The Existing Evidence-Base about the Effectiveness of Outdoor Learning

October 2015

Caroline Fiennes,[1] Elizabeth Oliver,[2] Kelly Dickson,[2] Diego Escobar,[2] Amy Romans,[3] Sandy Oliver[2]

[1] Giving Evidence
[2] UCL Institute of Education, EPPI-Centre
[3] Secondary school teacher

About The Authors

Giving Evidence is a consultancy and campaign, promoting charitable *giving* based on sound *evidence*.

Through consultancy, Giving Evidence helps donors and charities in many countries to understand their impact and to raise it. Through campaigning, thought-leadership and meta-research, we show what evidence is available and what remains needed, what it says, and where the quality and infrastructure of evidence need improving.

Giving Evidence was founded by Caroline Fiennes, a former award-winning charity CEO, and author of *It Ain't What You Give*. Caroline speaks and writes extensively about these issues, e.g., in the Stanford Social Innovation Review, Freakonomics, and the Daily Mail. She is on boards of The Cochrane Collaboration, Charity Navigator (the world's largest charity ratings agency) and the US Center for Effective Philanthropy.

The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) is part of the Social Science Research Unit (SSRU), UCL Institute of Education, University College London.

The EPPI-Centre was established in 1993 to address the need for a systematic approach to the organisation and review of evidence-based work on social interventions. The work and publications of the Centre engage health and education policy makers, practitioners and service users in discussions about how researchers can make their work more relevant and how to use research findings.

Founded in 1990, the Social Science Research Unit (SSRU) is based at the UCL Institute of Education, University College London. Our mission is to engage in and otherwise promote rigorous, ethical and participative social research as well as to support evidence-informed public policy and practice across a range of domains including education, health and welfare, guided by a concern for human rights, social justice and the development of human potential.

***

This work was a team effort involving: designing, co-ordinating and publicizing the study and drafting the report (CF, SO), designing and implementing the search strategy (KD), retrieving reports (DE), coding and analyzing systematic reviews (LO) and UK primary studies (DE, SO), and drawing out implications for schools and study centres (AR). All authors approved the final report.

The views expressed in this work are those of the authors and do not necessarily reflect the views of the EPPI-Centre or the funder. All errors and omissions remain those of the authors.

The study was guided by a Steering Group[1] drawn from the outdoor learning sector, and we are grateful to them for their input. It was conducted from April to September 2015.

Contact:

Caroline Fiennes, Director, Giving Evidence

+44 7803 954512, caroline.fiennes@giving-evidence.com

---

[1] This comprised: Andy Robinson (Chief Executive, the Institute for Outdoor Learning) – Chair; Jo Wells (Director, Blagrave Trust); Lucy Maynard (Head of Research, Brathay); Mike King (Releasing Potential and trustee of the Institute for Outdoor Learning); Justin Dillon (Bristol University and trustee of the Council for Learning Outside the Classroom); Emma Ferris (Head of Impact Evaluation, Outward Bound); Lyndsey Nassim (Head of Sales & Marketing, Scouts)

Table of Contents

Summary

Education programmes vary in their effectiveness. Mindful of this, the Blagrave Trust, whose areas include outdoor learning, wanted to understand what is known about the effectiveness of the various types of outdoor learning programmes. In partnership with the Institute of Outdoor Learning, the Blagrave Trust commissioned Giving Evidence and The Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI-Centre) at UCL Institute of Education to produce a systematic review of the existing literature about outdoor learning.

By searching the academic literature systematically, and inviting submissions of research ('crowdsourcing') from outdoor learning organisations, we aimed to:

1. Categorise the various outdoor learning (OL) activities being run in the UK, in order to provide a coherent sense of the sector as a whole;

2. Identify the various outcomes which organisations running outdoor learning activities are measuring, i.e., identify the outcomes which providers seem to be seeking to achieve; and

3. Assess the designs of individual evaluations (while aware that study designs vary in their openness to bias and hence inaccuracy) and the standard of evidence generally available for different types of outdoor learning.

This review benefited from a knowledgeable Steering Group with members drawn from provider and research organisations. It employed a systematic search of the academic literature and crowdsourcing of UK studies.

Summary of findings

We found that:

1. ***A sense of the sector as a whole***: There is no comprehensive or regular (repeated) survey of the scale of outdoor learning in the UK. There are some studies of specific outdoor learning activities (e.g., of particular types, or in particular parts of the UK). In these, some authors express concern about barriers to delivering outdoor learning and a reduction in outdoor learning.

2. ***The current research base:***

- Crowdsourcing UK research revealed an enthusiasm for research and sharing of knowledge amongst people who deliver outdoor learning activities. However, some of the material submitted were data or reflections which included named individuals, rather than anonymized research reports. This raises some issues around practitioners' understanding of research ethics.

- There is a growing body of individual studies and systematic reviews about the development and effectiveness of outdoor learning. We found 15 systematic reviews of the effects of outdoor learning. They provide extensive evidence of the effects of outdoor learning. However, the set is somewhat confusing because many of them overlap in terms of the primary studies they include. Moreover,

some systematic reviews include other systematic reviews, or are an update of an earlier review. This overlap therefore repeatedly reports the same evidence without necessarily strengthening it.

- Distinctions between types of interventions and outcomes employed to categorise studies are not always clear. For instance, 'healthy lifestyles' and 'health and well-being' were part of the 'learning and development' domain, while 'health behaviour' and 'health, physical / mental' were part of the 'health' domain.

- We found **58 primary UK studies**. Four features of them are striking:

  a. **They are spread thinly** across many populations (types and age groups), interventions, settings and outcomes, such that few topics have been researched more than a handful of times. This leads to our suggestion that the sector collectively identify and prioritise the important unanswered questions, and then focuses its (presumably limited) research resources on those priority questions.

  b. **The activities and participants on which studies focus** may not be where the sector would choose that research should focus. For example, the most common study topics are: adventure or residential activity; 11-14 year olds; and the general population. This leaves very few studies on (and hence little insight about) other age groups, popular activity such as Scouts or Ramblers, or people who are not in employment, education or training (NEET), have disabilities or are post-trauma.

  c. That there seems surprisingly **little linkage between the outcomes measured by the studies and the agenda of 'customers' and funders**. The outcomes measured are mainly around 'character development-type' outcomes (communication skills, teamwork, self-confidence etc. Very few studies addressed interventions with strong links to core curriculum subjects. There was only one primary study of educational outcomes at Key Stage 1 (5-7 year olds), few of educational outcomes at Key Stages 2, 3 and 4, and none at or beyond Key Stage 5 (sixth form). There is also a mismatch with the interests of employers: 'employability' is only measured in relation to offenders but not young people generally. Looking internationally, only six of the 15 systematic reviews looked at educational attainment, and only one addressed employability.

  d. **Safety is little covered** in the systematic reviews and was not measured as an outcome in any of the primary studies. Safety is obviously a major issue in outdoor learning since it can be dangerous: few social interventions can result in broken limbs or fatalities. Even if safety isn't the primary focus of a study, data could be gathered about safety: this is often how patient safety data and insights are gathered in medical research.

3. ***Outcomes assessed***: This evidence, both in the UK and internationally, and in both primary studies and systematic reviews, is very varied in terms of the populations who are offered outdoor learning, the type of outdoor learning and the outcomes assessed. The categorisation that informed this study captured some interventions and outcomes, but others emerged from the literature. Generally, there is considerable consensus in the general aims of interventions, but little consensus on the outcomes for assessing their effects.

4. ***The designs of individual evaluations***:

- We compared reports of UK studies in terms of attributes on a scale developed by Project Oracle, which looks at the extent of plans for an intervention and the evidence for it (described further in the document). Using this scale was challenging because the Project Oracle scale was designed for organisations to plan and assess their own interventions and evaluations, rather than to assess research reported elsewhere.

- Many UK studies did not reach Level One of the Project Oracle scale, normally because they did not cite or appear to use a Theory of Change (also known as a logic model: an articulation of the inputs, the intended outcomes, how the inputs are meant to produce those outcomes, and assumptions about context, participants or other conditions). Clear theories of change serve a couple of useful purposes: first, they demonstrate that the practitioners understand their intervention; and second, they are invaluable for other practitioners reading the research in estimating whether they will achieve the same outcomes with those interventions in their contexts. To be clear, a practitioner may have a theory of change but not cite it in their research, but (a) citing it in the research is useful and (b) experience from many other social sectors suggests that practitioners may need support to develop or articulate their theories of change.

- No UK study, or set of studies, featured the more demanding attributes of Levels Four or Five, around the intervention having been replicated in several places.

Implications for practice and policy

The study did not set out to look at implications of the research for practice and policy. Nonetheless, we found:

- Almost all outdoor learning interventions have a positive effect.

- The effect attenuates over time: the effect as measured immediately after the intervention is stronger than in follow-up measures after a few months. This is common for social interventions. However, one meta-analysis found that effects relating to self-control were high and were normally maintained over time.

- Evidence for the value of longer interventions. The systematic reviews found that overnight and multi-day activities had a stronger effect than shorter ones. While this is perhaps unsurprising, it does pose a challenge for funders / funding since it obviously forces a trade-off with the number of participants.

Recommendations

For providers of outdoor learning

Outdoor learning organisations can refer to systematic reviews of research about outdoor learning when planning their programmes. Careful reading is required to (a) check the rigour of each review and the studies they include (for instance, did the review include a systematic search and critical appraisal of the studies included?); and (b) check the precise types of programmes, populations and outcomes they studied.

Implications for the outdoor learning sector about developing its research

Because the existing research is spread quite thinly, few questions about effectiveness are yet answered reliably. We therefore recommend that the outdoor learning sector collectively prioritise the various unanswered questions in order to focus its research resources on those which are most important.

We recommend that the outdoor learning sector:

1. **Types and volume of activity**: Pull together the various data sources on this to give the current picture, and create a system to regularly capture data on the types and volumes of activity.

2. Improve practitioners' **theories of change**, enabling practitioners' to both create and to use them. Theories of change are explained in Box 4: they are invaluable for understanding why an intervention works and hence whether it is likely to work in other contexts, but only few evaluations of UK outdoor learning activity cited them.

3. Convene practitioners, researchers and others to **prioritise research topics**.

4. **Manage the resulting sector-wide research agenda**, through relationships with funders, and possibly by creating partnerships between practitioners and researchers.

5. Ensure that both **interventions and research are described clearly, fully and publicly.**

These recommendations are discussed more in Section 6.

Outdoor learning organisations need to have systems in place to support ethical practices for monitoring and research, particularly the storage and sharing of data from evaluations.

Greater consensus about the important outcomes of interest would allow research findings from different studies to be pooled more easily, and thereby facilitate accumulating knowledge to inform better the whole field.

Introduction

Professions and charities working for the public good have much to gain from sharing research about the issues they address, their efforts and achievements. The past 35 years has seen a growing emphasis on conducting research, and sharing and using the findings for improving, in particular, health, education and environmental conservation. Where these areas overlap is where many public, charitable and commercial services offer opportunities for outdoor learning.

The Blagrave Trust, which funds outdoor learning, recognised the value of underpinning outdoor learning with sound research and so commissioned this project to assess the nature and scale of research available to inform outdoor learning in the UK. The Trust, together with the Institute for Outdoor Learning, wanted to have a coherent picture of outdoor learning activities across the UK, and how they are being evaluated.

Giving Evidence, which works to enable charitable *giving* and charitable activity to be based on sound *evidence*, and the EPPI-Centre at University College London, which is committed to informing policy and professional practice with sound evidence, joined forces to provide this picture.

We made three important assumptions. First, that outdoor learning in the UK can be informed by research conducted outside as well as inside the UK, where contexts and activities are similar. Second, that research for informing outdoor learning comes from both organisations that specialise in research, such as universities, as well as organisations that specialise in delivering outdoor learning programmes. Third, that developing and evaluating outdoor learning suits the stepwise process recommended by Project Oracle[2] which is being increasingly adopted for youth development. Project Oracle's scale 'rates' what we know about interventions on whether there are: (1) detailed project descriptions and logic models; (2) before and after studies; (3) evaluations with a control group, which one would expect for interventions beyond the pilot stage; (4) replicated evaluations of impact; and (5) multiple independent evaluations in different settings, which may imply that further evaluations are less useful.

With these assumptions in mind, we searched sources of international research and invited UK outdoor learning organisations to contribute their own research. In doing this we aimed to:

1. Categorise the various outdoor learning (OL) activities being run in the UK, in order to provide a coherent sense of the sector as a whole;

2. Identify the various outcomes which organisations running outdoor learning activities are measuring, i.e., identify the outcomes which providers seem to be seeking to achieve; and

---

[2] "Project Oracle is a children and youth evidence hub that aims to improve outcomes for young people in London. We do this by building the capacity of providers and funders to develop and commission evidence-based projects, creating an ecosystem in which evidence is widely gathered, used and shared. We also work with specific "cohorts" or sub-sets of the sector to embed good practice, and at a national and international level to promote the wider use of evaluation and evidence. Project Oracle is funded by the Greater London Authority (GLA), the Mayor's Office for Police and Crime (MOPAC) and the Economic and Social Research Council (ESRC)."

3. Assess the designs of individual studies for what knowledge they can contribute about the development and evaluation of interventions in terms of the Project Oracle scale.

Our precise scope and the details of how we did this are described in Appendix 2. Here we describe what we found in terms of:

- o   Outdoor learning activity in the UK

- o   Overview of the international research about outdoor learning

- o   The effectiveness of various outdoor learning activities

- o   Coverage, design and findings of the primary research of activity in the UK.

We then discuss the implications for practice, policy, and guiding future research.

1. Outdoor learning in the UK

The diversity and scale of outdoor learning in the UK, being an overlapping patchwork of interests that differ in what is offered, to whom and where, is not captured by a single, unifying survey of activities. However, there are sources of information about various types of outdoor learning offered to different populations.

Scale of outdoor learning in the UK

The scale of outdoor learning was investigated by several surveys by different authors which spanned the past 45 years, three published since 2003, and an analysis of 13 UK surveys published between 1963 and 2009 (Lock 2010).

Historical data (1970 – 2000) and a survey (2002) of biology and geography teachers visiting one of 17 Field Study Centres (FSCs) (Tilling 2004) revealed **geography having ten times the amount of fieldwork at Key Stage**[3] **4 as biology**, and a **decline in the proportion of biology groups using the FSCs**; more specifically there was a 6% fall between 1998 and 2003. The factors most commonly identified by biology teachers were curriculum pressure, cost of courses, timetabling problems and a reduction in fieldwork opportunities in the curriculum.

Glackin (2007 and 2012) reports on a 2006 survey of how often **south London secondary schools use local green spaces and parks for teaching science**. Fewer than half the schools reported using fieldwork in teaching science; and only one school provided a residential science fieldtrip for either Key Stage 3 or 4. Frequent reasons for fieldtrips having a low priority were time, access to local sites, limitations of the National Curriculum, funding and safety.

When surveying fieldwork and outdoor visits and activities, but not the more local school grounds and community projects, Taylor et al. (2009) found an uneven distribution of school provision and local authority support for out-of-school learning at Key Stages 3 and 4 (high school). They drew on research from two projects: one investigating the availability of local authority outdoor education centres across England, and the other looking at the participation in out-of-school learning within secondary schools across the UK.

Most recent was Lock's (2010) investigation of the amount and type of biology fieldwork opportunities available to pre- and post-16 students (Key Stage 5) in the UK between 1963 and 2009. Thirteen pre-existing studies of the general school population were analysed.

Lock's (2010) findings suggest that there was a decline in biology fieldwork provision during the period studied, though there was not clear evidence that the number of habitats studied declined with this. Variability within the data-sets created difficulty in determining whether residential study had declined but the evidence from the Field Studies Council (FSC) strongly suggested this was the case over the last 30 years.

---

[3] UK education is divided into: Key Stage 1 (Years 1 and 2, when pupils are aged 5-7); Key Stage 2 (Years 3-6, when pupils are aged 7-11); Key Stage 3 (Years 7-9, when pupils are aged 11-14); Key Stage 4 (Years 9-11, when pupils are aged 14-16); and Key Stage 5 (years 12-13, when pupils are aged 16-18).

