Appendices

Appendix A. Example search strategy – ERIC

S1	DE "Preschool Children"
S2	TI child* OR AB child*
S3	TI (boy* OR girl*) or AB (boy* OR girl*)
S4	TI toddler OR AB toddler
S5	TI young N1 child* OR AB young N1 child*
S6	TI early N1 child* OR AB early N1 child*
S7	TI early N1 year* OR AB early N1 year*
S8	TI “pre-primary” or AB “pre-primary”
S9	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8
S10	DE "Nursery Schools" OR DE "Preschool Education" OR DE "Outdoor Education" OR DE "Adventure Education"
S11	TI nurser* OR AB nurser*
S12	DE "learning" OR TI early N1 learning OR AB early N1 learning
S13	TI (“preschool” or “pre-school”) OR AB (“preschool” or “pre-school”)
S14	TI kindergarten OR AB kindergarten
S15	TI (childcare OR child N1 care) OR AB (childcare OR child N1 care)
S16	TI (daycare OR day N1 care) OR AB (daycare OR day N1 care)
S17	TI education OR AB education
S18	DE "Play" OR TI (Play OR “play-based learning”) OR AB (Play OR “play-based learning”)
S19	TX (Waldkindergartens OR udeskole OR friluftsliv OR peuterspeelzaal OR kinderopvang OR bush N1 kinder*) OR TI (forest N1 kindergarten* OR forest N1 school*) OR AB (forest N1 kindergarten* OR forest N1 school*)
S20	S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19
S21	TI outdoor* OR AB outdoor*
S22	TI (nature OR “nature-based”) OR AB (“nature-based”)
S23	TI environment* OR TI outdoor N1 environment* OR AB outdoor N1 environment*
S24	TI (forest* OR wood* OR park* OR recreation* OR landscape* OR tree* OR hill* OR garden* OR beach* OR eco)
S25	AB (forest* OR wood* OR park* OR recreation* OR landscape* OR tree* OR hill* OR garden* OR beach* OR eco)
S26	TI (green OR greenspace or green N1 space) OR AB (green OR greenspace or green N1 space)
S27	TI (loose N1 parts OR “loose-parts”) OR AB (loose N1 parts OR “loose-parts”)
S28	TI (adventure* OR wild OR “open-air”) OR AB (adventure* OR wild OR “open-air”)
S29	S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28
S30	S9 AND S20 AND S29

Appendix B. Modified quality appraisal tools
EPHPP Quality Assessment Tool

Modifications in bold red

A) SELECTION BIAS
(Q1) Are the individuals selected to participate in the study likely to be representative of the target population? **(i.e. children aged 2-7 years not in formal education yet)**

1. Very likely
2. Somewhat likely
3. Not likely
4. Can't tell

(Q2) What percentage of selected individuals **consented to the research?**

1. 80 - 100% agreement
2. 60 – 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN
Indicate the study design:

1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify _____
8. Can't tell

Was the study described as randomized? If NO, go to Component C.
No **Yes**

If Yes, was the method of randomization described? (See dictionary)
No **Yes**

If Yes, was the method appropriate? (See dictionary)
No **Yes**

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS
(Q1) Were there important differences between groups prior to the intervention?

1. Yes
2. No
3. Can't tell

The following are examples of confounders:

1. **Gender**
2. **Age**
3. **Socio economic status (SES – e.g. Parental education, deprivation status)**

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

1. **All confounders**

2. **Two confounders**
3. **One confounder**
4. Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING
(Q1) Was (were) the outcome assessor(s) **and/or analysists** aware of the intervention or exposure status of participants?
1. Yes
2. No
3. Can't tell

(Q2) **Were outcome assessors aware of the research question?**
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS
(Q1) Were data collection tools shown to be valid?
1. Yes
2. No
3. Can't tell

(Q2) Were data collection tools shown to be reliable?
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS
(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1. Yes
2. No
3. Can't tell
4. Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).
1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK

See dictionary	1	2	3

COMPONENT RATINGS
Please transcribe the information from the grey boxes on pages 1-3 onto this page. See dictionary on how to rate this section.

