

## ***Case Study 1: An Evidence-Based Practice Review Report***

***Theme: How effective is outdoor education (including forest school) at improving social and emotional skills in pupils with Special Educational Needs?***

### **1. Summary**

Forest School and outdoor education are adaptable and flexible approaches to learning which have been applied to education settings and are increasingly being used with a range of pupils in Britain, including those with special educational needs. Outdoor education, such as Forest School uses regular sessions of hands-on learning experiences in a natural environment in order to encourage physical and emotional development.

This systematic review was conducted to examine the effectiveness of Forest School and outdoor education at improving social and emotional skills for children with special educational needs. Searches in Web of Science, PsychINFO and ERIC identified five studies relevant to this review, which were evaluated using Gough's (2007) Weight of Evidence Framework.

The studies showed that the effect sizes of outdoor education on social and emotional outcomes ranged from medium to large, suggesting that it is an effective intervention. Methodological issues are discussed along with recommendations for further research in order to increase the validity and generalisability of findings.

## **2. Introduction**

### ***2.1 Summary of intervention***

Forest School was originally developed in Denmark in the 1950s for the pre-school age group. It moved over to the United Kingdom (UK) in the 1990s and is now a form of outdoor education which is popular in the UK (Forest School Association, 2016). It offers children a chance to engage in a regular and sustained programme of child-initiated learning and play using the outdoors as a classroom, designed to encourage physical and emotional development through outdoor learning opportunities. Forest School is delivered on a regular basis, for example one morning per week, for at least a half term, but often throughout the year.

Apart from Forest School, in the UK, the Learning Outside the Classroom manifesto (Department for Education and Skills, 2006) supported use of more diverse learning sites, including the outdoors. Outdoor learning was made compulsory for Early Years Foundation Stage pupils in 2007 and is now provided through free flow between indoors and outdoors areas.

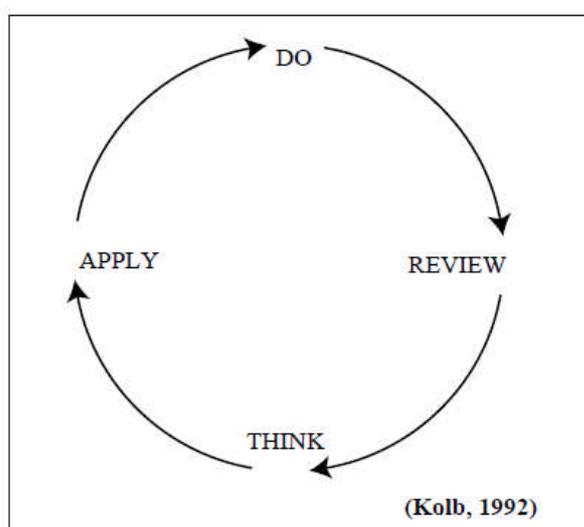
An evaluation carried out by the New Economics Foundation (NEF) on behalf of the Forestry Commission suggests that Forest School is effective in increasing children's confidence, social skills, communication, motivation, physical skills (stamina and fine and gross motor skills) and knowledge and understanding about the environment (O'Brien & Murray, 2006). Similarly a recent qualitative study found that children with Autism Spectrum Disorders (ASD) noted that Forest School allowed them to experience challenge and risk taking, make friends, achieve learning outcomes and experience success

(Bradley & Male, 2017) highlighting the positive impacts that the intervention could have.

## 2.2 Psychological basis

The term social, emotional and mental health was introduced into the most recent Code of Practice (Department for Education & Department of Health, 2015) and refers to difficulties which a pupil or young person may have in managing their emotions or behaviour. Social and emotional wellbeing is necessary for healthy behaviours and educational attainment (Gutman & Vorhaus, 2012), as well as preventing problems in later life. This highlights the importance of focussing interventions towards improving social and emotional skills of children and young people.

Outdoor education and Forest School are based on several psychological theories, including experiential learning (Kolb, 1992), Gardner’s multiple intelligences and theories around developing self-esteem. Firstly, the fact that outdoor education programmes are so interactive is considered beneficial as they encourage learning through the experiential learning cycle (Kolb, 1992). This model (Figure 1) involves a cycle of taking part in an



activity, analysing which parts were fun and which were difficult, thinking about why some parts were challenging and how fears were overcome, and making the activity relevant to everyday life (Farnham & Mutrie, 1997). The resulting immediate consequence of actions is thought to help to overcome short attention spans, making it particularly relevant to children with Special Educational Needs (SEN).

Figure 1: Experiential learning cycle (Kolb, 1992)

The relevance of this theory to outdoor education can be seen through group leaders attempting to maximize generalisable understanding from real life practical experiences.

Outdoor education also taps into the idea of multiple intelligences. Gardner proposed that intelligence can fall under 8 different categories: linguistic, logical-mathematical, spatial, bodily-kinesthetic, musical, interpersonal, intrapersonal, and naturalist intelligences. Gardner believes that schools focus most of their attention on linguistic and logical-mathematical intelligence, which results in labels being applied to children or children underachieving. A more balanced curriculum such as that delivered through outdoor education may be helpful to develop children's abilities in different areas (Gardner, 1983, 1999).

Forest school gives equal importance to these different types of intelligence. For example naturalistic intelligence will be supported through activities which give pupils a chance to explore and investigate their natural world. Also, both interpersonal relationships (how people get along in groups) and

intrapersonal relationships (relationships with themselves for example confidence, self-concept, self-efficacy) can be developed through individual and group problem solving and collaboration which occurs during outdoor education. For example, research has shown that forest school can aid the development of social skills and citizenship skills (Swarbrick, Eastwood, & Tutton, 2004).

Outdoor education also takes a positive psychology approach to education focussing on children's interests and strengths in a natural environment. In Forest School children are allowed to set their own goals based on their interests, allowing them to achieve what they themselves see as important. This is in turn valued by the group, building the self-esteem and confidence of individuals when working both independently and in groups on practical activities. Indeed Forest School has been shown to be associated with an increased self-esteem (O'Brien, 2009). The importance of self-esteem is highlighted by Maslow who placed self-esteem high up in his hierarchy of human needs (Figure 2), which need to be met in order for people to be motivated towards self-actualisation (Maslow, 1943).

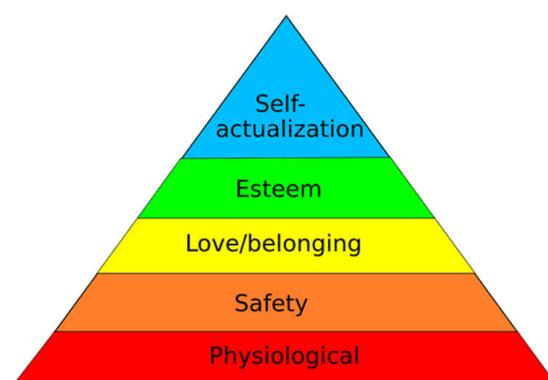


Figure 2: Maslow's Hierarchy of need

Self-esteem has been closely linked with more general psychological wellbeing and it has been highlighted as an influential predictor of lifelong outcomes such as relationships, job satisfaction and health (Orth, Robins, & Widaman, 2012). This highlights the potential benefits of interventions such as Forest School, through targeting self-esteem.