Lock (2010) concluded that there were eight factors related to the decline in fieldwork in the UK. Time and cost were the most strongly evidences factors. The more recent included studies cited the curriculum (syllabus, specification, scheme of work) and its assessment to be definitive negative factors. Lock (2010) reported that health and safety factors, teachers' enthusiasm and teachers' expertise were the remaining key factors contributing to the decline in biology field study in the UK.

Nature of outdoor learning in the UK

The nature of outdoor learning was investigated by two county surveys.

The National Federation of Education Research (2013) collected case studies during 2004-5 of special schools in Hampshire using (i) school grounds and gardens (ii) farms and city farms and/or (iii) field study / nature centres. The authors described five different approaches to outdoor learning, each of which places a different educational emphasis on: the experience of being outdoors; traditional learning practices linked to the formal curriculum; integrating themes or theories linking to general education; and learning unique to the outdoor environment.

Waite (2011) surveyed the practice and aspirations for learning outdoors for children aged 2-11 in one county in South West England. Her paper 'critically evaluates the implications of personal values associated with the outdoors including freedom and fun; ownership and autonomy; authenticity; love of rich sensory environment and physicality for pedagogical practice'.

The increased role of schools as the way that young people access the outdoors – e.g., since Youth Services ceased to exist – may mean that data from local authorities come to provide a more complete picture of outdoor activity. Over time, if the EVOLVE dataset covers all school activity, it may become very useful.

Nonetheless – and given that nobody would want a situation in which schools were the sole route for accessing the outdoors – it might be useful to create a complete and regular survey of outdoor learning activity. It might show that particular types of activity are rising or declining, or geographically patchy, and these insights might aid planning and policy discussions.

2. International research about outdoor learning activity

Focus of international research about outdoor learning

Research about these types of activities is commonly published in specialist journals of outdoor learning. The extent and type of this research was described in a study that analysed all the peer reviewed papers published 1998 – 2007 in the Australian Journal of Outdoor Education, the Journal of Adventure Education and Outdoor Learning and the Journal of Experiential Education (Thomas 2009).

Thomas et al. (2009) found 343 studies across the three journals and entire period (1998 - 2007). The focus of those papers is described by the papers' authors as outdoor education (21% of the papers), followed by adventure education (19%), adventure / wilderness therapy (14%), experiential education (13%), outdoor leadership (12%), service learning (8%), outdoor environmental education (7%), outdoor recreation (3%) and lastly expeditions (2%).

Explicitly contributing to the evidence about developing feasible, acceptable and effective programmes were studies focusing on 'programme design / facilitation' (16%) and studies focusing on 'outcomes / effects / participant experiences' (also 16%). Other areas of focus were teaching and teacher issues (14%), relationship with nature / self / others (10%), curriculum issues (9%), and safety and risk management (3%).

Outdoor learning research addressing effectiveness

As well as studies of children or young people (primary studies) we also found studies of prior research (systematic reviews of research literature) – see Box 1. We found 15 systematic reviews of the effects of outdoor learning. Not anticipating any systematic reviews specifically about outdoor learning, our original intention had been to include only studies published after 2003 (the date of the search by Rickinson et al (2004), in order to dove-tail with it). However, as the systematic reviews we found varied in scope and included studies published well before 2003, ultimately we chose not to exclude earlier reviews (Casson (1994); Hattie (1997); and Jill Dando Institute (2015)) which summarised the evidence presented by Wilson and Lipsey (2000) and Bedard (2004).

Systematic reviews cover much more ground than individual studies – for example, one of the systematic reviews we found analysed 150 studies, another included 58 studies, another covered 35 studies, etc. – and hence they are less open to bias than are individual studies, so we used them heavily though not exclusively. These systematic reviews, by including studies both from the UK and from other countries (e.g., Australia and North America), give a good picture of the global research base for outdoor learning. The list of systematic reviews is in Appendix 8, and each is outlined in Appendix 9.

> *Box 1: What are primary research and systematic reviews?*
>
> **Primary research** is a study of people. It can involve questionnaires, surveys or interviews, or other measurements about people such as their income or height.
>
> **A systematic review** is a study of studies. It is a structured investigation to find, critically appraise and synthesise all the relevant primary research on a specific topic. Systematic reviews are stronger than non-systematic 'literature reviews' in that they: (i) can reconcile differences in the conclusions of different studies by looking across a larger set of participants, (ii) identify gaps to inform further research, (iii) are more transparent and hence can be reproduced by other researchers in future and (iv) are less prone to bias, as science writer and doctor Ben Goldacre (2012) explains:
>
> *"Instead of just mooching through the research literature consciously or unconsciously picking out papers that support [our] pre-existing beliefs, [we] take a scientific, systematic approach to the very process of looking for evidence, ensuring that [our] evidence is as complete and representative as possible of all the research that has ever been done*.
>
> Thus a systematic review is more likely to be *accurate* and hence useful to practitioners for informing research and programme design than non-systematic literature. It is also more *credible* and hence useful in terms of convincing funders and policy-makers.

Each systematic review defines a ***scope*** (the topics, geography and timescale of interest) and the way that it will search for studies with that ***remit*** (the 'search strategy'). Most set some threshold for the ***quality*** of the primary studies they include in their analysis (the importance of quality of primary studies is discussed in Box 3). This is significant because the systematic review process is not magic: if the primary studies on which a systematic review is based are unreliable, the review's results will be unreliable. As a Yale cardiologist wrote recently on Twitter (Krumholz 2015): '*You can't just combine weak evidence and pretend that when mushed together it is strong. [Rather] it is meta-mush*.'

Scope of the systematic reviews

The scope of the systematic reviews reflects the prior interests of their authors, and possibly the topics addressed in the primary studies they included. Four systematic reviews were very specific in their focus. Stott (2013) assessed expeditions in terms of young people's personal growth. Two other reviews assessed adventurous activities and bushcraft for young offenders in terms of similar broad sets of outcomes (Stott (2013); Jill Dando Institute (2015)). Bowen (2013) addressed adventure therapy. Other systematic reviews were broader, and included a range of populations.

Five broad systematic reviews assessed outdoor learning in the areas of: mathematics (Hattie (1997); Rickinson (2004); Neill (2008a)); science (Rickinson (2004); Gill (2011)); reading / writing / language (Hattie (1997); Rickinson (2004); Gill ADD; Neill (2008a)); and problem solving (Hattie (1997); Neill (2008a)). Other areas of interest were: PSHE; environmental sustainability and design and technology (Rickinson (2004)); and creativity (Davies (2013)). Specific measures were rarely reported. An exception was Grade Point Average (Neill 2008a).

More interest was shown in a broad range of other learning and development outcomes. The range of outcomes employed in reviews for assessing different types of interventions appear in Table 1. The reviews addressed not only outcomes identified by the prior categorisation but also many others (in italics in Table 1). Outcomes have been clustered (as indicated by shading of rows) into those for personal / family development, social interaction, education and employment, relationship with nature, and health. Some reviews focused on particular areas in great detail. For instance, Davies (2013) addressed: motivation, engagement, enthusiasm, enjoyment, concentration, attention and focus associated with creativity initiatives.

Table 1 - Coverage of the included systematic reviews

Outcome	Outdoor learning for various populations	Expeditions for general population	Adventurous activity for higher education	Adventurous activity / bushcraft for young offenders	Outdoor learning for people with additional health needs
Self-awareness	Cason; Coalter; Gill; Gillis; Hattie; Neill a; Neill b; Rickinson	Stott		SMCI	
Self esteem				Jill Dando	
Self-control				*Jill Dando*	
Self-responsibility	Coalter; Gillis; Hattie; Neill a; Rickinson	Stott		SMCI	
Self-reliance				*Jill Dando*	
Self-concept					*Bowen*
Physical and Social resilience		*Stott*			
Independence				*Jill Dando*	
Persistence				*Jill Dando*	
Resourcefulness				*Jill Dando*	
Social development					*Bowen*
Family development					*Bowen*
Communication or teamwork	Coalter; Gill; Gillis; Hattie; Neill a; Neill b; Rickinson	Stott	Cooley	SMCI Jill Dando Institute	
Community integration	Coalter; Gillis; Hattie; Neill a; Rickinson	Stott	Cooley	SMCI	
Community leadership	Hattie	Stott	Cooley		
Youth leadership	Hattie; Neill b; Rickinson				
Employability				SMCI	

Outcome	Outdoor learning for various populations	Expeditions for general population	Adventurous activity for higher education	Adventurous activity / bushcraft for young offenders	Outdoor learning for people with additional health needs
Behaviour					*Bowen*
School adjustment				*Jill Dando*	
Educational attainment/ attendance	Cason; Gill; Neill a; Neill b		Cooley		
Attention/ creativity	Davies				
Recidivism				*SMCI*	
Relationship with nature	Gill; Hattie; Neill b; Rickinson	Stott			
Health lifestyles	Coalter; Gill; Rickinson			SCMI	
Health & well-being	Coalter; Gill; Gillis; Hattie; Higgins; Neill b; Rickinson				Bowen
Rehabilitation					*Bowen*

Although evidence of the effects of outdoor learning was beyond the scope of our study, once systematic reviews had been identified, the Steering Group expressed an interest in their findings. We therefore outline some insights here.

Headline findings

Although the systematic reviews varied in size and scope they reported consistent positive effects of the impact of outdoor learning activities, although their choice of outcomes varied. **They almost all report that the various outdoor learning activities have positive effects on all their various outcomes,** e.g., attitudes, beliefs, interpersonal and social skills, academic skills, positive behaviour, re-offending rates and self-image.

Longer programmes tend to be more effective than shorter ones. This fits with practice-based knowledge that length can allow for a more intensive and integrated experience and is obviously important given the pressure to cut length in order to reduce costs.

Strong benefits are also associated with well-designed preparatory work, and follow-up work.

Other findings

One review (Higgins (2013)) found adventure learning interventions to consistently show **positive benefits on academic learning** and wider outcomes such as self-confidence. On average, pupils who had participated in adventure learning interventions appeared to make approximately three additional months' progress. However, one analysis of 44 studies of 'ropes courses' (Gillis (2008)) found therapeutic and developmental effect sizes (see Box 2) of 0.53 and 0.47 respectively, but educational effects rather lower at 0.17. The highest effect sizes occurred in studies conducted in therapeutic settings, which the authors, Gillis & Speelman, hypothesised may be due to the nature of the populations studied.

There are some links between intervention and effect for which the **evidence is weak or absent**, for example on the psychological effects of mountaineering. Some interventions in some populations are shown by some studies to be harmful. For instance, 'research on mountaineering and rock climbing highlights the potential negative physiological impact of the activities (which can be minimised problems through adequate nutrition and hydration, and appropriate fitting and use of equipment)' (Coalter et al (2010)). Evaluations show considerable **variation in the sizes of effect** they find. This variation is probably explained by differences in programme type, duration and participants' age and characteristics. One assessment is that 65% of participants benefit (from adventure and bushcraft activity).

Two studies found a **decline in the amount of fieldwork**, despite evidence that it improves memory and social skills.

A review of 61 studies found evidence linking **forest schools** with improved social skills, self-control, self-confidence, language and communication (Gill (2011)).

A recent SR (Davies et al, (2013)) looked at 58 studies of school aged children and found that taking pupils out of the classroom and working **outdoors for part of their school time can foster creative development**. It

found other evidence that creative learning environments can aid children and young people's emotional development and social skills.

Box 2: Effect sizes

The term 'effect size' is used to compare the effects of interventions which can't be compared directly because they use different scales.

"Consider an experiment conducted by Dowson (2000) to investigate time of day effects on learning: do children learn better in the morning or afternoon? A group of 38 children were included in the experiment. Half were randomly allocated to listen to a story and answer questions about it (on tape) at 9am, the other half to hear exactly the same story and answer the same questions at 3pm. Their comprehension was measured by the number of questions answered correctly out of 20.

The average score was 15.2 for the morning group, 17.9 for the afternoon group: a difference of 2.7. But how big a difference is this? If the outcome were measured on a familiar scale, such as GCSE grades, interpreting the difference would not be a problem. If the average difference were, say, half a grade, most people would have a fair idea of the educational significance of the effect of reading a story at different times of day. However, in many experiments there is no familiar scale available on which to record the outcomes. The experimenter often has to invent a scale or to use (or adapt) an already existing one - but generally not one whose interpretation will be familiar to most people.
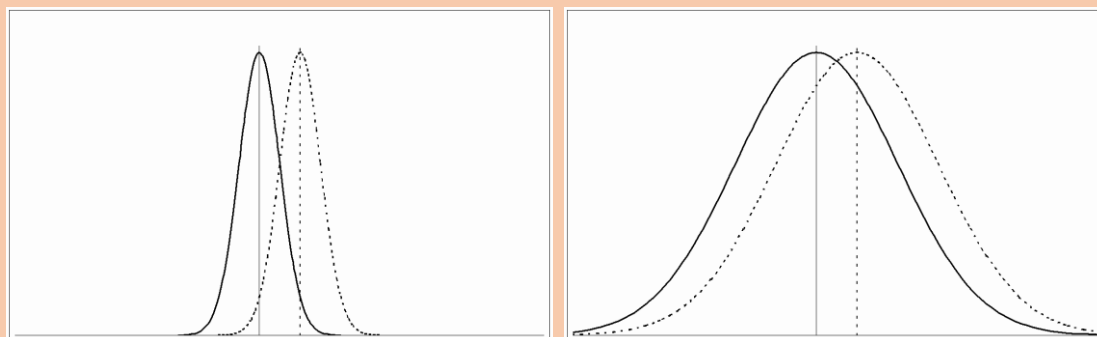
One way to get over this problem is to use the amount of variation in scores to contextualise the difference. If there were no overlap at all and every single person in the afternoon group had done better on the test than everyone in the morning group, then this would seem like a very substantial difference. On the other hand, if the spread of scores were large and the overlap much bigger than the difference between the groups, then the effect might seem less significant. Because we have an idea of the amount of variation found within a group, we can use this as a yardstick against which to compare the difference. This idea is quantified in the calculation of the *effect size*. The concept is illustrated in the figure above, which shows two possible ways the difference might vary in relation to the overlap. If the difference were as in the left graph it would be very significant; in the right graph on the other hand, the difference might hardly be noticeable." (Coe 2012)

Using 'effect size' allows comparison between the effects of interventions. In general, effect sizes of 0-0.2 are considered small, 0.5 is considered moderate, 0.8 or more is considered large (Research Rundowns,

undated). (An effect size of 0.2 means that the intervention increased the average score of the group by 0.2 standard deviations, i.e., 0.2 times the width of the distribution curve within the group. In the graph below, the area in dark blue marks one standard deviation either side from the average.)
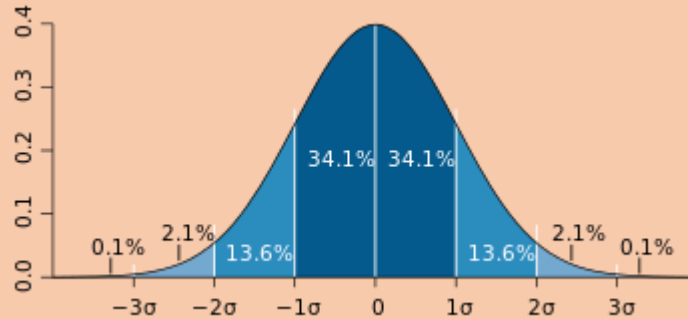
How dramatic is the effect on young people?

Most of the effect sizes reported are appreciable, and positive (examples in Figure 1). Two systematic reviews of adventure and bushcraft activity found effects varying in size: one found effect sizes ranging from -1.48 (i.e., a marked deterioration) to 4.26 (a vast improvement). The average effect size of 0.31 which it found was much more normal. The average effect sizes found in three systematic reviews are shown below.