A	**SELECTION BIAS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
B	**STUDY DESIGN**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
C	**CONFOUNDERS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
D	**BLINDING**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
E	**DATA COLLECTION METHOD**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3
F	**WITHDRAWALS AND DROPOUTS**	**STRONG**	**MODERATE**	**WEAK**
		1	2	3

Overall Grade (based on above six criteria):

Scored 1 for study design (i.e. controlled studies); AND **Scored 1 or 2 in at least three other important components, including:** **selection bias** **confounders** **blinding** **withdrawals and drop-outs.**	**STRONG** **1**
Scored 1 for study design; AND **Scored 1 or 2 in two other important components, including:** **selection bias** **confounders** **blinding** **withdrawals and drop-outs.**	**MODERATE** **2**

OR **Scored 2 for study design; AND** **Scored 1 or 2 in at least three other important components, including:** **selection bias** **confounders** **blinding** **withdrawals and drop-outs.**	
Scored 1 for study design; AND **Scored 3 in more than <u>two</u> other important components, including:** **selection bias** **confounders** **blinding** **withdrawals and drop-outs.** **OR** **Scored 2 for study design; AND** **Scored 3 in more than <u>one</u> other important components, including:** **selection bias** **confounders** **blinding** **withdrawals and drop-outs.** **OR** **Scored 3 for study design**	**WEAK** **3**

Dixon-Woods (2004) checklist

Question 1	Are the research questions clear?
Question 2	Are the research questions suited to qualitative inquiry
Question 3	Are the following clearly described? - Sampling - Data collection - Analysis
Question 4	Are the following appropriate to the research question? - Sampling - Data collection - Analysis
Question 5	Are the claims made supported by sufficient evidence?
Question 6	Are the data, interpretations, and conclusions clearly integrated?
Question 7	Does the paper make a useful contribution **to the review question**?

Appendix C. Characteristics of included studies

Table 1. Characteristics of included quantitative studies						
Author, year and country	**Study design**	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Follow-up time point**	**Outcome(s)**	**Data analysis**
Nature-based ELC						
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled before & after	E: Age: 47.2 months ± 6.52 Gender: 13m/28f C: Age: 46.75 months ± 6.95 Gender: 29m/23f SES not reported.	E: Teachers underwent special training in outdoor education over one year including (15 days). ELC consisted of a green park with some centuries-old trees (e.g., firs, willows, maples), plants and flowers, and without any play structures. C: ELC contained grass and cement without larger plants, trees, and play structures.	T1= Jan 2014 T2= May 2014 T3= Oct 2014 T4= May 2015 (16 months from baselines)	Motor skills Cognitive Social and Emotional Nature Connectedness Play	Mixed-Model Repeated Measures analysis of variance (ANOVA)
Cooper (2018), United Kingdom (England). E: 13 children C: 11 children Children from the same school	Controlled before & after	E: Age: 47 months (range 45-48) Gender: 7m/4f C: Age: 44 months (range 41-47) Gender: 7m/4f SES was noted as being “generally above average" for both groups.	E: Forest school sessions run by two trained leaders which operate for 10 week cycles on Tuesday AM and PM (2 hours each). Children attend either the AM or PM session. The forest school consists of trees and vegetation, a seating area made from logs, a mud kitchen using old crates and a tyre, a greenhouse and pond. The forest school is located on site and when children do not have forest school sessions outdoors, they have a “ free flow” environment where children are allowed outside when they want. C: Usual nursery practice which also involves a large amount of outdoor exploration. Children also participated	10- weeks	Cognitive Social and Emotional	Wilcoxon Signed-Rank Test; Mann-Whitney U test.