Commonly interventions aimed at children and young people with SEMH are aimed at the child or young person. Despite the fact that environment-centred interventions may produce larger benefits for children, it is less common for interventions to be aimed at the environment (Berryhill & Prinz, 2003). Outdoor education is an environment-centred approach to education which aims to nurture learning through the interactions between emotions, actions and thoughts, in practical situations (Gustafsson, Szczepanski, Nelson, & Gustafsson, 2012).

### ***2.3 Rationale for review/ relevance to EP practice***

Educational Psychologists (EPs) work with children and young people with a range of Special Educational Needs and Disabilities (SEND), including pupils with Social, Emotional and Mental Health needs. The recent government Green paper (Department of Health & Department for Education, 2017) reported that 1 in 10 young people experience some form of diagnosable mental health condition, meaning that it is currently a focus for schools.

In order to be able to recommend interventions for pupils that they are working with, or to comment on whether an existing intervention is appropriate, EPs need to have a good understanding of the research base and be able to clearly rationalise the implementation of interventions with

specific pupils. This is particularly important at a time when schools have to make decisions about how to use their limited resources effectively.

The purpose of this review is to provide a quantitative systematic review of the existing literature around forest school or outdoor education improving social and emotional outcomes for pupils with SEN. To the extent of my knowledge there is currently no recent systematic review in this area and so this systematic review will review the existing literature and bring together the findings in this area. In this way, the evidence found in this review can be used to help EPs and schools make decisions about intervention implementation.

### **2.4 Review question**

How effective is outdoor education (including forest school) at improving social and emotional skills in pupils with Special Educational Needs?

## **3. Critical review of evidence base**

### **3.1 Literature search**

A systematic review of existing literature was conducted between 12 December 2017 and 2 January 2018. The following online databases were searched through the online UCL library: PsychInfo, ERIC (Education Resources Information Center) and Web of Science. The search terms shown in Table 1 below were used.

Table 1: *Search Terms Used*

1	2	3
"special educational needs" OR autis* OR disabilit* OR "learning difficult*" OR ADHD	"forest school*" OR "outdoor learning" OR "outdoor curriculum" OR "outdoor education"	school* OR education OR college* OR young people OR children OR child OR boy* OR girl*

---

OR behavio* OR "learning disabilit**"	OR "outdoor adventure programme"	OR young person OR student* OR pupil* OR adolescent* OR teenager*
--	-------------------------------------	--

---

*Note.* The asterisk is a wildcard symbol meaning that any word that contains the letters up to the asterisk will be included.

Depending on the options available for database search options, searches were limited to the title and abstract, or topic. Where possible database tools were used to limit the searches to peer reviewed journals and articles published in the English language as per the inclusion criteria. The three searches above were combined using “AND”. Searches from the three databases were combined giving a total of 109 papers (PsychInfo - 38, ERIC - 27, Web of Science - 44) and then duplicates were removed to give 83 papers. The remaining studies were title and then abstract screened (using the exclusion criteria detailed in Table 2) to give 11 studies. The full text of these was read whilst applying the exclusion criteria and seven of the studies were excluded to leave four. Articles excluded during the full text screening can be found in Appendix A with the reason for exclusion. An ancestral search of the reference lists of the included papers was carried out and one further relevant study was found. This process is outlined in Figure 3 and gave five studies which were the focus of this review (see Table 3).