Figure 1: Effect sizes for example interventions covered in selected systematic reviews

Average effect size reported in various systematic reviews of outdoor learning activity
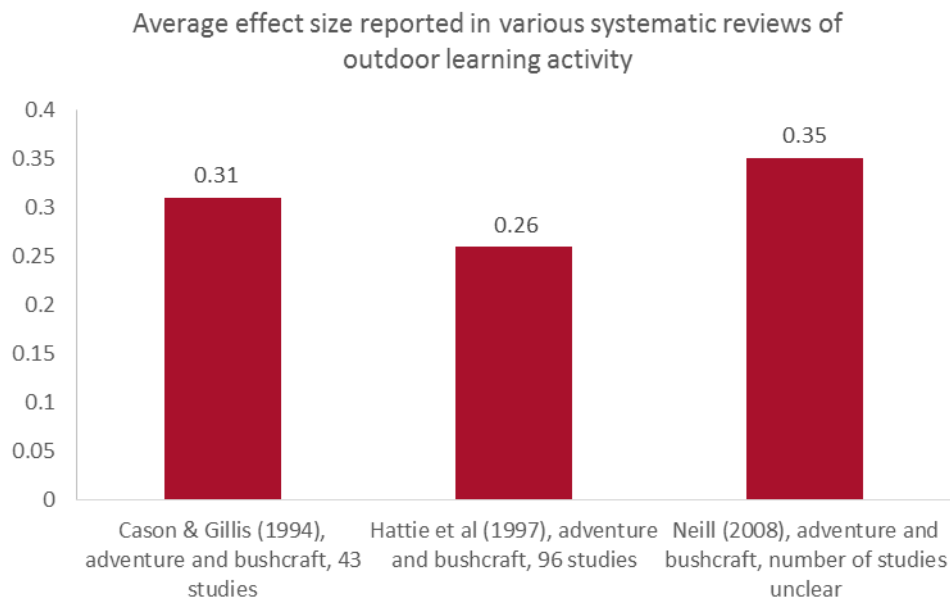
Does the effect on the young people endure?

The effect of most interventions attenuates over time: they show a stronger effect when measured immediately afterwards than when participants are followed up later. For instance, Hattie et al. (1997)

looking at adventure and bushcraft found outcome effects as shown below. However, that study also found that effects relating to self-control were high and were normally maintained over time.

Figure 2: Effect sizes for various outcomes immediately and at follow-up

Effect sizes reported in Hattie et al. (1997)
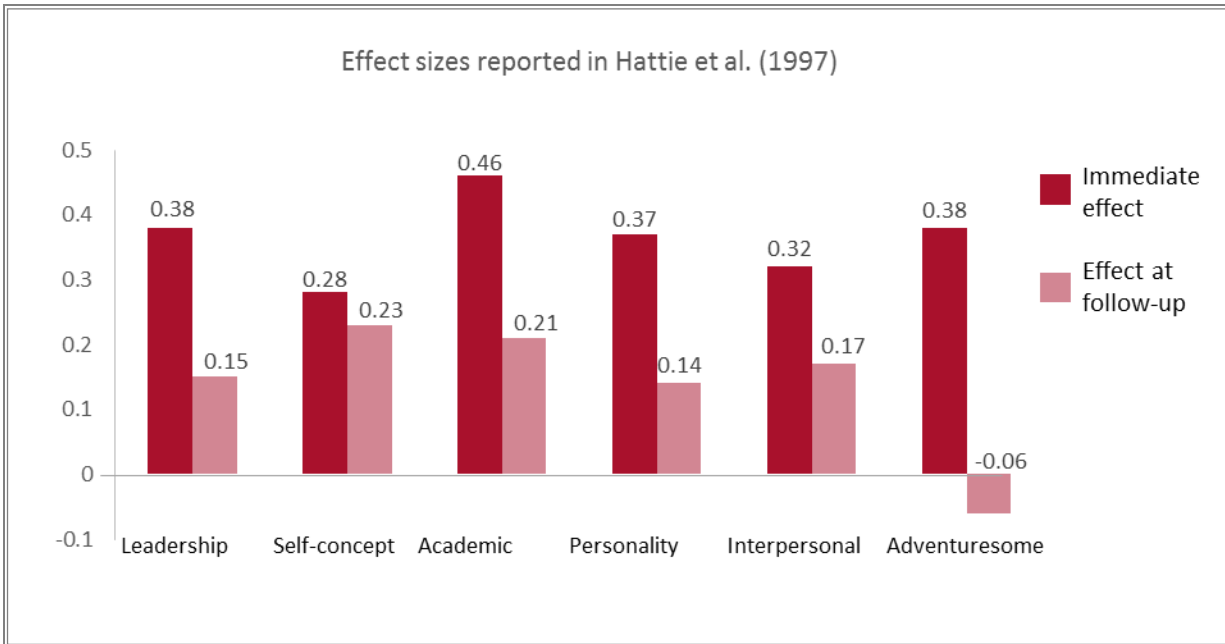
4. Outdoor learning research in the UK

Volume of the UK primary studies

As mentioned, we searched for material by (a) searching journals and databases for published studies, and (b) 'crowdsourcing' additional material through the Steering Group and blog posts. We then filtered for relevance.

Journal searches always produce masses of material, much of which turns out not to fit within the review's criteria (e.g., an article's title makes it look relevant but the abstract shows that it is not). The crowdsourced material was a mixture of published and unpublished material, evaluations, student theses, handbooks and guides for running interventions, and some raw data from participants (e.g., 'happy sheets'). Much of this material was illuminating and useful context even if it didn't fit the scope and therefore wasn't formally included in the analysis. For example, some submitted material was 'barely studies', e.g., teachers reflecting on practice and other teacher-to-teacher communications, which may be helpful for its intended audience, but doesn't fit *our particular purposes*. Appendix 5 shows the number of pieces of material filtered for the various reasons. Equally, there may be other relevant material which we didn't find, e.g., a couple of unpublished Masters theses came from the University of Edinburgh, suggesting that they may have more, but we did not go to them to seek out more.

We found 57 relevant primary studies which fitted our criteria around types of outdoor learning activity with people aged 5-25 normally resident in the UK, published since 2003, and which contribute to knowledge

about the effects of outdoor learning, from investigating theories of change or logic models, to assessing effects with a rigorous study design. These are in addition to the systematic reviews already discussed.

Focus of the UK primary studies

Below are tables which show the spread of those 58 primary studies by the type of activity, participant, setting, outcome etc. which they cover. A couple of comments on these:

First, these tables show that the existing research is spread fairly thinly: few topics have been researched many times. This is a problem because studies often vary in their answers – because they vary in their sample size, how they are done, their robustness, some are unlucky and get fluke results etc. – and hence the most reliable answers come only from combining multiple studies of the same topic.

Second, the spread of research may well reflect the funding and ease of conducting research. For example, few studies assess outcomes at primary school, which may simply reflect scarcity of funding at that stage. Residential activities are relatively well-studied, perhaps because they are a convenient population to study.

Third, many studies come under 'other'. This reflects the earlier comment about there being a wide range of outdoor learning activities and no standard categorisation of them.

Curriculum

Very few studies were of interventions with strong links to core curriculum subjects (Table 2).

Table 2: No. primary studies by type of activities and outcome (evaluated for impact on UK populations)

	Geography	Science	Maths	Physical education	Personal, health, social ed'n	Environmental sustainability	Other (specify)	Unclear (specify)
Field studies	1	3	0	0	0	0	2	2
Expedition(s)	0	0	0	0	3	0	0	3
Adventurous activity	1	1	1	0	3	0	3	7
Frequent adventurous activity (e.g. scouts, ramblers)	0	0	0	0	0	1	0	2
Nature visits	0	1	0	0	2	1	1	4
Other outdoor learner centred	1	3	2	0	4	2	2	13
Unclear	0	0	1	0	3	1	2	5

Setting

Most research looks at adventure activity, and/or residential experiences. Very little addresses 'frequent' adventurous activity, such as Scouts. 'Other' here includes studies on: childhood bereavement support, astronomy, developing group work skills, recreational activities.

Table 3: No. primary studies by type of activity vs settings (evaluated for impact on UK populations)

	School grounds	Residential facility	Local community	Other (specify)	Unclear
Field studies	5	4	5	3	1
Expedition(s)	0	6	0	0	0
Adventurous activity	2	12	1	1	0
Frequent adventurous activity (e.g. Scouts, Ramblers)	1	1	0	1	0
Nature visits	0	5	1	1	0
Bushcraft	0	0	0	0	0
Other outdoor learner centred	3	14	7	1	0
Unclear	2	7	1	2	1

Participants

Table 4 below shows that the UK studies are spread across many different types of people. Most concern the general population and only one looked at young people not in employment education or training (NEETs).

Table 4: No. primary studies by type of activities and participant (evaluated for impact on UK populations)

	NEETS	Non-engaged learners	With physical/ intellectual disabilities	Post trauma	Other special needs (specify)	General population	Other (specify)	Unclear (specify)
Field studies	0	1	0	0	1	2	1	1
Expedition(s)	0	1	0	0	2	0	4	0
Adventurous activity	0	3	1	1	3	3	1	3
Frequent adventurous activity (e.g. scouts, ramblers)	0	0	1	0	0	1	0	1
Nature visits	0	1	0	0	3	2	3	1
Bushcraft	0	0	1	0	2	5	5	1
Other outdoor learner centred	0	5	2	1	3	8	5	3

| Unclear | 1 | 1 | 0 | 0 | 2 | 1 | 1 | 5 |

Most studies concern 11-18 year olds, as shown in Table 5 and Table 6 below.

Table 5: No. primary studies by age of participant and setting

Code	School grounds	Residential Facility	Local community	Other (specify)	Unclear
5 – 10	3	11	5	9	0
11 – 14	3	22	4	4	1
15 – 18	3	18	2	2	0
18 – 25	0	10	2	1	1
Not stated	0	1	0	1	1

Table 6: No. primary studies by age of participant and type of activity

	Field studies	Expedition(s)	Adventurous activity	Frequent adventurous activity (e.g. scouts, ramblers)	Nature visits	Beach schools	Bushcraft	Other outdoor learner centred	Unclear
5 – 10	2	0	3	1	2	0	0	14	6
11 – 14	4	3	10	2	4	0	0	16	5
15 – 18	2	5	8	1	4	0	0	11	2
18 – 25	1	3	4	0	3	0	0	6	0
Not stated	0	0	0	1	0	0	0	0	2

Outcomes measured in UK primary studies

Strikingly few studies looked at educational attainment. None measured attainment in further education, higher education or vocational training. A few studies measured attainment at one or more of the Key Stages in primary or secondary education, up to age 16 (see Table 7) and these were spread over different types of interventions (Table 10).

Other studies addressed: school attendance, exclusion, school grades (unspecified), national exams, predicted and actual grades, knowledge and understanding, and skills achievement.

Table 7: Number of UK primary studies assessing educational attainment at each stage

Educational attainment	Count
Key Stage 1	1
Key Stage 2	5
Key Stage 3	7
Key Stage 4	3
Key Stage 5	10
Undergraduate degree	0
Postgraduate degree	0
Other (specify)	10
Unclear (specify)	4

Non-educational outcome have received much more research interest (Table 9) but are also widely spread across types of interventions (Table 10).

Table 8: Number of UK primary studies assessing non-educational outcomes

Non-educational development outcomes	Count
Curiosity	1
Relationship with nature	3
Self-awareness	2
Self-esteem	6
Self-responsibility	2
Communication or teamwork	13
Health & well being	10
Healthy lifestyles	3
Employability	3
Youth leadership	3
Community integration	3
Community leadership	1
Other (specify)	19
Unclear (specify)	3

Other outcomes included: creativity, commitment to learning, respect for self / others, sense of social responsibility, sense of belonging, addressing fear, tenacity, confidence, social skills, motivation, concentration, physical skills, resilience, social behaviour, direction, mindset, enjoyment, inspiration, impact on schools, family and community, critical thinking, self-determination, competence, relatedness, task approach, task avoidance, ego approach, ego avoidance, Relative Autonomy Index (RAI), interest effort, value autonomy-support, metacognition, problem-solving skills, optimism, pedagogical skills.

Table 9: No. UK primary studies by type of intervention and educational stage

	Key Stage 1	Key Stage 2	Key Stage 3	Key Stage 4	Key Stage 5	Further education	Vocational qualifications	Undergraduate degree	Postgraduate degree	Other (specify)	Unclear (specify)
Field studies	0	1	1	0	0	0	0	0	0	2	0
Expedition(s)	0	0	0	1	0	0	0	0	0	2	1
Adventurous activity	0	1	2	0	0	0	0	0	0	1	1
Frequent adventurous activity (e.g. Scouts, Ramblers)	0	0	0	0	0	0	0	0	0	1	0
Nature visits	0	0	1	1	0	0	0	0	0	3	1
Other outdoor learner centred	0	2	2	1	0	0	0	0	0	6	2
Unclear	1	2	2	1	0	0	0	0	0	0	0

Table 10: No. UK primary studies by intervention type and other learning / development outcome

Code	Curiosity	Relationship with nature	Self-awareness	Self esteem	Self-responsibility	Communication or teamwork	Health & well being	Healthy lifestyles	Employability	Youth leadership	Community integration	Community leadership	Other (specify)	Unclear (specify)
Field studies	0	1	0	0	0	2	0	0	0	1	0	1	1	1
Expedition(s)	0	0	0	2	0	1	0	0	0	1	0	0	2	0
Adventurous activity	1	1	1	2	1	6	1	0	1	2	2	1	6	0
Frequent adventurous activity (e.g. scouts, ramblers)	0	0	0	0	0	0	0	0	0	0	0	0	2	0
Nature visits	0	1	0	1	0	1	1	0	0	0	0	0	2	0
Other outdoor learner centred	0	0	0	4	1	4	2	0	2	0	1	0	8	0
Unclear	1	1	2	3	0	3	0	0	0	0	1	0	2	2

Discussion about the coverage and focus of the existing research

Four features of the UK-generated studies of outdoor learning are striking.

First, that they are **spread across many topics**. This precludes aggregating studies to reach more reliable answers. It leads to the suggestion that the sector collectively identifies and prioritises the important unanswered questions, and then focuses its (presumably limited) research resources on those priority questions. This is discussed further below.

Second, **the activities and participants on which studies focus** may not be where the sector would choose that research should focus. For example, as mentioned, the most common study topics are: adventure or residential activity; 11-14 year olds; and the general population. This leaves very few studies on (and hence little insight about) other age groups, popular activity such as Scouts or Ramblers, or people who are NEET, have disabilities or are post-trauma.

Third, that there seems surprisingly **little linkage between the outcomes measured by the studies and the agenda of 'customers' and funders** – notably employers, schools, and local authorities. The outcomes measured are mainly around 'character development-type' outcomes (communication skills, teamwork, self-confidence etc.). Schools in England are expected to promote 'spiritual, moral, social and cultural development' of their pupils (Ofsted 2015), and there could be a more explicit link between this type of development and the outcomes of outdoor learning.

Similarly, Science, Technology, Engineering and Maths (STEM subjects) are a priority of the government in schools (Department for Innovation, Business and Skills 2015). Yet, very few studies addressed interventions with strong links to schools' core curriculum subjects and there was only one primary study of educational outcomes at Key Stage 1 (5-7 year olds), fairly few of educational outcomes at Key Stages 2, 3 and 4, and none at or beyond Key Stage 5 (sixth form). Equally there could perhaps be some studies of the extent to which outdoor learning promotes engagement with and performance in STEM subjects.

Lastly, as well as being responsible for their student cohort as a whole, teachers have a particular responsibility for groups of pupils including: those with special educational needs (SEN); those of high ability (gifted and talented, G&T); those with English as an additional language (EAL); those with disabilities; and those who receive Pupil Premium (PP) funding (e.g., those who get free school meals). Teachers would therefore benefit from research reports that explicitly link their study populations with these characteristics.

We take no view here on whether non-educational outcomes are important, but rather notice the mismatch between research topics and the pressure schools face to achieve those educational outcomes. There is also a mismatch with the interests of employers: 'employability' is a measured outcome highlighted in relation to offenders but not young people generally. Various other employment-relevant outcomes are measured (e.g., team-work, communication) but the studies don't link them to employability. Looking internationally, only six of the 15 systematic reviews looked at educational attainment, and only one addressed employability.

Overall, we wondered whether the studies reflect the outcomes about which practitioners are enthusiastic – and where research can be funded and run easily – rather than where the policymakers focus. Charitable funders sometimes focus away from educational attainment, to build more rounded characters, deliberately because public policy tends to leave this gap.

Fourth, **safety is also little covered in both the systematic reviews and the primary studies**. It was not measured as an outcome in any of the primary studies. Safety is obviously (presumably) a major issue in outdoor learning since it can be dangerous (many social interventions can be harmful, though few can result in broken limbs or fatalities); perception of danger deters many schools, parents and other potential customers (it is identified in the systematic review by Rickinson et al (2004) as one of key five barriers to outdoor learning, and again in a 2006 survey of London schools (Glackin (2007)); and many children and young people are perhaps raised with inadequate exposure to risk which outdoor learning can usefully provide. Safety does feature in some of the material we received and found, such as handbooks and practitioners descriptions of how they organise their activities, which was interesting and useful even though these materials were not 'studies' suitable for inclusion in our formal analysis. Even if safety isn't the primary outcome of a study (which would probably be rare), it would be possible to gather data about safety: this is often how patient safety data and insights are gathered in medical research.