			in a one hour per week nature play session which incorporated elements of the forest school and included gardening, litter picking and PA. Staff have created an engaging multi-sensory outdoor environment for children which includes a sand pit area, water features and climbing apparatus. The nursery has an allotment system for children to plant fruit trees.			
Cordiano et al (2019), USA. E: 12 children / 1 ELC class. C: 14 children / 1 class. Children from the same school.	Controlled before & after study	Age: 51.5 months (4.3 years) Gender: 26f SES: 46% of students attending the ELC receive financial assistance	E: Outdoor pre-primary programme involved children spending five mornings per week at the school's outdoor campus. The children were outdoors in the forest for 90% of the school day. C: Traditional prekindergarten programme involved children spending five mornings per week at the school's main campus. This involves an Eco!Wonder curriculum that teaches all children about nature and sustainability. Children also visited the outdoor campus one morning per week and spent one immersion week at the outdoor campus in the spring. The remainder of their outdoor time was spent in built environments.	8 months	Cognitive Social and Emotional Play	Mixed-model analysis of covariance (ANCOVA) Covariates: age
Choi et al (2014), South Korea. E: 18 children / 1 ELC C: 19 children / ELC	Controlled before & after	E: Age: 4.2 ± 1.1 Gender: 11m/7f SES: all middle class C: Age: 4.0 ± 1.1	E: Children attend forest kindergarten 5 days per week, year-round, regardless of weather conditions. Children are outdoors more than 80% of the day and usually play, walk, run, and observe various things in the forest. C: Regular kindergarten (not described)	8 months	Sleep	Wilcoxon signed rank test.

		Sex: 11m/8f SES: all middle class				
Elliot et al (2014), Canada. E: 21 children / 1 ELC C: 22 children / 2 ELC	Controlled before & after (mixed-methods)	E: Age: 5.3 years (0.5 SD) Gender: 10m/11f SES not reported. C: Age: 5.3 years (0.3 SD) Sex: 7m/15f SES not reported.	E: A two-year pilot project in which 22 students would spend the mornings from 9:00 to 11:45 outside their school, exploring their local natural environment. C: not described	6 months	Nature Connectedness	ANOVA
Ene-Voiculescu & Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001), Norway. E: = 46 children / 1 kindergarten C: 29 children, / 2 kindergartens	Controlled before & after	Age: 6.1 years Gender: 38m/37f SES not reported.	E: Children used the forest every day for 1-2 hours throughout the year when they attended kindergarten. Occasionally they used the outdoor playground inside the kindergarten fence. The small forest (7.7 hectares) consisted of mixed woodland vegetation, some open spaces of rocks and open fields and meadows in between. C: Children used the traditional outdoor playground for 1-2 hours a day and visited natural sites only occasionally.	10 months	Motor skills	T-test.
Ernst & Burcak (2019), USA. E: 34 children / 2 ELC C: 43 children / 2 ELC	Controlled before & after	E: Age: 4 years Sex: 50%m/ 50%f C Age: 4 years	E: The nature-preschools utilised a combination of wild natural settings spaces that were minimally managed and natural playscapes designed specifically for nature play. The majority of time spent was in free play outdoors in unmaintained or minimally	9 months	Cognitive (all 5 papers) Social and emotional (Ernst & Burcak, 2019l Ernst et al., 2019)	GLM Covariates: pre-test scores, age, gender, prior participation