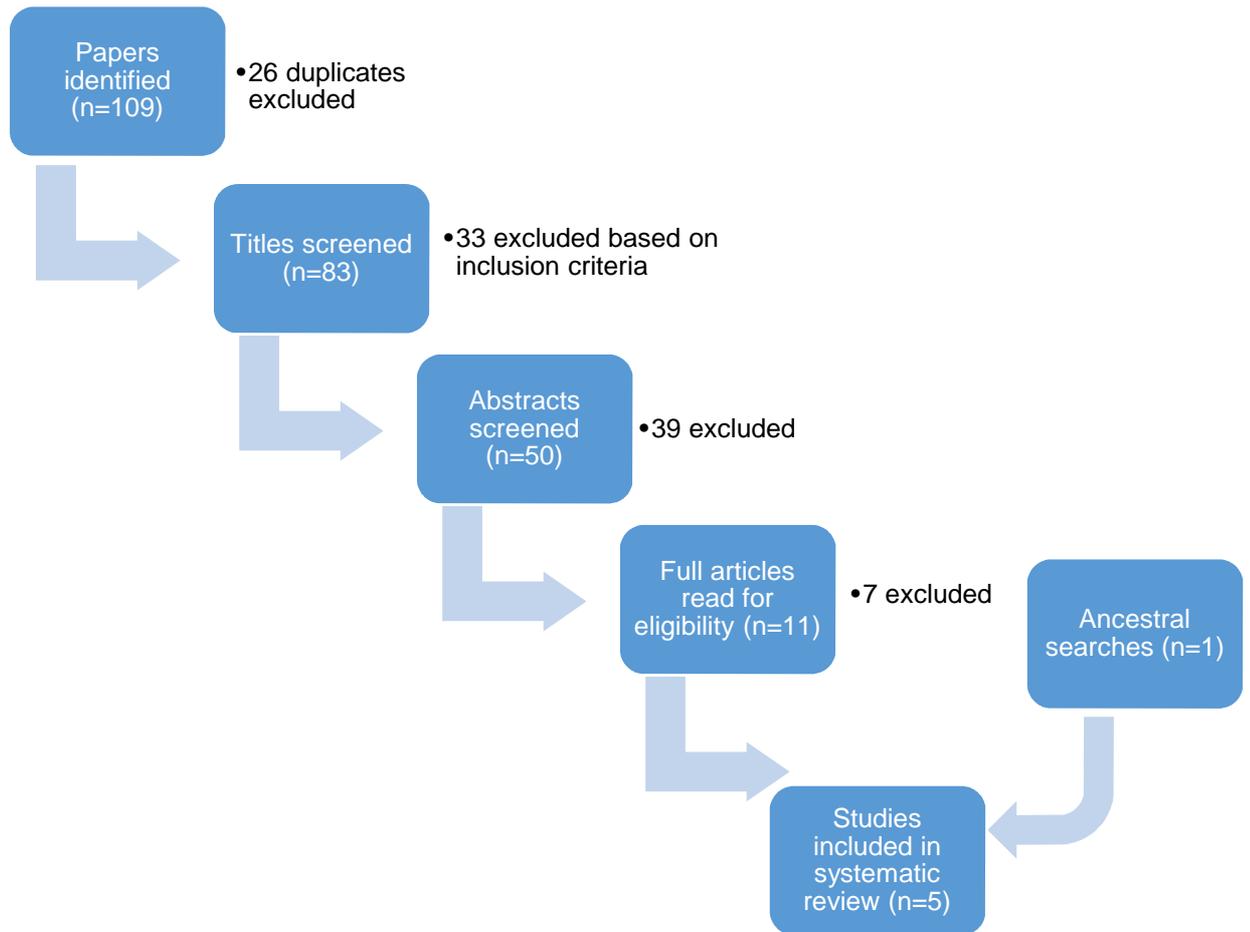


Figure 3: Flow diagram of the literature search and screening process

Table 2: *Inclusion and Exclusion Criteria*

		Inclusion criteria	Exclusion criteria	Rationale
1	Type of publication	Study is published in a peer-reviewed journal	Study is not published in a peer-reviewed journal	Ensures increased credibility and rigour of the research process
2	Type of study	Study includes evaluation of an intervention	Study does not include evaluation of an intervention	Research question is looking at effectiveness of intervention
3	Type of intervention	Intervention is a Forest School intervention or lessons delivered through an outdoor education curriculum, not specific to any subject	Interventions that were not based on principles of Forest school or outdoor education or were specific to one particular subject	Research question was looking at development of social and emotional skills rather than academic attainment in a particular subject.
4	Setting	Pupils aged between 0-25 in an education-based setting including special schools and nurseries	Intervention carried out in a non-education-based setting	Ensures findings are relevant to Educational Psychology practice
5	Location	Study conducted in an Organisation for Economic Co-operation and Development (OECD) country	Study conducted in a non-OECD country	Economical and educational similarities ensure that research findings are generalisable to UK schools
6	Participants	Participants had a special educational need	Participants did not have special educational needs	Research question is considering impact on children and young people with special educational needs
7	Outcome	Outcomes measured included quantitative measures of social and emotional skills	No quantitative outcomes were measured	Research question is looking at the effectiveness of the intervention
8	Language	Study is published in the English Language	Study is not published in the English Language	Ensures accurate translation from original language is not needed
9	Date	Study published before 2 January 2018	Study published after 2 January 2018	Literature searches took place up to this date before analysis took place

Table 3: *Studies Included in this Systematic Review*

Full study reference
Farnham, M., & Mutrie, N. (1997). Research Section: The Potential Benefits of Outdoor Development for Children with Special Needs. <i>British Journal of Special Education</i> , 24(1), 31-38.
Fox, P. & Avramidis, E.. (2003). An evaluation of an outdoor education programme for students with emotional and behavioural difficulties. <i>Emotional &amp; Behavioural Difficulties</i> , 8, 267-283.
Price, A. (2015). Improving school attendance: can participation in outdoor learning influence attendance for young people with social, emotional and behavioural difficulties?. <i>Journal of Adventure Education &amp; Outdoor Learning</i> , 15(2), 110-122.
Roe, J., & Aspinall, P. (2011). The restorative outcomes of forest school and conventional school in young people with good and poor behaviour. <i>Urban forestry &amp; urban greening</i> , 10(3), 205-212.
Zachor, D. A., Vardi, S., Baron-Eitan, S., Brodai-Meir, I., Ginossar, N. & Ben-Itzhak, E. (2017). The effectiveness of an outdoor adventure programme for young children with autism spectrum disorder: A controlled study. <i>Developmental Medicine &amp; Child Neurology</i> , 59, 550-556.

### **3.2 Mapping the field**

Appendix B shows the key features of the five studies identified through the systematic literature search.

### **3.3 Weight of Evidence**

Gough's Weight of Evidence (WoE) framework was used to critically evaluate the five studies identified for review (Gough, 2007). Criteria were used to assign the studies ratings over three different categories: methodological quality (WoE A), relevance of methodology (WoE B) and relevance of focus of the study to the review question (WoE C). These three ratings can be averaged to give Weight of Evidence D which gives an overall judgement of the value of each study to answering the review question.

Weight of Evidence A was judged using published coding protocols for research. The Gersten et al. (2005) protocol was used for three of the studies which had group experimental or quasi-experimental designs and Horner et al. (2005) was used for the other two studies which followed a small-N / single group research design. Full details can be found in Appendix C, and Appendix D contains the completed coding protocols.

WoE B evaluated how appropriate the methodologies of the studies were for answering the review question. This was done by considering the research design and a typology of evidence criteria was applied (Petticrew & Roberts, 2003). Details can be found in Appendix E.

WoE C examined how relevant the topics of the studies were to the research question. Criteria for this rating can be found in Appendix F.

The three Weight of Evidences calculated were averaged with equal weighting to give a WoE D score. A qualitative descriptor was given to this score based on the following ranges.

Low = less than 1.5

Medium = 1.5 – 2.5

High = above 2.5

Table 4: Overview of the Weight of Evidence Ratings

Study	WoE A	WoE B	WoE C	Overall WoE D
Farnham & Mutrie (1997)	1 (Low)	2 (Medium)	1.75 (Medium)	1.58 (Medium)
Fox & Avramidis (2003)	1 (Low)	1 (Low)	2.5 (High)	1.5 (Medium)
Price (2015)	1 (Low)	1 (Low)	2.25 (Medium)	1.42 (Low)
Roe & Aspinall (2011)	1 (Low)	2 (Medium)	2 (Medium)	1.67 (Medium)
Zachor et al (2017)	3 (High)	3 (High)	2 (Medium)	2.67 (High)

### 3.4 Participants

The participants included in this review all had some kind of special educational need; however the details of this differed between studies. Four of the studies recruited participants from special educational provisions which catered for pupils with social, emotional and behavioural difficulties (Farnham & Mutrie, 1997; Fox & Avramidis, 2003), autism spectrum conditions (Zachor et al., 2017), or either of these (Price, 2015). None of these studies explicitly measured the social and emotional skills of the pupils pre-intervention; however it was assumed that the pupils all had some level of need given that they were in a specialist provision. The remaining study recruited their participants from both a mainstream and a residential school, splitting participants into ‘good’ and ‘bad’ behaviour groups. They did not state the specific difficulties that pupils had due to confidentiality issues (Roe & Aspinall, 2011). This lack of specificity of the needs of the participants resulted in a lower WoE C rating for this study.

The age of participants varied between studies. Four of the studies involved secondary aged pupils (Farnham & Mutrie, 1997; Fox & Avramidis, 2003; Price, 2015; Roe & Aspinall, 2011) and one involved nursery aged children (Zachor et al., 2017). This limits the generalisability of the results found to the primary age group, as none of the pupils studied were in this age group.

The Horner et al. (2005) coding protocol used for WoE A looks at whether the selection of participants was described with replicable precision. Neither Price (2015) nor Fox and Avramidis (2003) included details about how the participants were selected, apart from using a convenience sampling approach. This was reflected in the low WoE A ratings for these studies.

For the study which included a comparison group (Zachor et al., 2017) the authors analysed measures pre-intervention to check that the two groups did not differ significantly from each other. This ensures that relevant characteristics of the samples are comparable across conditions, reducing the likelihood that results could be due to a confounding factor, and so this study was rated higher for WoE A.

Four of the studies were carried out in settings in the UK, whereas Zachor et al. (2017) was assumed to take place in Israel. This was not explicitly stated in the paper but inferred from the authors' correspondence addresses. Therefore the generalisability of the results from Zachor et al. (2017) to UK settings is reduced. This is reflected in the study's WoE C rating.