5. UK primary studies contributing to the development and evaluation of interventions

We found 57 primary studies of outdoor learning in the UK, and compared their features with Project Oracle's stepwise process for developing and evaluating interventions. Project Oracle's five levels are:

1. Project model and evaluation plan (We know what we want to achieve)

2. Indication of impact (We have seen there is a change)

3. Evidence of impact (We believe there is a change caused by us. We can make this happen consistently)

4. Model ready (We know why and how the change happened. This works elsewhere)

5. System ready (We know why and how the change happened. This works everywhere).

Further detail on Project Oracle's levels is in Appendix 10.

In summary, reports did not feature all the attributes of any individual level (1-5). Attributes at Levels One or Two were more common than attributes matching higher levels. This is not altogether surprising, since Project Oracle has (at the time of writing) yet to find any interventions which attain Levels Four or Five. Below, we illustrate each level of Project Oracle's scale by citing one or more outdoor learning studies, describing how they met key attributes of Project Oracle's levels[i]. First though, Box 3 explains why research methods and quality matter.

UCL
Institute of Education

Institute for Outdoor Learning

THE Blagrave TRUST
Supporting young people

Giving Evidence
Enabling **giving** based on sound **evidence**

Box 3: Why we evaluate research methods

i) Because different research methods give different answers

"Two men say they're Jesus: One of them must be wrong" (Dire Straits lyric!)

The table below shows the effect of a reading programme in India measured using several research methods (Innovations for Poverty Action). These methods all used the same outcome measures, but the experimental designs were different.

The answers vary widely: some suggest that it works well, others show it to be detrimental. Clearly there can only be one correct answer! All the other answers are incorrect: and could mislead donors or practitioners to implement this programme at the expense of another which might be better.

Method	Impact Estimate
(1) Pre-post	26.42*
(2) Simple Difference	-5.05*
(3) Difference-in-Difference	6.82*
(4) Regression	1.92
(5) Randomized Experiment	5.87*

*: Statistically significant at the 5% level

The answers vary because research methods vary in how open they are to biases (i.e., systematic errors). For instance, suppose that a medical trial involves giving patients a drug for two years. Suppose that that drug has horrible side-effects such that during the two years, some patients can't stand taking it so they drop out of the trial (or worse, perhaps the drug kills some of them). If the trial only collects data on patients who are still in the trial after two years, it will systematically miss the important insights about those side-effects. This 'survivor bias' will make the drug look more effective than it really is.

Somebody reading the trial results without knowing that detail wouldn't be able to distinguish the actual effect of the drug from that of this survivor bias. Similarly, if a study only looks at the outcome (in the example above, it's reading level) before the programme and then afterwards (i.e., is a pre-post study), it won't be possible to distinguish whether any improvement in reading levels was due to the programme or just to the fact that children learn over time anyway.

{As an aside, contrary to popular myth, it is not invariably the case that robust research is more expensive than unreliable research, nor that randomised controlled trials (the most reliable design for a single primary study) are invariably terribly expensive: many are cheap or free. See Appendix 12.}

ii) Because weaker research methods allow for more positive findings

The UK National Audit Office searched for literally every published evaluation of a UK government programme (National Audit Office 2013). Of those, it chose a sample, and ranked on one hand, the quality of the research method ('robustness' on the x axis, i.e., how insulated the study is from bias), and on the other, the positive-ness of the programme ('claimed impact').

The trend line on the resulting graph below would slope diagonally downwards. It shows that more robust research only allows for modest impact claims whereas weak research allows much stronger claims.

Bad research can be persuaded to say almost anything, and won't allow researchers to distinguish the effects of a programme from other factors (e.g., the passage of time, the mindset of participants, other programmes) nor from chance.

Most social interventions have a small effect and a reliable research method will show what that is: bad research is likely to overstate it. The highest estimate for the reading programme above is from the pre-post study which is a weak study design.

Relationship between robustness and claimed impacts in evaluations

Note
1   Robustness assessed on Maryland Scale. Assessed effectiveness, rated low to high.
    Low = Small or insignificant effects.
    2   = Mixed effects, positive for some, negative or insignificant for others.
    3   = Positive effects, with some caveats or uncertainties noted.
    High = Significant positive impacts, no or only minor caveats or uncertainties noted.

Source: National Audit Office analysis of external assessment by London School of Economics

This relationship between weak research methods and positive findings has been shown also in medical research. We found it in the studies of outdoor learning too.

Project model and evaluation plan (Level One)

Sandford et al (2008, 2010) is an example which shares the attributes of Project Oracle's Level One. It studied an Outward Bound outdoor physical activity programme that sought to promote the personal, social and educational development of disaffected or disengaged young people with the help of volunteer mentors as informal educators. Their outcomes of interest (behaviour, attendance and self-esteem) were all assessed through teachers' perception. They had referred to existing research to develop their programme and had drawn a clear logic model (see Box 4) that linked:

* resources (finances, people, networks and equipment) to activities
* activities (set up and delivery of courses)
* outcomes (young people's behaviour)
* impact (the fundamental changes in organisations, systems, communities etc.)

Observations, informal conversations, interviews and focus groups confirmed the potential for mentors to function as informal educators in such youth programmes. They also highlighted areas for improvement before the programme was ready for a rigorous evaluation of its effects, namely better planning and understanding of how mentors and other youth professionals can work with each other.

Another Outward Bound programme (O'Brien (2014)) had a similarly clear aim, 'to unlock the potential of young people through discovery and adventure in the wild', by applying a series of mindset'[4] interventions during a five day outdoor personal development programme. Although the logic model was less clear than that of Sandford et al (2010), the three psychological measures employed had been validated by prior research: the General Self Efficacy scale for measuring a sense of personal competence to deal effectively with a selection of stressful situations; the Resilience Scale for measuring five facets of resilience including perseverance and self-reliance; and the Implicit Self Theories scale for measuring the participants' theory of intelligence (Mindset).

Indication of impact (Level Two)

The same authors presented before and after data (Sandford et al 2008) in terms of teachers' perceptions that indicated potential beneficial effects.

Evidence of impact (Level Three)

O'Brien's (2014) study progressed further than reporting indicative impact. Not only did she apply measures before and after the intervention, she also compared these measures with young people allocated randomly to receive or not the Mindset interventions.

Another small randomised controlled trial (White (2012)) similarly specified a rigorous research design and measure, but without an explicit logic framework.

Model evidence and system ready (Levels Four and Five)

We were unable to find replicated studies that took into account differing contexts and that were sufficiently well documented for wider implementation.

**Box 4: Theories of change**

What is a theory of change?

A theory of change (or logic model: we use the terms interchangeably) is what is meant by Project Oracle's Level 1's 'we know what we want to achieve' and 'project model' (i.e., articulation of how the activities are supposed to create the intended impact). It lays out the assumptions behind an intervention, and links between activities and intended impacts (i.e., how the activities are supposed to produce those impacts, and what is assumed, e.g., parental engagement, weather…) They allow organisations to find and cite evidence suggesting that their activities are likely to produce their target outcomes.

A clear theory of change also helps other organisations considering running the intervention to see whether the assumptions are likely to hold in their contexts, i.e., whether they're likely to get similar results. It also helps other organisations make good decisions about what outcomes to try to achieve by showing what's involved in the interventions which ostensibly deliver them.

---

[4] Mindset is a positive psychology theory. It will be given a capital throughout to emphasise that it is the theory which is being referred to rather than a more generalised conception of the term.

The diagram below shows the constituent pieces of a logic model:
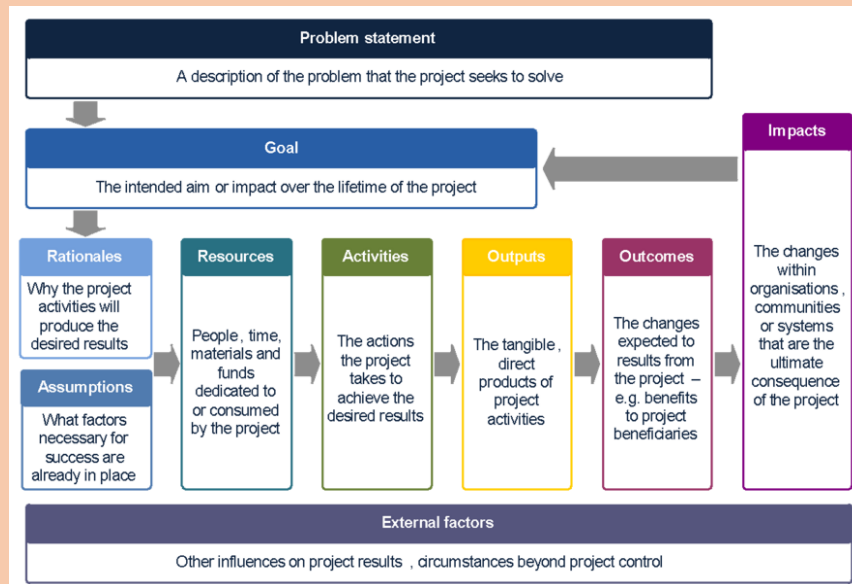
Why does having a clear logic model matter?

A clear logic model is important / essential to intelligent programme design because it enables predictions about whether a type of intervention is likely to work (for a specific population). An evaluation *without* a clear logic model simply shows *whether* a programme worked and the extent to which it worked: it gives no indication of *why* it worked (or not) – why it gets those results. That is, without a logic model, the intervention is like a black box: we gain no insight into whether it's likely to achieve those results again, nor elsewhere. It adds nothing to the 'science' (i.e., understanding) of these interventions. By contrast, if a provider starts with a clear logic model, they can use the existing research to see which parts are likely to be true, which are not evidenced, and therefore can:

(a) make an educated estimate of whether, when and for whom the intervention is likely to work,

(b) identify major risks and unsupported assumptions,

(c) change the design to make it more likely to succeed. It may transpire that the proposed logic model is totally fanciful and implausible, and hence this work will prevent them running a pointless intervention, or even a harmful intervention. And

(d) identify what needs testing. Maybe very little needs testing and so the practitioner is spared all the cost and hassle of evaluating.

In short, it enables practitioners to *use* existing research, rather than solely to *produce* research. Clearly this is more efficient. The focus on impact has led many organisations (particularly charities) to produce research often bad quality), when (i) they are not set up nor incentivised to be researchers, and (ii) it might be more useful for them to leverage the (better quality) research which already exists[5].

---

[5] Caroline Fiennes has written about this elsewhere, e.g., www.giving-evidence.com/M&E

6. Discussion

Implications for practice, and for funding practice

The clearest implications for delivering and funding practice is around the value of longer interventions. The systematic reviews found that overnight and multi-day activities had a stronger effect than shorter ones. (For example, see Hattie et al. (1997).) While this is perhaps unsurprising, it does pose a challenge for funders / funding since it obviously forces a trade-off with the number of participants.

Implications for research and the sector more widely

The major insights are around various improvements which could be made to the research-base for outdoor learning in the UK: a better research-base would inform practice and discussions with policy-makers, and could potentially increase funding by enabling the sector to make a stronger case about its value.

Categorising types of activity

As mentioned, creating a more shared language around the categories of activity might be useful. It might help practitioners find research relevant to their work which can inform their work, help funders to find practitioners whose work fits their goals and areas of interest, help with comparing interventions (e.g., for effectiveness, or value-for-money) and help with assessing the volume of activity (see below). The categories we used may be a good starting point (though the sector might want some subdivisions to disaggregate the 'other' category).

Assessing the amount of the various types of activity

Again, a regular assessment of the amount of outdoor learning activity of various types – encompassing school-based and non-school-based – would probably be a useful addition. It would inform policy discussions by showing if the amounts of an effective activity are declining or patchy, can inform practitioners as to geographic regions which are particularly under-served etc. The assessment could be a survey though other tools may be appropriate too.

Practitioners' clarity about their operational models

Few of the research reports cite, or appear to draw on, a clear logic model / theory of change. It may be that practitioners have theories of change but which don't appear in the research reports, though experience in other sectors suggests that probably few practitioners have them. It might be useful to support practitioners to create them. The process of creating a theory of change can improve performance by making practitioners clear about what they are trying to achieve (an internal Ford Foundation document says that 'people rarely succeed in achieving difficult things they never commit to achieving'). It also exposes the assumptions underpinning the work (e.g., around child development and psychology) which can be tested for their validity (e.g., by using the existing literature). The theory of change itself is useful in highlighting activities which are run but don't contribute to success, or it is essential to replicating and scaling up because they help other practitioners see whether the intervention is likely to work in their contexts. They are essential to good

evaluation, i.e., not only gaining insight into *whether* something works, but also *why* it works, and thereby building a cumulative 'science' of these interventions.

Hence outdoor learning providers should be encouraged to create clear logic models, and trained to do so if necessary.

However, it is not unusual for practitioners to be reticent about sharing the details of their interventions and their theories of change. Normally, practitioners have no incentive to do so, and indeed have a strong incentive against sharing them: intervention (which produces the alleged outcomes) is how practitioners compete for funding. Hence not sharing is a rational response to the incentives on the individual organisation, but it prevents the sector as a whole from learning, improving, replicating successful work, and making a strong case for increasing its collective funding. The outdoor learning sector may be able to solve this incentive problem by making accreditation dependent on sharing details of interventions.

Box 5 - Describing an intervention

Medical research has guidelines for describing interventions such that somebody else can replicate them accurately. They have a 12-point checklist for describing interventions, the Template for Intervention Description and Replication (TIDieR) (Hoffman et al 2014), which is helpful and could easily be adapted for outdoor learning. It has been adapted elsewhere, e.g., by mental health charities (Kent County Council, 2014):

- The name of the intervention (brief name or phrase)
- The way it works (rationale, theory, or goal of the essential elements)
- What materials and procedures were used (physical or informational)
- What (each procedure, activity, and/or process)
- Who provided the intervention (e.g., nurse, psychologist, and give their expertise and background)
- How was it delivered (e.g., face to face, online, by phone, and whether it was provided individually or in a group)
- Where it took place
- When and how much (the number of sessions, schedule, dosage and duration)
- Tailoring (what if anything could be adapted to the individual, why and by how much)
- Modifications which happened after the study started
- How well was adherence to the plan assessed (i.e., the process for assessing adherence)
- The extent to which implementation adhered to the plan.

Prioritising research topics to create a sector-wide research agenda

Many sectors have a mismatch between what researchers research and what practitioners want researched. (It is terrible and well-documented in medical research, for example see a series of papers in The Lancet (2014).) This arises because researchers and practitioners are often somewhat separate, their incentives are around different topics, and they rarely intuit each other's interests. Hence a formal process can be needed to determine and agree research priorities.

We recommend that the sector decide what activities, participants and outcomes are most important for it to understand, and focus research on those. To be clear, we are not here advocating that those be educational attainment or employability, but are suggesting that **the decision of where research focuses be made deliberately**.

Box 6: Prioritising research topics: lessons from the James Lind Alliance

In medical research, there is a terrible mismatch between the topics which get researched and the topics which concern patients. Acne is an example, having been largely overlooked by researchers.

To address this systemic problem, The James Lind Alliance[6] (JLA), a UK not-for-profit organisation, consults with patients affected by a particular condition (say, asthma or Parkinson's) together with their carers and clinicians, and runs a structured consultation exercise to identify 'unanswered questions' about effective treatments, and to collectively prioritise the top 10 questions[7]. For example, in its work on asthma, the JLA found that patients were keen to know if there is value in the breathing exercises they're asked to do, which research didn't answer.[8]

These Priority Setting Partnerships (PSPs) are often initiated by patient groups, or by doctors or hospitals interested in an area underrepresented in health research.