Burgess & Ernst (2020) E: 84 children / 4 ELC C: 24 children / 2 ELC Zamzow & Ernst (2020) E: 78 / 4 ELC C: 44 children / 2 ELC Ernst et al (2019) E: 78 children / 4 ELC Wojciehowski & Ernst (2018) E: 75 children / 4 ELC	Uncontrolled before & after	Sex:64%m/ 36%f SES not reported	maintained natural settings regardless of weather conditions (approximately four to five hours per day). C: Non-nature preschools emphasised child-directed play. The majority of time was spent indoors in free or loosely guided play (four to five hours), with about one hour daily of teacher-led playful learning. Children at both groups had one to two hours of daily outdoor playtime (weather permitting) in a maintained outdoor space that contained playground equipment.		Play (Burgess & Ernst, 2020)	t-test
Müller et al (2017), Canada. E: 43 children / 1 ELC C: 45 children / 1 ELC	Controlled before & after	Age: E: 63.56 months (3.33 SD) C: 64 months (3.56 SD) Gender not reported. SES not reported.	E: “nature kindergarten” C: “traditional kindergarten” Neither are described.	9 months Sep/Oct-May	PA Motor skills Cognitive Social and Emotional Nature Connectedness	Analyses of Covariance (ANCOVA)
Nazaruk & Klim-Klimaszewska (2017), Poland. E: 90 children (50 urban / 40 rural)	Uncontrolled before & after	Age: 6 years Gender not reported. SES not reported.	Teachers arranged trips in the forest, the park, the allotment garden, the meadow, the agritourism farm, animals at the zoo.	6 months	Nature connectedness	Mann-Whitney U test; Pearson Chi test

Yilmaz et al (2020), Turkey. 40 children / 1 ELC	Uncontrolled before & after	Age: 72 months (6 years) Gender: 14m/26 SES not reported.	E: Children visited a natural, unstructured area for one day in a week for four consecutive weeks. The education programme consisted of 12 semi-structured activities (3 per week). In addition, children also had 30 minutes walk near a natural pond when they visit the setting each week and each week, children had 30 minutes unstructured free play time to discover the natural environment.	4 weeks (1 session per week - 1 full day) conducted in spring 2018	Nature connectedness	Paired sample t-test; ANOVA
Barrable et al (2020), UK (England, Scotland, Wales). E: 141 /12 ELC C: 110 children / 6 ELC	Controlled cross-sectional	Age: 4.53 ± 1.39 Gender: 127m/89f SES not reported.	E: ELC's that have a continuous outdoor provision, with no permanent indoor access and children are outdoors for the whole duration of the ELC day. C: ELC's that are predominately indoor and have variable outdoor provision.	N/A	Nature connectedness	GLM with a binomial error distribution Covariates: Parental NC scores, sex, exposure
Frenkel et al (2019), USA. E: 71 children / 5 ELC C: 70 children / 4 ELC	Controlled cross-sectional	Age: 4.3% = 2 years, 29.1% = 3 years, 50.4% = 4 years, 16.3% = 5 years Gender: 82m/59f SES: 103, 036 USD (median zip code Income)	E: All nature ELC sites were located in parks with distinct areas marked off with rocks and other natural features for daily activities. Children were encouraged to play in the natural environment, which included grassy areas, areas with dirt, and tree cover and to play with natural features such as sticks, rocks, and mud. C: Traditional ELC were primarily held indoors and had outdoor play areas built on concrete. children spending less than 1.5 hr outdoors each day.	N/A	Harms	Poisson regression models Covariates: age
Fyfe-Johnson et al (2019), USA.	Controlled cross-sectional	Age: 3-5 years Gender:	E: The nature ELC occurs outdoors in a forested park where most children attend 5 days per week from 9 am to 1	N/A	Physical activity Cognitive Social and emotional	Descriptives only.

E: 20 children / 1 ELC C: 13 children (waitlist control or 2-hour nature-based, outdoor enrichment class provided by experimental ELC		E: 11m/9f C: 9m/4f SES: E: 18 > $90,000 C: 8> $90,000	pm; 2-day and 3-day per week options are available on a limited basis. The physical environment consists of dedicated classroom areas in the forested areas. Children use logs and tree stumps to sit; portable canopies are used during inclement weather. Most of the day is spent hiking and exploring the surrounding forest. No traditional play structures or pre-fabricated playgrounds are utilized. C: 2 hour nature-based outdoor enrichment class was offered once weekly by the same nature ELC the intervention group children attended. Classes were led by a teacher and attended by both child and caregiver. The classes consisted of science-based exploration through outdoor play in a forested park and involved: circle time, station time (learning stations that emphasize sensory and fine motor skills, creativity, and numerical and literacy skills), short stories, and hikes. Others were included in a wait-list control			
Giusti et al (2014), Sweden. E: 11 children / 2 ELC C: 16 children / 5 ELC	Controlled cross-sectional	Age: 5 years Gender not described. SES not reported.	ELC were assessed on their frequency of natural experiences. Each ELC was ranked according to the highest frequency of use of the greatest variety of nature experiences in its surroundings E: The ten ELC with the most frequent use of all nature experiences. C: The ten ELC with the least frequent use of all nature experiences.	N/A	Nature connectedness	t-test
Lysklett et al (2019), Norway.	Controlled cross sectional	Age: 5.1-6 years	Nature-based ELCs located close to a large recreational area, with woods,	N/A	Motor skills	T-test