### ***3.5 Research design***

Only quantitative studies were included in this review due to the research question looking at effectiveness. One of the studies used an experimental study design comparing outcomes for an intervention group and a control group with random

assignment to groups (Zachor et al., 2017). This random assignment to groups meant that the likelihood of bias was reduced in this study (Barker, Pistrang, & Elliott, 2015). Therefore, this is reflected by the high WoE B score for this study. Quasi-experimental designs with within-subject approaches were used by two studies (Farnham & Mutrie, 1997; Roe & Aspinall, 2011). One of these followed a within-subjects approach comparing outcomes between settings (Roe & Aspinall, 2011), and the other used a within-subject approach comparing scores pre- and post-intervention (Farnham & Mutrie, 1997). Because both of these studies followed a quasi-experimental design and did not either involve a control group or random allocation between groups, it resulted in them being coded a 2 for WoE B. This is because there may have been a compromise to the internal validity due to the outcomes measured being affected by variables other than just the presence of the intervention. Farnham and Mutrie (1997) do acknowledge that the lack of a control group means that a causal relationship cannot be concluded. The other two studies used small-N designs with post-test data only (Fox & Avramidis, 2003; Price, 2015). According to the Horner et al. (2005) coding protocol, the dependent variable should be measured at multiple points during the baseline in order to be considered adequate. This did not happen in either of these studies and so they received Low ratings on both WoE A and WoE B.

### **3.6 Intervention**

The studies covered a range of interventions under the topic of outdoor education, with only one study looking specifically at a Forest School (Roe & Aspinall, 2011). Irrelevant of what the intervention was, both of the coding protocols included criteria about whether the intervention was described with replicable precision. It was felt that only Zachor et al. (2017) and Price (2015) included clear and specific details

about the intervention, for example frequency of sessions, what was covered, who ran the sessions, with Price (2015) including the programme aims and structures in the appendices. However, Zachor et al. (2017) was the only study that mentioned how fidelity of implementation to the intervention was ensured. This resulted in a higher WoE A score for this study.

When two different interventions were being compared (Roe & Aspinall, 2011; Zachor et al., 2017) details of what the control intervention consisted of were lacking which is reflected in the WoE A scoring.

The ease of implementation of the intervention was covered by one of the criteria for WoE C, as it was felt that this was important to increase the likelihood that the intervention would be implemented if it could be done so within existing school resources. Fox and Avramidis (2003) and Roe and Aspinall (2011) scored highest for this criteria of WoE C, as the interventions in these studies could be implemented by school staff using the school's existing resources.

Price (2015) was also the only study which ran the intervention for a full school year. This is in line with the ideas of Forest School that the intervention should run continuously, rather than being one off more like an outward bound course. As WoE C covered two aspects of intervention implementation, a combination of the above reasons resulted in Price (2015) and Fox and Avramidis (2003) scoring highly for WoE C. Farnham and Mutrie (1997) scored lowest for WoE C due to it including a one off intervention which needed specialist staff and resources to implement.

### **3.7 Measures**

The studies in this review all used different measures to evaluate the effectiveness of the interventions on social and emotional skills. Farnham and Mutrie (1997)

measured tension and anxiety using the Unipolar Profile of Mood States (McNair, 1971) and group cohesion using the Group Environment Questionnaire (Carron, Widmeyer, & Brawley, 1985). Roe and Aspinall (2011) used the Mood Adjective Checklist (Matthews, Jones, & Chamberlain, 1990) to measure hedonic tone, energy, stress and anger levels of participants, and Zachor et al. (2017) used the Social Responsiveness Scale (Bölte, Poustka, & Constantino, 2008) to measure social impairment. Fox and Avramidis (2003) used an un-validated school measure of pupil behaviour and Price (2015) analysed pupil attendance as a measure of their social and emotional skills. This range of measures used means that the range of benefits of this type of intervention can be seen; however it also means that it is less easy to compare between studies or pinpoint the precise areas in which benefits may be greatest.

The coding protocol used for the experimental / quasi-experimental studies looks at whether there were multiple measures used within each study and whether outcomes were measured at appropriate time points (Gersten et al., 2005). Both of these were present in all three of the studies with which this protocol was used (Farnham & Mutrie, 1997; Roe & Aspinall, 2011; Zachor et al., 2017). Zachor et al. (2017) and Farnham and Mutrie (1997) stated the internal consistency reliability for one of their measures. In the same way, the importance of reliable, valid and precise measurement of dependent variables is highlighted by the coding protocol used for the single group research designs (Horner et al., 2005). In general this was lacking from the studies and the two studies with which this protocol was used contained unstandardised measures with unknown reliability and validity. This was reflected in the reviewer judgements of WoE A.

### 3.8 Findings

Effect sizes were calculated for all findings relevant to social and emotional outcomes, as this relates to the review question. For Farnham and Mutrie (1997) where pre- and post- data was available the effect size was calculated using the formula for Standardised Mean Difference (SMD).

$$\text{SMD} = \frac{\text{Mean}_{\text{post}} - \text{Mean}_{\text{pre}}}{\text{SD}_{\text{pre}}}$$

The Campbell Collaboration Effect Size Calculator was used to calculate Cohen's d for the two studies where F statistics were quoted (Roe & Aspinall, 2011; Zachor et al., 2017).

Effect size descriptors are given using Cohen's guidance (Cohen, 1992), see Table 5. These are shown along with the effect sizes in Table 6.

Table 5: *Effect Size Descriptors (Cohen, 1992)*

Effect size	Effect size descriptor
0.2-0.49	Small
0.5-0.79	Medium
0.80-1.0	Large

Effect sizes could not be calculated for two of the studies included in this paper. Fox and Avramidis (2003) used the school’s own measure of pupil behaviour to evaluate the impact of the outdoor education programme and quantitative data was only available for three of the pupils who took part in the study. As this is not a valid or reliable measure, effect sizes could not be calculated for this study. The authors concluded that the outdoor education programme was successful in promoting positive behaviour for this group of pupils with severe emotional and behavioural difficulties. However, the quality of the data provided meant that this could not be confirmed. Price (2015) evaluated the impact of outdoor learning by using pupil attendance data. Due to the quality of the data available, an effect size could not be calculated for this study either. Price (2015) reported that pupil attendance on outdoor learning days was higher (79.8%) than on days when outdoor learning did not occur (74.6%), however the statistical significance of this is unknown. These two studies were also the lowest in terms of WoE D.

The effect sizes calculated ranged from medium to extremely large, with the largest effect sizes seen for Roe and Aspinall’s data. The outcomes which were found to have improved were energy ( $d=1.46$ ), stress ( $d=-.993$ ), hedonic tone ( $d=1.49$ ) and anger ( $d=-1.22$ ). This study received an overall weight of evidence score (WoE D) of 1.67 which was categorised as Medium. This WoE D rating was mainly due to the

authors not providing reliability and validity measures for the outcome measures and a lack of description and clarity over what the intervention and control conditions looked like. This lack of description about the intervention means that it is hard to determine the construct validity and the effects seen may be due to other confounding variables or the Hawthorne effect, in which participation in research itself produces beneficial change in behaviour (Barker et al., 2015).

Four of the studies included in this review used within-subject designs comparing pre- and post- intervention. The lack of a control group poses a threat to internal validity with the possibility that the effect seen is purely down to another factor apart from the intervention, for example maturational trends or reactivity of measurement (Barker et al., 2015). The only study which followed a RCT design (Zachor et al., 2017) showed a large effect ( $d=.810$ ) of outdoor education in terms of social responsiveness in children with ASD. This study received the highest WoE D rating (2.67; high) suggesting that it provides the best evidence for suggesting that outdoor education is effective at improving social and emotional skills.