Key insights from the James Lind Alliance's process are:
- It is essential to **define the stakeholders who get a say** in the process: who can propose and prioritise research topics. For JLA, that excludes researchers, because JLA was set up to increase the influence of non-researchers in the choice of research topics – but for other sectors, the 'answer' here may be different.
- Also, **define the types of eligible questions**. JLA only considers 'treatment effectiveness': it does not consider causes of diseases, nor questions like 'why is my doctor horrible' or 'why isn't this better funded'.
- It is important to **define what is meant by 'unanswered'**. For JLA, 'unanswered' means that there has been no systematic review of it. (Presumably there must be a threshold for the quality of the systematic review, and the number and quality of the primary studies included in the review. JLA has a handbook on how to do PSPs[9].)

The process for a PSP is:

---

[6] The James Lind Alliance. *Priority Setting Partnerships* [Online]
http://www.lindalliance.org/Priority%20Setting%20Members.asp [Accessed: 22.10.13]
[7] The James Lind Alliance. *Welcome to The James Lind Alliance* [Online] http://www.lindalliance.org/ [Assessed: 10.13].
[8] Cowan, K. (2013) Senior Advisor to the James Lind Alliance, Interview with Fiennes, C., 07.13.
[9] Co-authored by one of us, Sandy Oliver.

a. Survey, of anybody eligible. Often PSPs are funded by patient groups (e.g., Asthma UK, the Parkinson's Association). This survey generates a long list of ideas – some of which are research questions. A qualitative researcher goes through to extract the research questions. Funding normally determines the number of respondents they try to get / number of people to whom the survey is sent.

b. Determining what's in scope, i.e., extracting the questions about research or treatment. These are 'ostensibly unanswered questions'.

c. Check against the evidence base, i.e., what research is there into those questions already? Are they really outstanding? Enables *evidence-based research. Avoids waste* by avoiding asking questions to which the answer is already known. That gives a long list of unanswered relevant questions. Usually 50-100.

d. Get people to vote to whittle them down to 20-30.

e. Prioritisation: One day workshop of <30 people to get to 10. There is then work to turn these topics / questions into 'research questions', which normally need to be much more specific.

JLA has found value in – and recommends – having clarity about the criteria, boundaries and process.

Its model is light: each Priority Setting Partnership (PSP) is run by only one 'senior advisor', who may run five or six of them at any one time. The senior advisor is supported by the others listed plus some administrator time. The PSPs cost about £20-100k each.

The aim of prioritising research topics is to influence research, so it is important to engage researchers and research funders. The JLA does this by being somewhat integrated into the National Health System research funding system. For the outdoor learning sector, there may need to be some strategy of engaging with donors, academics (both in relation to their own research and that of graduate students), and bodies such as Natural England.

"Ask an important question and answer it reliably"

This is a central tenet of clinical research (Yusuf 1984). In many social sectors, research addresses many questions which are not particularly important, and provides only unreliable answers. On the first half of the tenet, prioritising research questions, and creating and executing a research agenda shared by the sector can ensure that research addresses only questions agreed to be important. And the second, providing reliable answers will be easier if practitioners have (and research cites) clear theories of change. Some co-ordination / 'policing' of research may be needed to ensure that provides reliable answers (to be precise, to ensure that the research designs are capable of providing reliable answers). The amount of research which the sector does should be governed by the need to produce reliable answers: two unreliable answers are worth much less than one reliable answer, so the sector should only allocate resource to research likely to be reliable.

Experience elsewhere suggests that few practitioners have the skills or funds (or incentive) to produce reliable research. Hence partnerships between practitioners and researchers can be useful: e.g., part of Project Oracle's model is brokering such partnerships. The Institute for Outdoor Learning or other sector body could perhaps play a similar role.

That research should be reported clearly. Several systematic reviews are unclear about surprisingly basic dimensions, such as their scope, the number of studies they included, and the age of young people studied.

7. Conclusion and summary of recommendations

We found more research into outdoor learning than we were expecting. That said, there remain many unanswered questions – more than the finite research resources available can reliably answer – and therefore we suggest that the outdoor learning sector build more 'infrastructure' to make best use of those research resources and rapidly build an evidence base which will serve it well. Specifically, we recommend:

6. **Types and volume of activity**: Pulling together the various data sources on this to give the current picture, and creating a system to regularly capture data on the types and volumes of activity.

7. Improving practitioners' **theories of change**: both enabling them to create them, and to use them.

8. Convening practitioners, researchers and others to **prioritise research topics**.

9. **Managing the resulting sector-wide research agenda**, through relationships with funders, and possibly creating partnerships between practitioners and researchers.

10. Ensuring that both **interventions and research are described clearly, fully and publicly.**

_____

References

Journal analysis

Thomas G, Potter T, Allison P (2009) A Tale of three journals: An investigation of the development and futures of AJOE, JAEOL and JEE. *Australian Journal of Outdoor Education* 13: 16-29.

UK surveys

Glackin M, Jones B (2012) Park and Learn: Improving Opportunities for Learning in Local Open Spaces. *School Science Review*. 93: 105-113.

Taylor C, Power S, Rees G, Taylor C (2009) *Out-of-school learning: the uneven distribution of school provision and local authority support*. United Kingdom.

Tilling S (2004) Fieldwork in UK Secondary Schools: Influences and Provision. *Journal of Biological Education*. 38: 54-58.

Waite S (2011) *Teaching and learning outside the classroom: personal values, alternative pedagogies and standards*. United Kingdom.

Systematic reviews

Bowen DJ, Neill JT (2013) A Meta-Analysis of Adventure Therapy Outcomes and Moderators. *The Open Psychology Journal* 6: 28-53.

Cason D, Gillis H (1994) A meta-analysis of outdoor adventure programming with adolescents. *Journal of Experiential Education* 40-47.

Coalter F, Dimeo P, Morrow S & Taylor J (2010) *The Benefits of Mountaineering and Mountaineering Related Activities: A Review of Literature: A Report to the Mountaineering Council of Scotland. Mountaineering Council of Scotland*. British Mountaineering Council. Department of Sports Studies, University of Stirling.

Cooley Sam, Burns Victoria, Cumming Jennifer (2015) The role of outdoor adventure education in facilitating groupwork in higher education. *Higher Education* 69: 567-582.

Davies D, Jindal-Snapeb D, Collier C, Digbya R (2013) Creative Learning Environments in Education – a systematic literature review. *Thinking Skills and Creativity* 8: 80-91.

Gill T (2011) *Children and nature. A Quasi-systematic review of the empirical evidence*. Greater London Authority.

Gillis Lee H; Speelman Elizabeth (2008) Are Challenge (Ropes) Courses an Effective Tool? A Meta-Analysis. *Journal of Experiential Education*. 31: 111-135.

Hattie John, Marsh H W; Neill James T; Richards Garry E; (1997) Adventure education and outward bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research*. 67: 43-87.

Higgins S, Katsipataki M, Kokotsaki D, Coe R, Major L E; Coleman R (2013) *The Sutton Trust-Education Endowment Foundation Teaching and Learning Toolkit: Technical Appendices*.

Jill Dando Institute (2015) *Wilderness Challenge Programmes*.

Lock Roger (2010) Biology Fieldwork in Schools and Colleges in the UK: An Analysis of Empirical Research from 1963 to 2009. *Journal of Biological Education*. 44: 58-64.

Neill J (2008) Meta-Analytic Research on the Outcomes of Outdoor Education. *6th Biennial Coalition for Education in the Outdoors Research Symposium*, Bradford Woods, IN, 11-13 January, 2002.

Neill J (2008) *Enhancing life effectiveness. The impacts of outdoor education programs*. University of Western Sydney .

Rickinson M, Dillon J, Teamey K, Morris M, Choi M Y; Sanders D, Benefield P (2004) *A review of research on outdoor learning*. National Foundation for Educational Research.

SMCI Associates (2013) *'Living Wild - Chance for Change' Highland LEADER Programme: Evaluation Report.* East Lothian: SMCI Associates.

Stott T, Allison P, Felter J, Beames S (2015) Personal development on youth expeditions: A literature review and thematic analysis. *Leisure Studies* 34(2): 197-229

UK primary studies

Amos R, Reiss M (2012) The Benefits of Residential Fieldwork for School Science: Insights from a Five-Year Initiative for Inner-City Students in the UK. *International Journal of Science Education*. 34: 485-511.

Barry M (2010) *An Evaluation of Venture Trust's Moving On – Willow Partnership Project*. University of Strathclyde, UK.

Beames S (2004) Overseas youth expeditions with Raleigh International: a rite of passage*? Australian Journal of Outdoor Education*. 8: 29-36.

Beames S (2005) Expeditions and the social construction of the self. *Australian Journal of Outdoor Education.* 9: 14-22.

Birdwell J, Wybron I (2014) *Scouting for Skills*. Demos, London.

Bowey L, McGlaughlin A (2006) The Youth Crime Reduction Video Project: An Evaluation of a Pilot Intervention Targeting Young People at Risk of Crime and School Exclusion. *Howard Journal of Criminal Justice*. 45: 268-283.

Bowker R (2004) Children's Perceptions of Plants Following Their Visit to the Eden Project. *Research in Science and Technological Education*. 22: 227-243.

Brewer J Sparkes AC; (2011) The Meanings of Outdoor Physical Activity for Parentally Bereaved Young People in the United Kingdom: Insights from an Ethnographic Study. *Journal of Adventure Education and Outdoor Learning* 11: 127-143.

Brown D, Neale N, Francis R (2011) Peak into the Past: An Archaeo-Astronomy Summer School. *School Science Review.* 93: 83-90.

Christie B, Higgins P, McLaughlin P (2014) 'Did you enjoy your holiday?' Can residential outdoor learning benefit mainstream schooling? *Journal of Adventure Education & Outdoor Learning*. 14: 1-23.

Cooley S, Holland M, Cumming J, Novakovic E, Burns V (2014) Introducing the use of a semi-structured video diary room to investigate students' learning experiences during an outdoor adventure education groupwork skills course. *Higher Education*. 67: 105-121.

Cramp A (2008) Knowing Me Knowing You: Building Valuable Relationships outside the Classroom. *Education* 3-13. 36: 171-182.

Crosbie J (2014) *The value of outdoor education for people with disabilities: An in-depth case study of the Calvert Trust.* Doctoral/PhD. University of Edinburgh.

Davis B, Waite S (2005) *Forest Schools: an evaluation of the opportunities and challenges in Early Years Final Report*. University of Plymouth.

Dillon J, Morris M, O'Donnell L, Raid A, Rickinson M, Scott W (2005) *Engaging and Learning with the Outdoors – The Final Report of the Outdoor Classroom in a Rural Context Action Research Project*. National Foundation for Education Research, Slough.

Dismore H, Bailey R (2005) "If Only": Outdoor and Adventurous Activities and Generalised Academic Development. *Journal of Adventure Education and Outdoor Learning*. 5: 9-19.

Dourneen J (2010) *Setting free the spirit of English in the Brecon Beacons: the value of learning outside the classroom*. United Kingdom.

Fabian H (2005) *Outdoor learning environments: easing the transition from the foundation stage to Key Stage One*. United Kingdom.

Fielding S, Jones M (2014) Children Go Bonkers about Bugs. *Primary Science*: 23-27.

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

Gerret K (2014) *Cedar Road Primary School Aberdovey Analysis* 2013 / 2014.

Halocha J (2005) Developing a Research Tool to Enable Children to Voice Their Experiences and Learning through Fieldwork. *International Research in Geographical and Environmental Education*. 14: 348-355.

Heeley S (2013) *The Value of Outdoor Education For Young People With Learning Difficulties: A Case Study of Arthog Outreach And Southall School*. [dissertation for BA Honours degree].

Hinds J (2011) Woodland Adventure for Marginalized Adolescents: Environmental Attitudes, Identity and Competence. *Applied Environmental Education and Communication*. 10: 228-237.

Horseman L (2010) *Reflections and Evaluation on Forest School projects 2010*. Kindling.

Humberstone B, Stan I (2009) Well-being and outdoor pedagogies in primary schooling: the nexus of well-being and safety. *Australian Journal of Outdoor Education*. 13: 24-32.

Humberstone B, Stan I (2011) Outdoor Learning: Primary Pupils' Experiences and Teachers' Interaction in Outdoor Learning. *Education* 3-13. 39: 529-540.

Humberstone B, Stan I (2012) Nature and Well-Being in Outdoor Learning: Authenticity or Performativity. *Journal of Adventure Education and Outdoor Learning*. 12: 183-197.

Loynes C (2010) Journeys of Transition: The Role of Narrative within the Stoneleigh Project. *Journal of Adventure Education and Outdoor Learning*. 10: 127-145.

Maynard T, Waters J (2007) Learning in the Outdoor Environment: A Missed Opportunity?. *Early Years: An International Journal of Research and Development*. 27: 255-265.

Maynard T, Waters J, Clement J (2013) Child-Initiated Learning, the Outdoor Environment and the "Underachieving" Child. *Early Years: An International Journal of Research and Development*. 33: 212-225.

McArdle K, Harrison T, Harrison D (2013) Does a Nurturing Approach That Uses an Outdoor Play Environment Build Resilience in Children from a Challenging Background?. *Journal of Adventure Education and Outdoor Learning*. 13: 238-254.

McManus J (2012) The Thurston Family Project: Working with families through outdoor activities and resiliency training. *Psychology of Education Review*. 36: 40-45.

Newman M, Higgins P, McLaughlin P, Allison P (2004) Outdoor Experiential Education Programmes for Socially Excluded Young People An Evaluation of the Venture Trust's 10-day Pilot Programme.

North Lanarkshire Council and the Outward Bound Trust (2014) *North Lanarkshire Council and the Outward Bound Trust Interim Evaluation. 2013-14*. The Outward Bound Trust.

O'Brien L (2009) Learning Outdoors: The Forest School Approach. *Education* 3-13. 37: 45-60.

O'Brien K (2014) *Developing a Growth Mindset through Outdoor Personal Development: Can an intervention underpinned by psychology increase the impact of an outdoor learning course for young people?* [MSc dissertation]

PACEC (2011) *Impact assessment evaluation of The Scout Association*. 84.

Peacock A (2006) *Changing Minds The lasting impact of school trips. A study of the long-term impact of sustained relationships between schools and the National Trust via the Guardianship scheme*. University of Exeter.

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: 110-122.

Heaven A, Browne A (2008) *A small scale evaluation of existing practice, monitoring systems and outcomes for vulnerable children.* Lindley Educational Trust.

Riley K, Ellis S, Weinstock W, Tarrant J, Hallmond S (2006) Re-engaging disaffected pupils in learning: insights for policy and practice. *Improving Schools*. 9: 17-31.

Kenny R (2010) *Forest School and the Early Years Foundation Stage – An explorative case study.*

Sandford RA; Duncombe R, Armour KM; (2008) The Role of Physical Activity/Sport in Tackling Youth Disaffection and Anti-Social Behaviour. *Educational Review*. 60: 419-435.

Sandford RA; Armour KM; Stanton DJ; (2010) Volunteer Mentors as Informal Educators in a Youth Physical Activity Program. *Mentoring & Tutoring: Partnership in Learning*. 18: 135-153.

Scrutton RA; (2015) Outdoor adventure education for children in Scotland: Quantifying the benefits. *Journal of Adventure Education and Outdoor Learning*. 15: 123-137.

SMCI (2013) *'Living Wild – Chance for Change' Highland LEADER Programme*. SMCI Associates.

Sharpe D (2014) Independent Thinkers and Learners: A Critical Evaluation of the "Growing Together Schools Programme". *Pastoral Care in Education*. 32: 197-207.

Slade M, Lowery C, Bland K (2013) Evaluating the Impact of Forest Schools: A Collaboration between a University and a Primary School. *Support for Learning*. 28: 66-72.

Social Finance, Adviza Partnership (2014) E*nergise Innovation Limited a Year in Review October 2012 – October 2013*.

Sproule J, Martindale R, Wang J, Allison P, Nash C, Gray S (2013) Investigating the Experience of Outdoor and Adventurous Project Work in an Educational Setting Using a Self-Determination Framework*. European Physical Education Review*. 19: 315-328.

Stott T, Hall N (2003) Changes in aspects of students' self-reported personal, social and technical skills during a six week wilderness expedition in Arctic Greenland. *Journal of Adventure Education and Outdoor Learning.* 3(2): 159-169.

Waller T (2007) The Trampoline Tree and the Swamp Monster with 18 Heads: Outdoor Play in the Foundation Stage and Foundation Phase. *Education* 3-13. 35: 393-407.