E: 43 children / 4 ELC C: 49 children / 4 ELC		Gender: 53m/39f SES not reported	lakes and tracks just outside the city centre. Both types of preschools used the nearby nature area for hiking and playing every week E: nature ELC at least three times, per week C: traditional preschools once per week.			
Meyer et al (2017), Canada. E: 46 children / 3 ELC C: n= 35 children / 2 ELC	Controlled cross-sectional	Age: 5-6 years Gender: 39m/42f SES: predominately middle-class children	E: Children spent every morning in nature participating in teacher-directed, nature-based learning activities. The nature kindergartens differed per site but included a beach, unmanaged wooded area, natural playground (trees and vegetation) and artificial playground. C: Children were assessed in their classrooms where they engaged some storytelling, singing, dancing, tai chi, reading, drawing, and art. They also took part in music and computer classes and science fair.	N/A	PA	Descriptives only.
Moen et al (2007), Norway. E: 267 children / 37 ELC C: 264 children / 32 ELC	Controlled cross-sectional	Age: 3-6 years. Gender not reported. SES not reported.	E: had ‘‘outdoor’’ or ‘‘nature’’ as part of their name, or emphasized outdoor pedagogy and children spent an average of 3.5–8 hours/day outdoors in winter. C: children spend on average spend 1.25–4.0 hours/day outdoors.	N/A	Harms	GLM
Rice & Torquati (2013), USA. E: 68 children / 6 ELC C: 46 children /4 ELC	Controlled cross-sectional	Age: 56.4 months (12.8 SD) Gender not reported. SES: 46.5% of participants	E: The nature programme featured: vegetation, gardens, areas for digging in soil, sand, and “loose parts” (sticks, seeds, pinecones etc) and other naturally occurring objects that children used in their play. Climbing structures and pretend play structures such as a	N/A	Nature connectedness	ANOVA and Chi square

		reported an annual income of $85,000 or more.	boat or a playhouse were also included. C: The non-nature programmes consisted of pretend play structures, sand and/or wood chips, and paved surfaces for wheeled toys, and had few natural elements such as trees or grass.			
Robertson et al (2020), Australia. E: 15 children / 1 ELC C: 15 children / 1 ELC	Controlled cross-sectional	Age: 4-5 years Gender not reported. SES not reported.	E: ELC is in a rural area and consisted of a small traditional playground area (sand pit, obstacle course etc.) and a larger open ended nature area consisting of trees, shrubbery, grass, natural loose-parts). It has a highly naturalised area towards the rear that was rich in natural elements including small and large shrubbery, and larger tree and vegetation C: ELC is located in a suburban area and consisted predominately of man-made structures (almost half the space). The playground also consisted of some nature such as trees and vegetable garden.	N/A	Play	Independent samples t-test
Scholz & Krombholz (2007), Germany E: 45 children / 10 forest kindergartens C: Rural = 42 children / 2 ELC; Urban = 42 children / 2 ELC	Controlled cross-sectional	Age: E: 5.5 (SD 0.4) C: Rural= 5.7 (0.4 SD); Urban= 5.7 (0.4 SD) Gender: 71 boys, 58 girls SES not reported.	E: forest kindergarten C: traditional rural and urban kindergarten	N/A	Motor skills	MANOVA Covariates: age