Table 6: *Effect Sizes for Outcomes*

Study	Sample size	Comparison	Measure	Effect size – Standardised Mean Difference (d)	Descriptor	WoE D
Farnham & Mutrie (1997)	19	Within participants (pre- and post)	Unipolar Profile of Mood States (POMS) – tension/anxiety	-.731	Medium	1.58 (Medium)
			Group Environment Questionnaire (GEQ) – group cohesion	.946	Large	
Fox & Avramidis (2003)	11 (quantitative results only reported for 3 pupils)	Within participants (pre- and post)	School measure of Pupil behaviour	-- <sup>a</sup>	-- <sup>a</sup>	1.5 (Medium)
Price (2015)	5	Within participants (outdoor learning vs. other lessons)	Learner attendance (%)	-- <sup>a</sup>	-- <sup>a</sup>	1.42 (Low)
Roe & Aspinall (2011)	18 including 6 pupils with 'good' behaviour and 6 pupils	Within group (forest school vs. conventional school)	Mood Adjective Checklist (MACL) – hedonic tone	Energy: 1.46	Large	1.67 (Medium)
				Stress: -.993	Large	

Study	Sample size	Comparison	Measure	Effect size – Standardised Mean Difference (d)	Descriptor	WoE D
	with 'bad' behaviour			Hedonic tone: 1.49 Anger: -1.22	Extremely Large Extremely Large	
Zachor et al. (2017)	51 (Intervention group = 30, control group = 21)	Between group (treatment vs. control)	Social Responsiveness Scale (SRS) – social impairment	.810	Large	2.67 (High)

<sup>a</sup> The effect size could not be calculated for these studies due to lack of quality of the data provided in the study.

## **4. Conclusion and recommendations**

### **4.1 Conclusions**

Five studies were included in this systematic literature review in order to evaluate the effectiveness of outdoor education and forest school on social and emotional outcomes for children with SEN. All studies included in the review found at least a medium effect size across a range of social and emotional outcomes for pupils with SEN. The smallest effect size was for Farnham and Mutrie (1997) which implemented the intervention for the shortest time period, suggesting that larger effects may be seen with longer intervention programmes.

The methodological issues discussed in the previous section may limit the validity and reliability of findings and this is reflected in the WoE ratings. Only one study received a 'high' WoE D score (Zachor et al., 2017), suggesting that in general the evidence available lacked quality and appropriateness for answering this specific review question.

None of the studies in this review involved pupils in the primary age group; one involved nursery aged children and the other four secondary aged children. This poses limitations to the generalisability of findings, and it would not currently be considered evidence-based to implement an outdoor education intervention to improve social and emotional skills in primary aged children with SEN. Also four of the studies included in this review recruited participants solely from schools for pupils with SEN. This therefore also limits the generalisability of the results to pupils with SEN in mainstream settings. In addition, the inclusion criteria of pupils with SEN means that a broad range of participants were included in the studies. Therefore it is uncertain how generalisable the results are to all pupils within this broad category of pupils with special educational needs.

#### **4.2 Recommendations for use in practice**

In summary, this review provides promising evidence towards the use of Forest School and outdoor education at improving social and emotional skills in pupils with SEN; however more research is needed as only a small number of studies have been reviewed and these included limitations. Outdoor education is an environment-centred intervention, adapting how the curriculum is taught, moving away from the view that the needs are within-child and may provide wider benefits. This review suggests that outdoor education is a useful strategy that EPs can recommend to school staff for pupils with social and emotional needs. However, issues around funding, training and fidelity of implementation will need to be considered by EPs when recommending this to schools.

#### **4.3 Areas of further research**

Only one study received a high WoD D rating (Zachor et al., 2017) with relevance to this review question. This suggests that there is a need for future studies to be more rigorous in their methodology, for example by using reliable and valid measures through a randomised research design involving a control condition, in order to measure the effectiveness of this intervention accurately.

As discussed above, further research is also needed into the effect of outdoor education at improving social and emotional skills for children with SEN in the primary age group. This would be beneficial to add to the evidence base for this type of intervention and will facilitate practitioners to make informed decisions about how resources are spent in order to achieve the best outcomes for this age group.

## 5. References and appendices

### References

- Barker, C., Pistrang, N., & Elliott, R. (2015). *Research Methods in Clinical Psychology*. Chichester, UK: John Wiley & Sons, Ltd.  
<https://doi.org/10.1002/9781119154082>
- Berryhill, J. C., & Prinz, R. J. (2003). Environmental Interventions to Enhance Student Adjustment: Implications for Prevention. *Prevention Science*, 4(2), 65–87. <https://doi.org/10.1023/A:1022994514767>
- Bölte, S., Poustka, F., & Constantino, J. N. (2008). Assessing autistic traits: cross-cultural validation of the social responsiveness scale (SRS). *Autism Research*, 1(6), 354–363. <https://doi.org/10.1002/aur.49>
- Bradley, K., & Male, D. (2017). Forest School is muddy and I like it': Perspectives of young children with autism spectrum disorders, their parents and educational professionals. *Educational & Child Psychology*, 34(2), 80–93. Retrieved from <http://www.netley.camden.sch.uk/welcome/wp-content/uploads/2017/06/Forest-School-research-paper-Kate-Bradley.pdf>
- Carron, A. V., Widmeyer, W. N., & Brawley, L. R. (1985). The Development of an Instrument to Assess Cohesion in Sport Teams: The Group Environment Questionnaire. *Journal of Sport Psychology*, 7(3), 244–266.  
<https://doi.org/10.1123/jsp.7.3.244>
- Cohen, J. (1992). A Power Primer. *Psychological Bulletin*, 112(1), 155–159. Retrieved from <http://www.bwgriffin.com/workshop/Sampling A Cohen tables.pdf>
- Department for Education & Department of Health. (2015). Special educational needs and disability code of practice: 0 to 25 years, (January), 35.  
<https://doi.org/DFE-00205-2013>
- Department for Education and Skills. (2006). *Learning Outside the Classroom MANIFESTO*. Retrieved from <http://www.lotc.org.uk/wp-content/uploads/2011/03/G1.-LOtC-Manifesto.pdf>
- Department of Health, & Department for Education. (2017). Transforming Children and Young People's Mental Health Provision: a Green Paper. Retrieved from [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/664855/Transforming\\_children\\_and\\_young\\_people\\_s\\_mental\\_health\\_provision.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/664855/Transforming_children_and_young_people_s_mental_health_provision.pdf)
- Farnham, M., & Mutrie, N. (1997). Research Section: The Potential Benefits of Outdoor Development for Children with Special Needs. *British Journal of Special Education*, 24(1), 31–38. <https://doi.org/10.1111/1467-8527.00008>
- Forest School Association. (2016). History of Forest School. Retrieved March 27, 2018, from <http://www.forestschoollassociation.org/history-of-forest-school/>
- Fox, P., & Avramidis, E. (2003). An evaluation of an outdoor education programme for students with emotional and behavioural difficulties. *Emotional and Behavioural Difficulties*, 8(4), 267–283.