White R (2012) A Sociocultural Investigation of the Efficacy of Outdoor Education to Improve Learner Engagement. *Emotional & Behavioural Difficulties.* 17: 13-23.

Williams R (2013) Woven into the Fabric of Experience: Residential Adventure Education and Complexity. *Journal of Adventure Education and Outdoor Learning.* 13: 107-124.

Williams R (2013) *Report to The Outward Bound Trust on an Evaluation of The Classic Course.*

York Consulting (2015) *Paul Hamlyn Foundation Evaluation of Learning Away: Final Report.*

Other references

Coe, R. (2012) It's the Effect Size, Stupid: What effect size is and why it is important [Online] http://www.leeds.ac.uk/educol/documents/00002182.htm [Accessed: 22.10.2015].

Goldacre B (2012) *Bad Science*. Fourth Estate.

Hoffmann, T.C., et al. (2014) Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *The British Medical Journal*. 348.

Innovations for Poverty Action

Kent County Council (2014). HeadStart Kent [Online] http://www.kelsi.org.uk/__data/assets/powerpoint_doc/0010/25588/Programme-Board-Main-Presentation-19-11-14-AF-UN.pptx [Accessed: 22.10.15].

Krumholz H. (@hmkyale) "You can't just combine weak evidence and pretend that when mushed together it is strong. It is meta-mush." 30 August 2015. Tweet. [Online] https://twitter.com/hmkyale/status/638013905055444992 [Accessed: 22.10.15].

The Lancet (2014). *Research: increasing value, reducing waste* [Online] www.thelancet.com/series/research [Accessed 22.10.15].

Marie Curie's Analytics Team (example of Theory of Change framework)

National Audit Office (2013) *Evaluation in Government.*

Research Rundowns. *Effect Size*. [Online] https://researchrundowns.wordpress.com/quantitative-methods/effect-size/ [Accessed: 22.10.15].

Yusuf S, Collins R, Peto R (1984) Why do we need some large, simple randomized trials? *Statistical Medicine* 3(4):409-22.

Appendices

Appendix 1: Composition of Steering Group

The Steering Group met once at the beginning of the process to discuss scope; once part-way through to discuss the findings emerging from the analysis and the implications for the outdoor learning sector; and once when the research was complete and we were starting to draw conclusions and draft this report. It was pivotal for crowdsourcing studies to include. It comprised:

Andy Robinson (Chief Executive, the Institute for Outdoor Learning) – Chair

Jo Wells (Director, Blagrave Trust)

Lucy Maynard (Head of Research, Brathay)

Mike King (Releasing Potential and trustee of the Institute for Outdoor Learning)

Justin Dillon (Bristol University and trustee of the Council for Learning Outside the Classroom)

Emma Ferris (Head of Impact Evaluation, Outward Bound)

Lyndsey Nassim (Head of Sales & Marketing, Scouts)

Appendix 2: Scope and methods

This appendix describes how we set about the following activities:

1. Categorise the various outdoor learning (OL) activities being run in the UK, in order to provide a more coherent sense of the sector as a whole;

2. Identify the various outcomes which organisations running outdoor learning activities are measuring, i.e., identify the outcomes which providers seem to be seeking to achieve; and

3. Assess the designs of individual evaluations (mindful that study designs vary in their openness to bias and hence inaccuracy) and the standard of evidence generally available for different types of outdoor learning.

First, the Institute of Outdoor Learning and the Blagrave Trust convened a Steering Group, which included funders, practitioners and academics. The members are listed in Appendix 1.

Second, we worked with the Steering Group to **define the scope** of the systematic review, i.e., the types of outdoor learning interventions of interest, the population of interest, and types of studies.

We included studies of:

- children and young people, aged 5 – 25 who are participants or volunteer helpers;
- outdoor learning in the education sector, or the health sector; and
- assessments of educational attainment, other learning or development.

We excluded:

- studies where outcomes relate solely to health
- studies of 'sport and leisure': this study is primarily about education-related outcomes.

The studies needed to be:

- systematic reviews of outdoor learning research; or
- primary studies of children and young people normally living in the UK (including studies addressing foreign trips or sailing trips), and which used the following designs:
  - o Detailed descriptions
  - o Before and after studies
  - o Controlled trials
  - o Reporting the delivery or receipt of programmes
  - o Surveys of practice
  - o Methods studied, particularly of outcome measure;
- published 2003 or later. This is because a similar systematic review of outdoor learning activity, published in 2004, had already synthesised the relevant primary research conducted until 2003, so this enabled us to dovetail with that.

Third, we **defined the search strategy** and **searched for studies** within that scope. As Goldacre says, we do that scientifically and systematically. The search process we defined had two parts. First, searching journals

and databases for relevant academic and published non-academic studies, for which we defined up-front the databases and search terms we would use. And second 'crowdsourcing' additional studies: this we did through the Steering Group and posts on various blogs.

Fourth, we **analysed the studies** which fit within our scope. For the systematic reviews we found, we looked at their coverage and implications. For the primary studies, we analysed their focus in terms of the types of interventions and outcomes they assess, and their quality.

This 'plan' for the research was published as a **protocol**.[10]

Towards the end of this study, Giving Evidence and the EPPI Centre started one on Sail Training[11]. It is a somewhat different design and purpose, though towards the same goal of understanding the effect of the interventions.

Appendix 3: Search strategy

We sought studies by searching bibliographic databases that record predominantly academic studies, and crowdsourcing studies from organisations active in the sector.

The first route involved choosing key words likely to describe relevant studies and combining them in a string to search databases electronically. Two sets of key words will describe (a) children and young people; and (b) outdoor learning. By combining the key words they identified studies that addressed outdoor learning with children and young people. We searched databases that compile research in the areas of education, psychology and other social sciences.

The second route involved circulating a request for studies through outdoor learning networks across the UK.

The third route was through searching the list of contents of two specialist journals:

- The Journal for Adventure and Experiential Outdoor Learning
- Field Studies.

Studies were checked against the inclusion criteria. Initially, two researchers checked studies and compared their decisions. Subsequently one researcher applied the inclusion criteria and another checked their decisions. They discussed discrepancies and, where necessary, amended the definitions of terms. Records of studies will be managed with software designed for systematic reviewing to maximize efficiency and accuracy when identifying and analyzing studies.

---

[10] http://eppi.ioe.ac.uk/cms/Default.aspx?tabid=3477
[11] More detail is at www.giving-evidence.com/sailing-start

Appendix 4: Electronic search strategy

We searched the following databases:

1. Education Resources Information Centre (ERIC) http://eric.ed.gov/
2. British Education Index (BEI) https://www.ebscohost.com/academic/the-british-education-index
3. AEI
4. Applied Social Science Index and Abstracts (ASSIA)
5. PsychInfo http://www.apa.org/pubs/databases/psycinfo/
6. Child and Adolescent development studies

The search strategy for ERIC is shown below. This was adapted for the other databases.

String	
Population S1	(((SU.EXACT("Children") OR SU.EXACT("adolescents") OR (SU.EXACT("Young Children") OR SU.EXACT("Young Adults")) OR (SU.EXACT("Out of School Youth") OR SU.EXACT("Disadvantaged Youth")) OR SU.EXACT("Special Needs Students") OR (SU.EXACT("College Students") OR SU.EXACT("Low Ability Students") OR SU.EXACT("Middle School Students") OR SU.EXACT("Elementary School Students") OR SU.EXACT("Junior High School Students") OR SU.EXACT("At Risk Students") OR SU.EXACT("Secondary School Students") OR SU.EXACT("High School Students") OR SU.EXACT("Minority Group Students")) OR SU.EXACT("Dropouts")) AND la.exact("English")) OR (((SU.EXACT("Parents") OR SU.EXACT("Teachers") OR SU.EXACT("Youth Leaders") OR SU.EXACT("School Counselors") OR SU.EXACT("Mentors")) AND la.exact("English")) OR (((ti((Child* OR "young person")) OR ab((Child* OR "young person")) OR ti(("young people" OR "young woman")) OR ab(("young people" OR "young woman")) OR ti(("young women" OR "young man")) OR ab(("young women" OR "young man")) OR ti(("young men" OR boy*)) OR ab(("young men" OR boy*))) OR (((ti((girl* OR adolescent*)) OR ab((girl* OR adolescent*)) OR ti((Teen* OR "school child*")) OR ab((Teen* OR "school child*")) OR ti(("school student*" OR "school child*")) OR ab(("school student*" OR dropouts)) OR ti(("special needs" N2 child* OR "special needs" N2 student*)) OR ab(("special needs" N2 child* OR "special needs" N2 student*))) AND la.exact("English")) OR (((ti(("Learning disab*" N2 child* OR "Physical disab*" N2 child*)) OR ab(("Learning disab*" N2 child* OR "Physical disab*" N2 child*)) OR ti(("Learning disab*" N2 student* OR "Physical disab*" N2 student*)) OR ab(("Learning disab*" N2 student* OR "Physical disab*" N2 student*))) OR (((ti((parent* OR family*)) OR ab((parent* OR family*)) OR ti((carers OR guardians)) OR ab((carers OR guardians)) OR ti((mother* OR father*)) OR ab((mother* OR father*))) OR (((ti((Teacher* OR "special education" N2 teacher*)) OR ti(("youth worker" OR "youth leader")) OR ti((Mentor OR "School Counsellor"))) OR (ti(therapist N2 youth) OR ti(("School Counselor*" OR Counselor N2 child*)) OR ab(("School Counselor*" OR Counselor N2 child*)) OR ti((counselor N2 student* OR counselor n2 youth)) OR ab((counselor N2 student* OR counselor n2 youth)) OR ti((counselor n2 "young people" or "young person" OR therapist

	n2 "young people" or "young person")) OR ab((counselor n2 "young people" or "young person" OR therapist n2 "young people" or "young person")) OR ti((therapist N2 child OR therapist N2 student*)) OR ab((therapist N2 child OR therapist N2 student*)) OR ab(therapist N2 youth))
Intervention S2	((SU.EXACT("Outdoor Education") OR SU.EXACT("Adventure Education") OR SU.EXACT("Field Trips") OR SU.EXACT("Museums") OR su((adventurebased therapy OR adventurebased counselling)) OR su((adventure therapy OR adventure counselling))) OR (ti((sport* OR activity*)) AND ti("experiential learning") OR ab((sport* OR activity*)) AND ab(("experiential learning" OR "experiential learning"))) OR (ti((adventure counseling OR adventure based counseling)) OR ab((adventure counseling OR adventure based counselling)) OR ti((adventure therapy OR adventure based therapy)) OR ab((adventure therapy OR adventure based therapy)))) OR (ti(("outdoor education" OR outdoor learning)) OR ab(("outdoor education" OR outdoor learning)) OR ti(("field trips" OR "field studies")) OR ab(("field trips" OR "field studies")) OR ti((adventure education OR adventure-based learning)) OR ab((adventure education OR adventure-based learning)) OR ti(("adventure learning" OR "outward bound")) OR ab(("adventure learning" OR "outward bound")))
Combined	S1 and S2

Appendix 5: Numbers of studies identified

The figure below shows the numbers of studies identified and how some that were not relevant were excluded from the analysis.

| 57 crowd sourced | → | 3536 titles & abstracts | ← | 3479 electronic search (+39 duplicates) |

Excluded: 308
182 non UK primary study
0 not outdoor learning
0 before 2003
17 no empirical data
38 no education/dev outcome
1 not age 5-25
12 only educators
0 not education/health sector
5 employee development
15 non-systematic review
38 could not access full report

385 titles & abstracts

Excluded: 3151
1157 non UK primary study
1110 not outdoor learning
300 before 2003
223 no empirical data
142 no education/dev outcome
97 not age 5-25
72 only educators
27 not education/health sector
22 non-systematic review
1 focus on play

77 full reports

4 UK surveys

16 systematic reviews

57 primary UK studies

Appendix 6: Outdoor learning activity of interest to the sector

The following categorisation is based on work from the Institute of Outdoor Learning and refined by the Steering Group. It informed our work and the categorisation we used.

Category	Targeted Outcome	Practice / Discipline / Activity
1. Education	a. Key Stage (KS) 1, 2 & 3 topics (geography, science, maths, PE, PSHE ...) b. Key Stage 4 GCSEs c. A-level, BTEC, Degree, PGCE, MA d. Environmental / sustainability e. Non-mainstream learners	i. Field Studies ii. Adventurous Activity iii. Nature visits iv. Residentials v. Forest Schools/Beach Schools vi. Bushcraft vii. Other outdoor learner centred viii. Forest schools ix. School grounds
2. Personal Development	a. Self-awareness b. Self-responsibility c. Communication & teamwork d. Health & Wellbeing	i. Adventurous Activity residentials ii. Frequent Adventurous Activity e.g. Scouts, Ramblers... iii. Expeditions iv. Duke of Edinburgh
3. Youth & Community	a. Employability b. Diversionary c. Youth leadership d. Community integration e. Community leadership f. Healthy lifestyles g. Rehabilitation	i. Outdoor sport training & qualifications ii. Adventurous activity days & residentials iii. Programmes of multiple activity interventions
4. Therapy	a. Physical / intellectual disabilities b. Family c. Post Trauma d. Rehabilitation	i. Prescribed outdoor activities ii. Group/family adventure based residential iii. One to one therapy outdoors

Not included:

Category	Targeted Outcome	Practice / Discipline / Activity
5. Early Years / Play	a. Curiosity b. Relationship with nature c. Risk management	i. Outdoor nursery ii. Forest Schools
6. Sport & Leisure	a. Competitive performance b. Introduction to sport c. Recreational competence d. Group leadership e. Train the trainer f. Family activities	i. NGB coaching programmes ii. Activity sessions in wider programmes iii. Mass participation events – e.g. adventure races, OPAL etc. iv. Local sports clubs

7. Adventure Tourism	a. Introduction to sport b. Personal development c. Environmental / cultural awareness	i. Expeditions ii. Holiday camps iii. Adventure holidays iv. Day events...
8. Teacher / Practitioner training	a. Understanding of usage options b. Personal design, delivery & evaluation competence	i. ITT college based outdoor activity ii. INSET days in school grounds iii. CPD workshops & conferences iv. FE & HE courses v. Professional accreditation schemes
9. Employee development	a. Teamwork b. Communication c. Planning & problem solving d. Innovation e. Leadership & management	i. 'Away Days' ii. Adventurous activity residentials for apprentice/graduate iii. Management development programmes with outdoor exercises

Appendix 7: Coding tool for describing outdoor learning in this study

Outdoor learning: coding tool

Questions	Answers	Definitions / notes / examples
Section A: Core keywords		
A.1 Name of reviewer	A.1.1 Details (specify)	
A.2 Linked reports	A.2.1 None / not known	
	A.2.2 Linked (specify)	
	A.2.3 Unclear (specify)	
A.3 Language of main report	A.3.1 English	
	A.3.2 Other (specify)	
Section B: Study characteristics		
B.1 Form of publication	B.1.1 Journal article	
	B.1.2 Technical report (specify)	
	B.1.3 Dissertation/thesis (specify)	
	B.1.4 Other (specify)	
B.2 Year of publication	B.2.1 Before 2001	
	B.2.2 2001-2005	
	B.2.3 2006-2010	
	B.2.4 2011-2015	
B.3 Broad aims of the study	B.3.1 Not stated	
	B.3.2 Explicitly stated (specify)	
	B.3.3 Implicit (specify)	
	B.3.4 Unclear (specify)	
B.4 Study funding	B.4.1 Not stated	
	B.4.2 Programme provider (specify)	
	B.4.3 Programme funder (specify)	
	B.4.4 Other (specify)	
	B.4.5 Unclear (specify)	

Questions	Answers	Definitions / notes / examples
B.5 When was the study conducted?	B.5.1 Not stated	Pertinent dates are for the start and end of data collection
	B.5.2 Initial year (specify)	
	B.5.3 Final year (specify)	
	B.5.7 Unclear (specify)	
Section C: Study population		
C.1 Age of children	C.1.1. Children: under 5/ pre-school	
	C.1.2 Children: 5-10	
	C.1.3. Children: 11-14	
	C.1.4 Young people: 15 - 18	
	C.1.5 Young people: 18+	
	C.1.6 Other	
	C.1.7 Unclear	
C.2 Other characteristics	C.2.1 NEETS	
	C.2.2 Non-engaged learners	
	C.2.3 Physical / intellectual disabilities	
	C.2.4 Post trauma	
	C.2.5 Other special needs (specify)	
	C.2.6 General population	
	C.2.7 Other (specify)	
	C.2.8 Unclear (specify)	
C.3 Who else is involved?	C.3.1 Parents	
	C.3.2 Family (specify)	

