and subsequent work with children and families. This is surprising considering the intuitive value of quality supervision on practice and drive for outcomes-based work throughout children’s services and education. In addition, this review has not identified any studies that have focused on the impact of group supervision on team relationships and effectiveness. Research conducted through EP practice could follow a Realistic Evaluation methodology such as Soni (2013) used, to explore both the ‘mechanism’ (process of supervision) and context as these are both important in the outcome (p.150).

The dominance of self-report measures and correlational studies means that it is not possible to be certain about causality and a further limitation is that self-report measures can be at risk of honesty and insight biases. None of the studies reviewed used an evaluation measure that could be used after each supervision session; using such a tool would provide information about how the content, process or facilitation of sessions changes over time, about the associated impact on supervisees and inform future supervisory practice.

The creative research design reported by Bostock, Patrizo, Godfrey and Forrester (2019) explored the direct impact of group supervision on social workers and their families and highlights how this might be possible in EP practice. This small-scale exploratory study found a strong association between the quality of supervision and resulting practice with families (Bostock, Patrizo, Godfrey & Forrester, 2019). However, it would have been useful to assess
social work practice before the systemic supervision had started, to explore pre- and post-supervision practice quality; it is possible that some of the practitioners who demonstrated highest quality practice were already doing so, and as the authors highlight, the social workers recruited the family participants, which may have contributed to selection bias. Furthermore, qualitative information from social workers exploring their perceptions of the supervision to identify the aspects found most useful (i.e., the ‘how’), would be helpful. Future research therefore should involve triangulation of methodology or data collection so that information is gathered from different sources to add strength to the findings.

This review has identified evaluation tools relevant for EP use. Data sources such as supervision records are readily available and can be used to explore how the time has been used to address the agreed functions of supervision, which could be agreed at the outset. Developing coding frameworks also offers potential for self-evaluation and for capturing practice change, perhaps by using video feedback of practice and/or gathering feedback directly from children and families to follow-up supervisee’s intended actions. Self-report questionnaires are a practical tool used to assess how supervision has achieved the intended outcomes for supervisees and could be used alongside other methods to overcome some of the limitations.

The evaluation of group supervision and the outcomes for supervisees and clients has been under-researched in psychology, and it is yet to become an integral part of supervisory practice despite recognition of the importance of this. This review has highlighted the need for further research exploring how supervision impacts on supervisees in terms of well-being, skill practice and their effectiveness within teams. Questions remain regarding how group dynamics change, what aspects are most useful, the skills supervisor require to facilitate effective sessions
and which functions of supervision appear to be most frequently addressed. Despite the absence of published tools to evaluate group supervision, this review has highlighted creative attempts to gather impact data from supervisees and their clients. Supervision is a dynamic, fluid process, therefore following a model of ‘double loop learning’ (Argyris & Schön, 1996) involving a process of ‘do’, ‘learn’ ‘review how we are doing’ and ‘review why we are doing what we do/are we doing the right things?’ is recommended. To create a safe climate for supervisees to learn from each other, there is a need to maintain a reciprocal dialogue about the process of supervision and seeking ways of demonstrating impact on practice, for the benefits of supervisees, their clients and for EP professional development.

This article is the original work of the author and not under consideration elsewhere.
References


Maxwell, T. (2013). A reflection on the work of an Educational Psychologist in providing supervision for a team of community based support workers, supporting families with