<https://doi.org/10.1080/13632750300507025>

- Gardner, H. (1983). *Frames of mind : the theory of multiple intelligences*. Basic Books. Retrieved from [https://books.google.co.uk/books?id=wxj6npSaykgC&dq=Gardner,+H.+\(1983\).+Frames+of+mind:+The+theory+of+multiple+intelligences.+New+York:+Basic+Books.&hl=en&sa=X&ved=0ahUKEwiL29aQ3ofZAhVBC8AKHXDcA-sQ6AEIJzAA](https://books.google.co.uk/books?id=wxj6npSaykgC&dq=Gardner,+H.+(1983).+Frames+of+mind:+The+theory+of+multiple+intelligences.+New+York:+Basic+Books.&hl=en&sa=X&ved=0ahUKEwiL29aQ3ofZAhVBC8AKHXDcA-sQ6AEIJzAA)
- Gardner, H. (1999). *Intelligence reframed : multiple intelligences for the 21st century*. Basic Books. Retrieved from [https://books.google.co.uk/books/about/Intelligence\\_Reframed.html?id=pU4gAQAAQBAJ&redir\\_esc=y](https://books.google.co.uk/books/about/Intelligence_Reframed.html?id=pU4gAQAAQBAJ&redir_esc=y)
- Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education. *Exceptional Children*, 71(2), 149–164. <https://doi.org/10.1177/001440290507100202>
- Gough, D. (2007). Weight of Evidence: a framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213–228. <https://doi.org/10.1080/02671520701296189>
- Gustafsson, P. E., Szczepanski, A., Nelson, N., & Gustafsson, P. A. (2012). Effects of an outdoor education intervention on the mental health of schoolchildren. *Journal of Adventure Education & Outdoor Learning*, 12(1), 63–79. <https://doi.org/10.1080/14729679.2010.532994>
- Gutman, L. M., & Vorhaus, J. (2012). The Impact of Pupil Behaviour and Wellbeing on Educational Outcomes Childhood Wellbeing Research Centre. Retrieved from [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/219638/DFE-RR253.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/219638/DFE-RR253.pdf)
- Horner, R. H., Nforfaun, V., Carr, E. G., Halle, J., Mcgee, G., Odom, S., ... Wolery, M. (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education. *Exceptional Children*, 71(2), 165–179. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.626.2457&rep=rep1&type=pdf>
- Kolb, D. G. (1992). The Practicality of Theory. *Journal of Experiential Education*, 15(2), 24–28. <https://doi.org/10.1177/105382599201500204>
- Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review*, 50(4), 370–396. Retrieved from [https://s3.amazonaws.com/academia.edu.documents/34195256/A\\_Theory\\_of\\_Human\\_Motivation\\_-\\_Abraham\\_H\\_Maslow\\_-\\_Psychological\\_Review\\_Vol\\_50\\_No\\_4\\_July\\_1943.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1517153775&Signature=6Gz%252FyZ4U%252FaB7M5jkTwOCcgdPKFw%253D&r](https://s3.amazonaws.com/academia.edu.documents/34195256/A_Theory_of_Human_Motivation_-_Abraham_H_Maslow_-_Psychological_Review_Vol_50_No_4_July_1943.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1517153775&Signature=6Gz%252FyZ4U%252FaB7M5jkTwOCcgdPKFw%253D&r)
- Matthews, G., Jones, D. M., & Chamberlain, A. G. (1990). Refining the measurement of mood: The UWIST Mood Adjective Checklist. *British Journal of Psychology*, 81(1), 17–42. <https://doi.org/10.1111/j.2044-8295.1990.tb02343.x>

- McNair, D. (1971). *Manual profile of mood states*. San Diego: Educational & Industrial testing Service. Retrieved from <http://www.worldcat.org/title/manual-profile-of-mood-states/oclc/500535397>
- O'Brien, L. (2009). Learning outdoors: the Forest School approach. *Education 3-13*, 37(1), 45–60. <https://doi.org/10.1080/03004270802291798>
- O'Brien, L., & Murray, R. (2006). *A participatory evaluation of Forest School in England and Wales. A participatory evaluation of Forest School in England and Wales A marvellous opportunity for children to learn*. Retrieved from [https://www.forestry.gov.uk/pdf/fr0112forestschoolsreport.pdf/\\$FILE/fr0112forestschoolsreport.pdf](https://www.forestry.gov.uk/pdf/fr0112forestschoolsreport.pdf/$FILE/fr0112forestschoolsreport.pdf)
- Orth, U., Robins, R. W., & Widaman, K. F. (2012). Life-Span Development of Self-Esteem and Its Effects on Important Life Outcomes. *Journal of Personality and Social Psychology*, 102(6), 1271–1288. <https://doi.org/10.1037/a0025558>
- Petticrew, M., & Roberts, H. (2003). Evidence, hierarchies, and typologies: horses for courses. *Journal of Epidemiology and Community Health*, 57(7), 527–9. <https://doi.org/10.1136/JECH.57.7.527>
- Price, A. (2015). Improving school attendance: can participation in outdoor learning influence attendance for young people with social, emotional and behavioural difficulties? *Journal of Adventure Education and Outdoor Learning*, 15(2), 110–122. <https://doi.org/10.1080/14729679.2013.850732>
- Roe, J., & Aspinall, P. (2011). The restorative outcomes of forest school and conventional school in young people with good and poor behaviour. *Urban Forestry & Urban Greening*, 10(3), 205–212. <https://doi.org/10.1016/J.UFUG.2011.03.003>
- Swarbrick, N., Eastwood, G., & Tutton, K. (2004). Self-esteem and successful interaction as part of the forest school project. *Support for Learning*, 19(3), 142–146. <https://doi.org/10.1111/j.0268-2141.2004.00337.x>
- Zachor, D. A., Vardi, S., Baron-Eitan, S., Brodai-Meir, I., Ginossar, N., & Ben-Itzhak, E. (2017). The effectiveness of an outdoor adventure programme for young children with autism spectrum disorder: a controlled study. *Developmental Medicine & Child Neurology*, 59(5), 550–556. <https://doi.org/10.1111/dmcn.13337>

**Appendix A: Articles excluded at full text screening with reason**

Study reference	Exclusion criteria used to exclude
Brodin, J. (2009). Inclusion through access to outdoor education: Learning in Motion (LIM). <i>Journal of Adventure Education &amp; Outdoor Learning</i> , 9(2), 99-113.	7 – study collects qualitative data
Bradley, K. & Male, D. B. (2017). 'Forest School is muddy and I like it': Perspectives of young children with autism spectrum disorders, their parents and educational professionals. <i>Educational and Child Psychology</i> , 34(2), 80-93.	7 – study collects qualitative data
Furman, N., & Sibthorp, J. (2014). The development of prosocial behavior in adolescents: A mixed methods study from NOLS. <i>Journal of Experiential Education</i> , 37(2), 160-175.	6 – participants did not have SEN
Hurley, L. P & Lustbader, L. L. (1997). Project Support: Engaging children and families in the educational process. <i>Adolescence</i> , 32(127), 523-531.	3 – did not look at an outdoor education intervention
Marx, J. D. (1988). An outdoor adventure counseling program for adolescents. <i>Social Work</i> , 33(6), 517-520.	2 – study contains no evaluation data
Mirrahimi, S., Tawil, N. M., Abdullah, N. A. G., Surat, M., & Usman, I. M. S. (2011). Developing conducive sustainable outdoor learning: The impact of natural environment on learning, social and emotional intelligence. <i>Procedia Engineering</i> , 20, 389-396.	2 – study is a review of existing literature
Schleien, S. J, Hornfeldt, D. A & McAvoy, L. H. (1994). Integration and environmental/outdoor education: The impact of integrating students with severe developmental disabilities on the academic performance of peers without disabilities. <i>Therapeutic Recreation Journal</i> , 28(1), 25-34.	6 – participants from which outcomes were collected did not have SEN