Questions	Answers	Definitions / notes / examples
	C.3.3 School teachers	
	C.3.4 Nursery / playgroup staff	
	C.3.5 Other adults (specify)	
	C.3.6 Unclear	
Section D: Outcomes		
D.1 Educational attainment	D.1.1 Key Stage 1	
	D.1.2 Key Stage 2	
	D.1.3 Key Stage 3	
	D.1.4 Key Stage 4/ GCSEs	
	D.1.5 Key Stage 5/ AS, A level	
	D.1.6 Further Education (HNC/HND)	
	D 1.7 BTECs, OCR Nationals and other vocational qualifications	
	D 1.8 Undergraduate Degree (BA, BSc)	
	D 1.9 Postgraduate Degree (MA, MSc)	
	D.1.5 Other	
	D.1.6 Unclear	
D.2 Educational sphere	D.2.1 Geography	
	D.2.2 Science	
	D.2.3 Maths	
	D.2.4 PE	
	D.2.5 PSHE	

Questions	Answers	Definitions / notes / examples
	D.2.6 Environmental sustainability	
	D.2.7 Other (specify)	
	D.2.8 Unclear	
D.3 Other learning and development	D.3.1 Curiosity	
	D.3.2 Relationship with nature	
	D.3.3 Self awareness	
	D.3.4 Self responsibility	
	D.3.5 Communication or teamwork	
	D.3.6 Health & well being	
	D.3.7 Healthy lifestyles	
	D.3.8 Employability	
	D.3.9 Youth leadership	
	D.3.10 Community integration	
	D.3.11 Community leadership	
	D.3.12 Other (specify)	
	D.3.13 Unclear	
D.4 Sport and leisure	D.4.1. Participation in sport (was introduction)	
	D.4.2 Competitive performance	
	D.4.3 Recreational competence	*How is this defined?*
	D.4.4 Group leadership	*How is this different from 'community leadership'?*
	D.4.5 Train the trainer	

Questions	Answers	Definitions / notes / examples
	D.4.6 Family activities	
	D.4.7 Other (specify)	
	D.4.8 Unclear	
D.5 Health	D.5.1 Healthy behaviour	
	D.5.2 Health, physical/ mental	Well being?
	D.5.2 Rehabilitation	
	D.5.2 Other (specify)	
	D.5.3 Unclear	
Section E: Practice/ Discipline/ Activity		
E.1 Formal name	E.1.1 Not applicable (no formal name)	
	E.1.2 Details (specify)	
	E.1.3 Unclear (specify)	
E.2 Dates of operation	E.2.1 Not stated (specify)	
	E.2.3 Details (specify)	
	E.2.4 Unclear (specify)	
E.3 Setting	E.3.1 School grounds	
	E.3.2 Residential facility	
	E.3.3 Local community	
	E.3.4 Other (specify)	
	E.3.5 Unclear	
E.3 Education	E.3.1 Field studies	

Questions	Answers	Definitions / notes / examples
	E.3.2 Expedition(s)	
	E.3.3 Adventurous activity	
	E.3.4 Frequent adventurous activity (e.g. scouts, ramblers)	
	E.3.5 Nature visits	
	E.3.6 Forest/ beach schools	
	E.3.7 Bushcraft	
	E.3.8 Outdoor nursery	
	E.3.9 Other outdoor learner centred	*What exactly does 'learner centred' mean?*
	E.3.6 Unclear	
E.4 Sports and Leisure	E.4.1 Sport training and qualifications	
	E.4.2 NGB coaching programmes	
	E.4.3 Local sports club(s)	
	E.4.4 Mass participation events (e.g. adventure races, OPAL etc	
	E.4.5 Other (specify)	
	E.4.6 Unclear	
E.5 Therapy	E.5.1 Prescribed outdoor activity	
	E.5.2 Group / family adventure based	*Was Group / family adventure based residential*
	E.5.3 One-to-one therapy outdoors	
	E.5.4 Other therapy (specify)	
	E.5.5 Unclear	

Appendix 8: The systematic reviews of the effects of outdoor learning

We found 15 systematic reviews of outdoor learning. Three were excluded as their focus was tangential to effectiveness: surveys of outdoor learning in the UK (Lock 2010); the role of the adventure counselor in facilitating successful interventions in adventure therapy programs for troubled adolescents (Puchbauer 2007); and experiences of Forest School practitioners, in their journeys from training to initial practice (McCree 2014). The remaining 15 systematic reviews of the effects of outdoor learning are listed here. Many of them included studies published well before 2003, even if the systematic review was recent.

First author	Date	Scope
1. Cason	1994	Outdoor adventure programming
2. Hattie	1997	Adventure education and outward bound
3. Rickinson	2004	Outdoor learning
4. Gillis	2008	Challenge ropes courses
5. Neill	2008	Outdoor education
6. Neill	2008	Outdoor education (5 meta-analyses)
7. Coalter	2010	Mountaineering
8. Gill	2011	Spending time in nature
9. Bowen	2011	Adventure Therapy
10. Davies	2013	Creative Learning Environments in Education
11. Higgins	2013	Outdoor adventure education
12. SMCI Assoc.	2013	Wilderness journeys / youth reoffending
13. Stott	2013	Personal development on youth expeditions
14. Cooley	2015	Adventure education for group work in higher education
15. Jill Dando Inst	2015	Wilderness Challenge Programmes

There is considerable overlap between the systematic reviews since several may include the same primary studies, or even other systematic reviews.

---

[i] We used Project Oracle's public materials. The way that we applied them may be different from how Project Oracle uses them.

Appendix 9: Characteristics of systematic reviews

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Bowen 2013	197	Based on statistical meta-analysis	Referred for psychological and/or behavioural therapeutic reasons	Adventure therapy	Academic Behaviour Clinical Family development Morality / spirituality Physical Self-concept Social development
Cason & Gillis (1994)	43	Based on statistical meta-analysis	General population With physical / intellectual disabilities Adjudicated youths / delinquents Emotional issues Inpatients 'at-risk' youths 11 - 14 years 15 - 18 years	Adventurous activity Bushcraft Residential facility Wilderness setting Therapy: group / family adventure based Some 'Outward Bound'	School grades School attendance Self-awareness Behavioural Attitude Locus of control Clinical scales
Coalter et al. (2010)	Not stated	Other study designs	General population ADHD 'at risk' youths Young offenders Age not stated	Adventurous activity Frequent adventurous activity Outdoor / countryside settings Therapy: prescribed outdoor activity (for young offenders) Mountaineering	Relationship with nature Self-awareness Self-responsibility Communication or teamwork Health and well being Healthy lifestyles Community integration Health, physical / mental -weight loss

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Cooley et al. (2015)	11	Experimental designs (not true exp. – 4 used control groups) Other study designs – surveys, reports, anecdotal recall.	General population 18 - 25 years (higher education only)	Outdoor learner centred – structured outdoor activities Residential facility	Undergraduate and Postgraduate degree Communication or teamwork Community integration Community leadership
Davies et al. (2013)	58 – only 4 included outdoor education	Experimental designs (only 2) Other study designs – case studies, surveys	General population Socially excluded young people 5 - 10 years 11 - 14 years 15 - 18 years	Field studies Nature visits Other outdoor learner – developing school grounds School grounds Local community	Motivation Engagement Enthusiasm Enjoyment Concentration Attention Focus ...associated with creativity initiatives
Gill (2011)	61	Experimental designs Other study designs – mostly cross-sectional	General population Mentions children with ADHD 5 - 10 years 11 - 14 years 15 - 18 years 18 - 25 years	Field studies Nature visits Other outdoor learner centred – conservation, gardening, play School grounds Local community Woodlands	Educational benefits mentioned but not categorised Relationship with nature Self-awareness Communication or teamwork Health and well-being Healthy lifestyles Healthy behaviour Health, physical / mental – motor fitness

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Gillis & Speelman (2008)	44	Based on statistical meta-analysis	General population With additional special needs (not stated) 11 - 14 years 15 - 18 years 18 - 25 years	Adventurous activity Therapy: group / family adventure based Challenge (ropes) courses	Self-awareness Self-responsibility Communication or teamwork Health and well-being Community integration Health, physical / mental
Hattie et al. (1997)	96	Based on statistical meta-analysis and other review findings	General population Delinquents Low achievers Managers / managements 5 - 10 years 11 - 14 years 15 - 18 years 18 - 25 years	Adventurous activity Bushcraft Wilderness settings Some 'Outward Bound'	Self-awareness Self-responsibility Communication or teamwork Health and well-being Youth leadership Community integration Community leadership Adventuresomeness Health, physical / mental
Higgins et al. (2013)	Four for adventure education	Systematic review of four meta-analyses	General population 11 - 14 years 15 - 18 years	Adventurous activity School grounds Residential facility	Health and well-being – self esteem
Jill Dando Institute (2015)	28 effects studies 23 implement-ation studies	Critical appraisal of statistical meta-analysis	Young offenders	Wilderness challenge programmes, either in isolation or with other therapeutic enhancements	Interpersonal skills (self-esteem, social skills, self-control, school adjustment) Offending Self-reported delinquent behaviour

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Neill (2008a)	6	Systematic review of systematic reviews / meta-analyses	General population Other special needs Unclear 5 - 10 years 11 - 14 years 15 - 18 years	Adventurous activity Residential facility Local community Other Outward Bound	Grade Point Average Self-awareness Self-responsibility Communication or teamwork Community integration
Neill (2008b)	5	Systematic review of five meta-analyses	General population 11 - 14 years 15 - 18 years 18 - 25 years	Adventurous activity Bushcraft Residential facility Wilderness setting Therapy: Group / family adventure based Some 'Outward Bound'	Grade Point Average Relationship with nature Self awareness Communication or teamwork Health and well-being Youth leadership Recidivism
Puchbauer (2007)	Could not access full report				

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Rickinson et al. (2004)	150	Some meta analyses included Experimental designs Other study designs	General population Emotional and behavioural difficulties Young offenders 5 - 10 years 11 - 14 years 15 - 18 years 18 - 25 years	Field studies Adventurous activity Nature visits Bushcraft School grounds Local community Wilderness settings Rural areas Therapy: Group / family adventure based	Curiosity Relationship with nature Self-awareness Self-responsibility Communication or teamwork Health and well-being Healthy lifestyles Youth leadership Community integration Healthy behaviour Health, physical / mental – reduction in anxiety
SMCI Associates (2013)	Unclear	Unclear	Young offenders 'at-risk' youth Disadvantaged youths 11 - 14 years 15 - 18 years 18 - 25 years	Adventurous activity Bushcraft Wilderness setting Therapy: therapeutic aspect around reducing re-offending	Self awareness Self responsibility Communication or teamwork Healthy lifestyles Employability Community integration Ethical and moral developments Recidivism

Lead Author	Number of included studies	Reliability of conclusions	Population characteristics	Practice / discipline / Activity	Outcomes
Stott et al. (2013)	35	Experimental designs Other study designs – observations, surveys, descriptive narratives	General population 11 - 14 years 15 - 18 years 18 - 25 years	Expedition Overseas Some 'Raleigh' expeditions	Relationship with nature Self awareness Self responsibility Communication or teamwork Community integration Community leadership

Appendix 10: Project Oracle standards of evidence in detail

To be validated as attaining the various levels, evaluations / organisations must meet the following criteria.[12] The third column indicates the number of organisations which have to date (September 2, 2015) been deemed to have achieved each level.[13] (The total number of organisations which have applied to be validated, and hence the number which failed to reach Level 1, isn't clear):

Level 1	Key requirements are: • You have developed a Theory of Change for your project. • You have developed an outline evaluation plan of when and how you will measure the impact of your project.	186
Level 2	The main part of this validation is undertaking an evaluation and writing a report that meets the following criteria: *Evaluation design*: Your evaluation measures changes in the outcomes in an appropriate way. This can include qualitative and quantitative methods. Control and comparison groups are not a requirement. The methods you use must: • Include pre and post analysis. • Use valid and reliable measurement tools which are appropriate for the participants. *Evaluation report content*: Your evaluation report must contain details such as description of how participants were selected and their consent obtained, how measurement tools were used (e.g., questionnaires used, how any survey was distributed, details of any statistical analyses). Also a review / critique of the limits of the methods.	24
Level 3	There has been at least one rigorous evaluation using a comparison group or other appropriate comparison data, ideally with long term follow up. Exceptions to this apply in cases where it is not possible, or extremely difficult, to set up suitable control groups or use appropriate comparison data, or where long term follow up is not feasible or appropriate. In these cases, the following aspects of the evaluation will form part of the validation: • The strength of the theoretical model underpinning the intervention. • The quality of the data used to assess impact.	4

---

[12] Project Oracle. *Standard One.* [Online] http://project-oracle.com/support/for-youth-service-providers/validation-against-the-standards/standard-one/ [Accessed: 22.10.15].

[13] Project Oracle. *A-Z of our projects*. [Online] http://project-oracle.com/projects/standards-of-evidence/ [Accessed: 02.09.15].

Level 4	You have undertaken two or more rigorous impact evaluations of the project, including at least one undertaken by an external evaluator. At least one of the evaluations should include comparison group or other appropriate comparison data, and the evaluations should provide a rounded picture of the impact of the project. This could include using different methods to understand your impact on certain outcomes, looking at the project's effects on different outcomes, or over different time frames.	0
	Within these evaluations you can provide:	
	- Evidence to support the causal mechanism: how does your project lead to changes in the outcomes? - Evidence on dosage: does doing more or less of your project, or parts of it, have better or worse results? - Analysis of the impact of your project on sub-groups in your target population: for example, do the results hold up for different age groups, boys and girls, ethnic minority groups? - Evidence that your project continues to be effective when replicated to other settings. - Evidence that the project is consistently delivered as planned, and is reaching the target groups.	
	You have also undertaken a cost benefit analysis, using methods that meet internationally recognised standards.	
Level 5	You have in place systems and documentation to support large-scale implementation, and you are able to transfer the running of the intervention to other agencies. These systems enable quality to be maintained and ensure that strong results are consistently delivered.	0

Appendix 11: Report on included systematic reviews

Sixteen systematic reviews were found to fit the criteria for inclusion in this review. The full text could not be accessed for one systematic review (Puchbauer 2007), and a meta-analysis of programmes primarily using adventure-based activities for psychological and/or behavioural therapeutic purposes (Bowen and Neill 2013) was found when searching for a subsequent systematic review. Although the scope was added to the table of studies in the main report, there was insufficient time to summarise the findings here. The findings from the remaining fourteen systematic reviews are reported here.

Adventurous activity

Five reviews (Cason & Gillis (1994); Hattie et al. (1997); Neill (2008a); Rickinson et al. (2004); SMCI Associates (2013)) included studies focusing on adventurous activity and bushcraft, which involves participants setting up their own overnight accommodation in a wilderness setting. Two reviews, Higgins et al. (2013) and Neill (2008b), included studies with adventurous activities but rather than the bushcraft experience these participants stayed overnight in a residential facility. An additional two reviews (Coalter et al. (2010) and Gillis & Speelman (2008)) included adventurous activities without an overnight element.

All, except SMCI Associates (2013) included participants from the general population. In addition, four reviews (Cason & Gillis (1994); Hattie et al. (1997); Rickinson et al. (2004); SMCI Associates (2013)) included young offenders, 'delinquent' or 'at-risk' youths; with Cason & Gillis (1994), Coalter et al. (2010) and Rickinson et al. (2004) also including some participants with physical / intellectual disabilities or with emotional and behavioural difficulties. Gillis & Speelman (2008) and Neill (2008a) also included participants with additional special needs but did not specify further.

Adventurous activity and bushcraft

Cason & Gillis (1994) conducted a meta-analysis including participants from the general population and other specific populations: adjudicated, delinquent or 'at-risk' youths; participants with physical or intellectual disabilities; inpatients and adolescents with emotional difficulties. All participants were between 11 years old and college freshman age. Forty-three studies were included in the meta-analysis, resulting in 147 effect sizes.

Effect sizes in these studies ranged from -1.48 to 4.26, with an average effect size of 0.31 and standard deviation of 0.62. This finding represented a 12.2% improvement for the average adolescent participating in the adventure programming, indicating they were 62.2% better off than those who did not participate.

The summary effect sizes of outcome measurement categories (e.g., self-concept, locus of control, clinical scales) were significantly different from each other and ranged from 0.30 to 1.05. Larger effect sizes were linked with longer programmes, younger participants and published studies, which produced significantly higher effect sizes than unpublished dissertations. More rigorous study designs were linked with lower effect sizes. This meta-analysis showed adventure programming to be equally effective with adjudicated adolescents as with other adolescent populations.

Hattie et al. (1997) conducted a meta-analysis and non-statistical review including participants ranging from school aged up to adults who were from the general population, were delinquents, low achievers or in management job positions. Ninety-six studies were included, resulting in 151 unique samples forming 1728 effect sizes.

The outcome effects were collated in six categories: leadership; self-concept; academic; personality; interpersonal and adventuresome. The average effect of attending an adventure programme was .34, with a follow up effect of .17. The outcome effects for the six categories are as follows: leadership (.38, follow up .15); self-concept (.28, follow up .23); academic (.46, follow up .21); personality (.37, follow up .14); interpersonal (.32, follow up .17) and adventuresome (.38, follow up -.06). A theme underlying the outcomes with the greatest effects relate to self-control: independence (.47); confidence (.33); self-efficacy (.31); self-understanding (.34); assertiveness (.42); internal locus of control (.30) and decision making (.47). Most of these effects are maintained over time, thus adventure programmes appear to be effective at providing participants with a sense of self-regulation. The three individual variables that explained most variance between adventure programmes were: age (adult or student), length of programme (longer >20 days or shorter <20 days) and whether the adventure programme was Australian Outward Bound or not. The most effective programmes were longer, adult, Australian programmes, whereas longer, adult, non-Australian programmes were the least effective. The mean effect size for all programmes with school-aged students and for all shorter programmes was 0.26. Hattie et al. (1997) noted that there is variance between all adventure programmes, with some proving to be more effective and some less so.

Rickinson et al. (2004) conducted a review including participants of all ages from the general population, young offenders and those with emotional and behavioural difficulties. One hundred and fifty studies were included in the review, including a number of meta-analyses.

Rickinson et al. (2004) concluded that outdoor adventure activities had both short term and continuous positive effects. Although considerable variation between different programmes and outcomes were noted, there was evidence of positive impacts on attitudes, beliefs, interpersonal and social skills. Academic skills, positive behaviour, re-offending rates and self-image were also shown to have been positively impacted. However, a strong positive link between outdoor adventure activities and environmental understanding was not evidenced.

Rickinson et al. (2004) also investigated the impacts of fieldwork and school ground or community projects (which are reported here later) along with the factors Rickinson et al. (2004) found to influence outdoor learning and its provision.

Neill (2008a) conducted a review of six traditional and meta-analytic reviews. Participants between 5 - 18 years old were included from the general population and with some additional special needs but this classification was not further specified.

The traditional reviews reported some positive outcomes, concluding in cautious but positive views about the personal and social developmental effects of outdoor adventure programmes. The meta-analytic studies supported this indicating a small-moderate effect size of 0.35 for short term effects. There was some limited but promising evidence for long term effects. These findings indicated that 64% of those who had

participated in adventure programmes were better off than non-participants. Neill (2008a) noted a considerable variation in the study results, partly explained by programme type and length and participant age. Australian Outward Bound longer programmes with adult participants were most effective. It was noted many adolescent programmes are compulsory whilst adult programmes tend to be voluntary, which may have been confounding.

The **SMCI Associates (2013)** conducted a quasi-systematic review including young offenders, 'at-risk' and disadvantaged youths of all ages. The number of included studies and their design types are unclear from the review report.

The outdoor and wilderness adventure programmes included in the review were shown to have positive outcomes on young offenders and 'at-risk' youths' recidivism rates and personal, social and employability development and skills. The nature of the relationship between participants and the adult staff was found to be a key factor in the efficacy of outdoor adventure programmes; wilderness programmes were seen as providing important opportunities for participants to develop new positive relationships with adults. The literature suggested a pro-social and assets-based approach may be most successful in reducing reoffending behaviour in young people, though the review noted that it may be challenging for justice authorities to move from seeing young offenders as liabilities to assets. The SMCI Associates (2013) concluded that although empirical evidence on the impacts of wilderness journeys is limited, substantial anecdotal evidence indicates the positive impacts of such programmes.

Adventurous activity in residential settings

Neill (2008b) conducted a meta-analysis which included participants from the general population. Although the specific age ranges included were not stated, the review mentions younger adolescents, older adolescents and adults. The conclusions drawn in this review relied on five included meta-analyses.

Neil (2008b) concluded that outdoor education programmes have a small-moderate impact for typically measured outcomes such as self-esteem, behaviour problems, and teamwork. One of the included meta-analysis, Hattie et al. (1997) suggested that 65% of participants were better off for having participated in outdoor education programmes. On average, the outdoor education participants experienced additional growth on returning to their home environments, though the generalisability of this finding was limited. The main influences on empirical outdoor education research outcomes were the outdoor education organisation running the programme, the age of participants, and the length of the programme. Other moderators of note were the quality of study and whether the programme was residential. Overall, it appears that the results of outdoor education programme research show that there are small-moderate average effects which vary considerably from participant to participant and from programme to programme. Neil's (2008b) findings would suggest that more therapeutically / development focused programmes, for young adolescents or adults, for longer periods of time (and residential stays) produce larger measures of change.

Higgins et al. (2013) also conducted a systematic review of meta-analyses including participants from the general population. The age of the participants was unclear but some reference was made to adolescents. Higgins et al.'s (2013) conclusions about adventure learning were taken from four included studies; though

Overall, studies of adventure learning interventions within this review consistently showed positive benefits on academic learning, and wider outcomes such as self-confidence. On average, pupils who had participated in adventure learning interventions appeared to make approximately three additional months' progress. The evidence suggested that the impact was greater for longer courses (more than a week) and those taking place in 'wilderness' settings, though other types of interventions still showed positive impacts. The indicative effect size was 0.23, with the included meta-analyses' effect sizes varying from 0.17 to 0.61. Higgins et al. (2013) noted that understanding why this is the case is underdeveloped and that the more recent available evidence is more robust. Nevertheless, all evidence showed positive effects.

Adventurous activity without overnight stay

Gillis & Speelman (2008) investigated the impact of challenge (ropes) course activities from 44 meta-analyses. Participants ranged in age from 11 - 25 years old and were included from the general population and populations with some special needs, although these were not further classified.

An overall effect size of 0.43 was calculated, with the highest effect size for outcomes calculated from studies based on family measures (0.86); a large effect size with practical significance. Medium effect sizes for outcomes with educational significance were reported for self-efficacy (0.48), behavioural observations (0.37), personality measures (0.29), self-esteem or self-concept (0.26) and academic measures (0.26). Outcome measures related to classroom environment were small (0.01).

Studies with therapeutic (0.53) or developmental foci (0.47) had higher effect sizes than those with educational foci (0.17). The highest effect sizes occurred in studies conducted in therapeutic settings, which Gillis & Speelman (2008) hypothesised may be due to the nature of the populations studied in these settings and their assessments.

Coalter et al. (2010) investigated the impact of mountaineering and other mountaineering related activities. It is unclear from the review how many studies were included or the age range of the participants. The sample of participants included members of the general population, young offenders, 'at risk' youths and participants with ADHD.

Coalter et al. (2010) found limited research on the economic and social impacts of mountaineering activities. Drawing on broader literature they concluded that mountaineering had a positive impact on physical health with the ability to impact on cardiovascular, musculoskeletal, metabolic, endocrine and immune systems. The research also indicated possible negative physiological impacts of mountaineering. The review concluded that although much of the evidence about the benefits of mountaineering on physical health were not fully evidenced it is still appropriate to advocate these activities for physical health reasons.

Direct evidence of the mental health and psychological benefits of mountaineering activities is limited too. Coalter et al. (2010) concluded it was reasonable to assume such activities could improve mental health and psychological wellbeing, such as through their outdoor aspects and opportunities to set and achieve goals.

Three reviews (Davies et al. (2013); Gill (2011); Rickinson et al. (2004)) included studies focusing on field studies, nature visits or outdoor activities related to the school grounds, such as conservation projects or gardening. Each of these reviews included participants from the general population with the addition of a few specific populations in some reviews: socially excluded young people (Davies et al. (2013)), children with ADHD (Gill (2011)), young offenders and those with emotional and behavioural difficulties (Rickinson et al. (2004)).

Rickinson et al. (2004) found evidence that fieldwork had positive impacts on long term memory, individual growth and social skills. The review concluded that these positive impacts were only possible with well planned, taught and followed up fieldwork. When this was the case, students could develop knowledge and skills to add to their classroom learning. Rickinson et al. (2004) noted the evidence showed a severely restricted amount of fieldwork in the UK, particularly in science.

School grounds and community projects were also found to have a positive impact with evidence of academic, social and personal outcomes. Academically, students showed positive gains in science process skills and design and technology issues. There was evidence of social development, greater community involvement and the development of more positive relationships. Greater confidence, renewed pride in the community, stronger motivation toward learning and greater sense of belonging and responsibility were also evidenced.

Rickinson et al. (2004) also investigated factors influencing outdoor learning and its provision. Five main barriers were identified: fear about health and safety; teachers' lack of confidence; curriculum requirements; lack of time, resources and support; and wider changes within and beyond the education sector. A number of opportunities were identified too, including new legislation and regulations, recent curriculum developments and initiatives and developments in UK higher education.

Programme, participant and place factors were found to facilitate or impede learning. The research indicated that longer programmes, with well-designed preparatory and follow-up work with a range of curriculum linked activities and assessments were the most valuable. Recognising and emphasising the role of facilitation in the learning process and developing close links between programme aims and practices was also concluded to be important.

Gill (2011) conducted a systematic review including 61 studies covering participants of all ages. All studies were assessed for quality, resulting in ten studies being excluded from analysis having been rated as 'poor' quality. The included studies were made up of experimental designs and other study designs, mostly cross-sectional.

Based on the quality assessments the conclusions of this review were separated in to claims that were well supported, claims that had some good support and claims with some support. The findings for each of these categories are presented below.

Claims that are well supported: Spending time in natural environments as a child was associated with adult pro-environment attitudes, feelings of being connected with the natural world and a stronger sense of place. Living near green spaces was associated with greater physical activity. Spending time in nearby nature lead to improvements in mental health and emotional regulation, both for specific groups of children (such as those with ADHD) and for children as a whole. Children who took part in school gardening projects improved in scientific learning more than those who did not, and had healthier eating habits. Experience of green environments was associated with greater environmental knowledge. Play in natural environments lead to improvements in motor fitness for pre-school children.

Claims that have some good support: Forest school and school gardening projects were associated with improved social skills. In addition, forest school lead to improved self-control and school gardening projects lead to increased self-awareness.

Claims with some support: Nearby nature was associated with more outdoor play and hence improved well-being. Forest schools were associated with improved self-confidence and language and communication. Conservation activities in school grounds and nearby open spaces were associated with improved psycho-social health.

Davies et al.'s (2013) review was focussed around creativity. Fifty-eight studies of school aged participants were used, mostly of case study designs but only four covered outdoor education, including Forest Schools.

There was reasonable evidence across the studies that taking pupils out of the classroom and working in an outdoor environment for part of their time in school can foster their creative development. The reasons for this may be connected with ownership and collaboration. There was reasonable evidence for increased pupil motivation, engagement, enthusiasm, enjoyment, concentration, attention and focus associated with creativity initiatives. Additionally there was reasonable evidence that creative learning environments can aid children and young people's emotional development and social skills.

Structured outdoor activities in residential facilities

One review, **Cooley et al. (2015)**, focussed on outdoor structured activities based at residential facilities. The eleven included studies were for participants enrolled in higher education courses only and the review's main focus was around group work outcomes. Four studies utilised control groups but were not true experiments; the remaining studies were made up of surveys, reports and anecdotal recall.

There was evidence that transferable group work skills were developed during outdoor adventure education and retained when students returned to higher education. Robust evidence was lacking, however, to show the extent to which students were able to apply these skills in different contexts. Studies in the review demonstrated that teambuilding occurs during outdoor adventure education. Although there was some evidence that groups returned to higher education displaying a more positive group environment and more effective group processes, there were mixed findings on whether this led to improved group performance. Students developed more positive attitudes towards group work in terms of seeing the benefits and feeling more confident in engaging in group work; although it was not clear how this change in attitude may have influenced students' approach to new group work situations. Lastly, there were claims of increased

integration and feelings of social support within peer groups. It should be noted however, the included studies contained a range of limitations such as non-validated questionnaires, weak study designs and analytical procedures, and a lack of focus on long-term behaviour change.

Expeditions overseas

The final review, **Stott et al. (2013)**, focussed on studies of overseas expeditions of 14 days or more, which must have been self-propelled and overseas (or out of state for Australia and North America). Participants were from the general population. No total specific age range was stated but the report mentioned two included studies had included participants aged 14 - 18 years old and 20 - 23 years old. Thirty-five studies were included, mostly of non-experimental design, gathering data from observations, interviews, descriptions and self-reports written by expedition members.

The findings were presented in four themes: upward personal growth (realising potential); outward personal growth (learning about others); inward personal growth (learning about self) and downward personal growth (learning about environment). The following were found to be associated with overseas youth expedition participation:

- Upward personal growth (realising potential): increased confidence; physical and social resilience; self-reliance and ability to overcome challenges.

- Outward personal growth (learning about others): improved social skills.

- Inward personal growth (learning about self): improved emotional stability; better able to reflect on events.

- Downward personal growth (learning about environment): increased environmental appreciation and awareness.

Stott et al. (2013) also noted that processes that were valued by participants in overseas expeditions include genuine independence; group isolation and self-sufficiency; person-centred leadership; positive responses to stress and physically demanding activity.

Appendix 12: RCTs can be cheap, easy and quick

From Third Sector magazine, June 2015, by Caroline Fiennes

June brings the Queen's birthday, and perhaps this time, you – like many charity sector people before you – will get lucky and be in the Birthday Honours List.

If so, then arise, Lady Reader, for I have an important task for you.

This auspicious occasion creates an opportunity, which is nothing less than discovering whether Her Majesty's gongs actually make any difference. We currently don't know that, despite all the sound and fury about them.

When Iain Chalmers was knighted for his role in creating The Cochrane Collaboration, the leading source of rigorous evidence in health-care, he wondered whether it would make people pay him more attention. So he asked a colleague – Mike Clarke – to do a little randomised trial. Iain's outgoing letters were, by random allocation, signed either 'Iain Chalmers' or 'Sir Iain Chalmers': Mike monitored response rates and response times. A clue to the answer is that the resulting paper is called 'Yes Sir, no Sir, not much difference Sir'.

This little story says a lot about evaluations and evidence. First, this trial was free. Randomised trials have a weird reputation for vast expense. That's garbage: there's nothing inherently expensive about having a control group, nor populating it at random. This trial used data that were cheap to collect or being collected anyway. So do most of the new-ish breed of low-cost public sector randomised trials, such as those by the government's Nudge Unit to assess how to get people to pay tax on time: HMRC already checks whether you pay tax.

Second, the trial was quick. Randomised trials are reputed to invariably take ages, which is also untrue. They take as long as it takes for the outcome of interest to appear. No sensible evaluation could do otherwise. If you're assessing whether learning a language at school affects a person's lifetime earnings, then you can't avoid a long wait. However if you're interested in whether (to cite a real study that we've seen before) including information about your charity's results in your fundraising solicitations increases donations within two months, then you only have to wait two months.

Low-cost and rapid randomised trials could often assess charities' work too. By not doing them, we're missing fantastic opportunities to find out what works best.

And lastly, maybe the 'Not much difference, Sir' part was because Iain was already so well known that people replied to him promptly anyway. What's true for the mighty Iain Chalmers may not be true for you: this experiment's results may not be 'externally valid', in the jargon. A sample of one doesn't tell us much, and hence science is fundamentally about repeatable results: nobody's impressed if you achieved cold fusion in your bathroom last Tuesday but can't do it again. To my knowledge, nobody else has scientifically studied the effect of a gong.

This is where you come in, Lady Reader. You can help to find out. Look up this study, run it on your own correspondence, report the findings. I'll happily compile the answers - and tell the Palace.