**Appendix B: Mapping the field – key features of the five studies included in this review**

Author	Sample size	Age, gender	SEN	Study location	Study design	Intervention	Delivery	Outcomes
Farnham & Mutrie (1997)	19	1 female, 18 males, 13-17 years old (average age 14)	Mild to moderate SEN and emotional / behavioural difficulties	Scotland	Quasi-experimental one group – within subject comparisons (pre-/post-test design)	Outdoor development programme at residential centre	5 day residential programme delivered by staff at the centre	Following the outdoor development programme children showed a significant decrease in anxiety/tension ( $t(18)=4.06, p<0.05$ ), a significant increase in 'individual attractions to the group task/group social' ( $t(18)=6.82, p<0.05$ ) and no significant difference in self-perceptions.
Fox & Avramidis (2003)	11	All male, Year 9 and 10 pupils	Pupils with severe behavioural difficulties	South-west England	Small-N within subject comparisons (intervention vs. no intervention)	Outdoor education programme	Outdoor education for 2 hours once per week within school day	The authors concluded that the outdoor education programme was successful in promoting positive behaviour for this group of pupils with severe emotional and behavioural difficulties, although no statistical analysis was carried out.
Price (2015)	6	Genders not stated, aged between 12 and 13 years (Year 8 pupils)	All had Statements of SEN for social, emotional and behavioural difficulties or autistic spectrum disorders.	Special school in England	Small-N within subject comparisons (intervention vs. no intervention)	One year outdoor learning programme (OLP)	The OLP ran on Tuesdays throughout the school year and incorporated a residential element each term.	The researchers did not do any statistics on the data. Attendance on outdoor learning days was found to be higher (79.8%) than on days when outdoor learning did not occur (74.6%)

Author	Sample size	Age, gender	SEN	Study location	Study design	Intervention	Delivery	Outcomes
Roe & Aspinall (2011)	18	15 males, 3 females. Average age 11 years old	Details unknown although group included teenagers with ADHD, those 'at risk' from exclusion or exhibiting withdrawn behaviour.	Deprived urban areas of Central Scotland	Quasi-experimental – within subject comparisons (intervention vs. no intervention)	Forest school compared to conventional school	In forest settings near the school	Researchers ran a repeated measures ANOVA and found that Forest School had significant restorative effect on energy ( $F(1, 17) = 9.65, p=0.007$ ), hedonic tone ( $F(1, 17) = 9.99, p=0.007$ ) and anger ( $F(1, 17) = 6.65, p=0.020$ ) and a borderline effect on stress ( $F(1, 17) = 4.44, p=0.052$ ).
Zachor et al. (2017)	51	40 males, 11 females; aged 3 years 4 months – 7 years 4 months	Diagnosis of ASD	ASD special education kindergartens in Israel?	Randomised Control Trial - grouped by kindergarten (intervention vs. control)	Intervention group (n=30) participated in the outdoor adventure programme for 13 weeks	Programme delivered by Not for Profit Organisation. 13 weekly 30 minute sessions were delivered in urban parks near to the kindergartens.	A significant improvement in scores on the Social Responsiveness Scale was seen in the outdoor adventure group compared to the control group ( $F(1, 29)=8.1, p=0.007$ )

**Appendix C: Weight of Evidence A criteria and rationale**

For the Gersten et al. (2005) protocol ‘essential; and ‘desirable’ criteria were used to rate the studies. For the Horner et al. (2005) coding protocol, scores from each of the 7 sections were averaged in order to give an overall numerical rating for methodological quality. The criteria for converting these ratings to WoE A scores is shown below.

	WoE A score	Criteria	Rationale
Gersten et al. (2005) Coding Protocol	3	Meets $\geq 9$ essential criteria and $\geq 4$ desirable criteria	Based on recommendations of Gersten et al. (2005)
	2	Meets $\geq 9$ essential criteria and $\geq 1$ desirable criteria	
	1	Meets $< 9$ essential criteria	
Horner et al. (2005) Coding Protocol	3	Average rating $\geq 2.5$	Criteria converts possible average ratings of 0 – 3 into score of 1-3
	2	Average rating between 2.49 and 1.5	
	1	Average rating $< 1.5$	

Group designs

Study	Essential criteria fulfilled	Desirable criteria fulfilled	Weight of Evidence A (WoE A)
Farnham & Mutrie (1997)	4	4	1
Roe & Aspinall (2011)	6	1	1
Zachor et al. (2017)	9	4	3

Single subject designs

Study	Description of participants and setting	Dependent variable	Independent variable	Baseline	Experimental control/internal validity	External validity	Social validity	Average	WoE A
Fox & Avramidis (2003)	1	1	1	0	2	2	2	1.29	1 (Low)
Price (2015)	0	1	2	0	1	2	1	1.14	1 (Low)

### Appendix D: Coding Protocols

Gersten et al., (2005). Quality Indicators for Group Experimental and Quasi-Experimental Research in Special Education

**Date:** 04/01/18

**Name of coder:**

**Study Reference:** Zachor, Ditzza A, Vardi, Shira, Baron-Eitan, Shani, Brodai-Meir, Inbal, Ginossar, Noa & Ben-Itzchak, Esther. (2017). The effectiveness of an outdoor adventure programme for young children with autism spectrum disorder: A controlled study. *Developmental Medicine & Child Neurology*, 59, 550-556.

**Intervention name:** Outdoor Adventure Programme

**Research Design:** Randomised Controlled Trial

**Publication Type:** Journal article

#### Essential Quality Indicators

##### *Quality Indicators for Describing Participants*

1. Was sufficient information provided to determine/confirm whether the participants demonstrated the disability/ies or difficulties presented?
  - Yes
  - No
  - N/A
  - Unknown
  
2. Were appropriate procedures used to increase the likelihood that relevant characteristics of participants in the sample were comparable across conditions?
  - Yes
  - No
  - N/A
  - Unknown
  
3. Was sufficient information given characterizing the interventionists or teachers provided? Did it indicate whether they were comparable across conditions?
  - Yes
  - No
  - N/A
  - Unknown

*Quality Indicators for Implementation of the Intervention and Description of Comparison Conditions*

1. Was the intervention clearly described and specified?
  - Yes
  - No
  - N/A
  - Unknown
  
2. Was the fidelity of implementation described and assessed?
  - Yes
  - No
  - N/A
  - Unknown
  
3. Was the nature of services provided in comparison conditions described?
  - Yes
  - No
  - N/A
  - Unknown

*Quality Indicators for Outcome Measures*

1. Were multiple measures used to provide an appropriate balance between measures closely aligned with the intervention and measures of generalized performance?
  - Yes
  - No
  - N/A
  - Unknown
  
2. Were outcomes for capturing the interventions effect measured at the appropriate times?
  - Yes
  - No
  - N/A
  - Unknown

*Quality Indicators for Data Analysis*

1. Were the data analysis techniques appropriately linked to key research questions and hypotheses? Were they appropriately linked to the limit of analysis in the study?
  - Yes
  - No
  - N/A
  - Unknown

2. Did the research report include not only inferential statistics but also effect size calculations?
  - Yes
  - No
  - N/A
  - Unknown

### Desirable Quality Indicators

1. Was data available on attrition rates among intervention samples? Was severe overall attrition documented? If so, is attrition comparable across samples? Is overall attrition less than 30%?
  - Yes
  - No
  - N/A
  - Unknown
  
2. Did the study provide not only internal consistency reliability but also test-retest reliability and interrater reliability (when appropriate) for outcome measures? Were data collectors and/or scorers blind to study conditions and equally (un)familiar to examinees across study conditions?
  - Yes
  - No
  - N/A
  - Unknown
  
3. Were outcomes for capturing the intervention's effect measured beyond an immediate post-test?
  - Yes
  - No
  - N/A
  - Unknown
  
4. Was evidence of the criterion-related validity and construct validity of the measures provided?
  - Yes
  - No
  - N/A
  - Unknown
  
5. Did the research team assess not only surface features of fidelity implementation (e.g., number of minutes allocated to the intervention or teacher/interventionist following procedures specified), but also examine quality of implementation?
  - Yes
  - No

- N/A
- Unknown

6. Was any documentation of the nature of instruction or series provided in comparison conditions?

- Yes
- No
- N/A
- Unknown

7. Did the research report include actual audio or videotape excerpts that capture the nature of the intervention?

- Yes
- No
- N/A
- Unknown

8. Were results presented in a clear, coherent fashion?

- Yes
- No
- N/A
- Unknown

Number of Essential criteria met: 9

Number of Desirable criteria met: 2

Overall Weight of Evidence A: 3

Horner et al., (2005). The Use of Single-Subject Research to Identify Evidence-Based Practice in Special Education

Quality Indicators Within Single-Subject Research

**Date:** 06/01/17

**Name of coder:**

**Study Reference:** Fox, Paul & Avramidis, Elias. (2003). An evaluation of an outdoor education programme for students with emotional and behavioural difficulties. *Emotional & Behavioural Difficulties*, 8, 267-283.

**Intervention name:** Outdoor education

**Publication Type:** Journal Article

*Description of Participants and Setting*

Participants are described with sufficient detail to allow others to select individuals with similar characteristics (e.g., age, gender, disability, diagnosis).

- Yes
- No
- N/A
- Unknown

The process for selecting participants is described with replicable precision.

- Yes
- No
- N/A
- Unknown

Critical features of the physical setting are described with sufficient precision to allow replication.

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*Dependent Variable*

Dependent variables are described with operational precision.

- Yes
- No
- N/A
- Unknown

Each dependent variable is measured with a procedure that generates a quantifiable index.

- Yes
- No
- N/A
- Unknown

Measurement of the dependent variable is valid and described with replicable precision.

- Yes
- No
- N/A
- Unknown

Dependent variables are measured repeatedly over time.

- Yes
- No
- N/A
- Unknown

Data are collected on the reliability or inter-observer agreement associated with each dependent variable, and IOA levels meet minimal standards (e.g., IOA = 80%; Kappa = 60%).

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*Independent Variable*

Independent variable is described with replicable precision.

- Yes
- No
- N/A
- Unknown

Independent variable is systematically manipulated and under the control of the experimenter.

- Yes
- No
- N/A
- Unknown

Overt measurement of the fidelity of implementation for the independent variable is highly desirable.

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*Baseline*

The majority of single-subject research studies will include a baseline phase that provides repeated measurement of a dependent variable and establishes a pattern of responding that can be used to predict the pattern of future performance, if introduction or manipulation of the independent variable did not occur.

- Yes
- No
- N/A
- Unknown

Baseline conditions are described with replicable precision.

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*Experimental Control/internal Validity*

The design provides at least three demonstrations of experimental effect at three different points in time.

- Yes
- No
- N/A
- Unknown

The design controls for common threats to internal validity (e.g., permits elimination of rival hypotheses).

- Yes
- No
- N/A
- Unknown

The results document a pattern that demonstrates experimental control.

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*External Validity*

Experimental effects are replicated across participants, settings, or materials to establish external validity.

- Yes
- No
- N/A
- Unknown

Overall weighting of evidence: 3  2  1  0

*Social Validity*

The dependent variable is socially important.

- Yes
- No
- N/A
- Unknown

The magnitude of change in the dependent variable resulting from the intervention is socially important.

- Yes
- No
- N/A
- Unknown

Implementation of the independent variable is practical and cost effective.

- Yes
- No
- N/A
- Unknown

Social validity is enhanced by implementation of the independent variable over extended time periods, by typical intervention agents, in typical physical and social contexts.

- Yes
- No
- N/A

Unknown

Overall weighting of evidence: 3  2  1  0

Average across the 7 judgement areas = Sum of X / N = 9/7 = 1.29

X = individual quality rating for each judgement area

N = number of judgement areas

Overall weighting of evidence: 3  2  1  0

**Appendix E: Weight of Evidence B criteria**

Petticrew and Roberts detail the types of evidence aside from systematic reviews which are most suited to answering research questions concerning ‘effectiveness’ (Petticrew & Roberts, 2003). This was used to give each study a rating based on the table below.

Weighting	Study design
3	Randomised Controlled Trials
2	Cohort studies and Quasi-experimental studies
1	Qualitative research, Surveys, Case-control studies and Non-experimental evaluations

**Appendix F: Weight of Evidence C criteria**

Criteria	Weightings	Rationale
A Intervention	3 Intervention implemented by school staff with school resources	Intervention is more likely to be implemented if it can be implemented within the existing school resources
	2 Intervention implemented by specialist staff but on school grounds	
	1 Intervention requires significant specialist resources, specially trained staff or a residential aspect	
B Educational setting	3 A UK education setting	Results will be more generalisable to the UK educational setting
	2 An OECD education setting (apart from the UK)	
	1 An education setting not within the OECD	
C Intervention frequency	3 Intervention carried out at least once per week for at least one school year	Intervention is more in line with the principles of Forest School rather than a one off outward bound course.
	2 Intervention carried out regularly for less than a year	
	1 Intervention occurred on a one off block	

Criteria	Weightings	Rationale
D Participants	3 Participants have a diagnosis or Statement of Special Educational Needs and are screened and found to have social and emotional skill levels which would benefit from intervention prior to the study	The research question refers to interventions that are effective at improving social and emotional skills of pupils with SEN
	2 Participants all have a diagnosis or Statement of Special Educational Needs	
	1 Authors do not specifically state that all of the participants have identified SEN	

Scores across the four criteria above were averaged to give the overall Weight of Evidence C.

Study	A	B	C	D	Overall WoE C
Farnham & Mutrie (1997)	1	3	1	2	1.75
Fox & Avramidis (2003)	3	3	2	2	2.5
Price (2015)	1	3	3	2	2.25
Roe & Aspinall (2011)	3	3	1	1	2
Zachor et al. (2017)	2	2	2	2	2