Weisshaar et al (2006), Germany. E: 506 children / 25 ELC C: 1201 children / 28 ELC	Controlled cross-sectional	Age: 4.9 (1.1 SD) Gender: 901m/803f SES not reported.	E: Forest kindergarten located in forested areas where children spend all-season full-time outdoors. C: Conventional kindergartens (not described)	N/A	Harms	Fisher test and logistic regression Covariates: age, sex, skin inspection, and recommended vaccination
Ernst (2014), USA. E: 46 educators	Cross-sectional	Not described.	Outdoor environments that range from relatively natural to wild spaces.	N/A	Motor skills Cognitive Social and emotional Nature connectedness	Multiple regression
Wright (2019), USA. 48 children / 2 ELC	Cross-sectional	Age: 3-5 years Gender not reported. SES not reported.	The 2 sites were located in a forested park/ They both consisted of large space (10,000Sq/ft), log borders, sloping areas, vegetation, large trees, natural loose parts. Manufactured supplies such as shovels, wheelbarrows, books, magnifying glasses were brought in. 4 hours of the school day is spent outdoors.	N/A	Physical activity	Descriptives only
Author, year and country	**Study design**	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Follow-up time point**	**Outcome(s)**	**Data analysis**
Naturalised playgrounds						
Brussoni et al (2017), Canada. E: 48 children / 2 ELC	Uncontrolled before & after (mixed methods)	Age: 4.28 (0.63 SD) Gender: 53% m/47%f SES not reported.	Playgrounds were improved using the Seven Cs which consists of 27 items, rated on a 5-point scale, for a maximum score of 135 Changes predominately involved inclusion of more natural elements such as, vegetation, boulders, rock, loose parts. Seven Cs scores	Data were collected at T2; May-July 2014) two-weeks after playground modification	Physical activity Social and emotional Play	Wilcoxon signed rank tests; General linear modelling. Covariates: age, gender, ELC

			increased from 44 to 97 in ELC A, and 35 to 125 in ELC B.			
Cosco et al (2014), USA. E: 804 / 27 ELC	Uncontrolled before & after	Age: 2-5 years Gender not reported. SES not reported.	Preventing Obesity by Design is an ELC outdoor renovation intervention. Prior to the intervention the space had few structures (slides, swings etc.) in a rectangle space enclosed by a fence. Whereas, post intervention, the space had more natural elements, including trees, garden, vegetation etc.	Not described.	Physical activity Social and emotional	Logistic regression and bivariate correlations Covariates: gender
Cloward Drown et al (2014), USA. E: 24 children / 1 ELC (observed in 2 different playgrounds, natural vs manufactured)	Controlled cross-sectional	Age: 4.5 years Gender: 7m/17f SES not reported.	E: The natural playground was characterised by a majority of natural surfaces (vegetation, boulders, grass etc.) This playground also consists of sandbox, bikes pathway and instruments. C: The manufactured playground is equipment-oriented with hard surfaces. Although it includes some vegetation, the main features are a xylophone, slide, and pit, a ball pit, water play area and concrete ramps leading to a plastic play castle and a spin chair.	N/A	Play	Chi-squared
Luchs, & Fikus (2013), Germany. E: 38 children / 1 ELC C: 21 children / 1 ELC	Controlled cross-sectional	Age: 5-6 years Gender: 33m/26f SES not described	E: the natural playground provides children with wild and natural areas, including trees, grass, flowers etc. There are also sandboxes, dirt, rock and water and mud area. C: the contemporary playground provides traditional man-made structures, such as slide, sandbox, playhouse, water area, seesaw, roundabout etc.	N/A	Play	t-test
Carrus (2012), Italy. E: 16 children / 1 ELC	Cross-sectional	Age: 18-36 months (1.5-3 years). Gender not reported. SES not reported.	Free play in garden and green spaces of the ELC compared to free play indoors.	N/A	Cognitive Social and emotional	mixed model ANOVA with 2-way interactions

Dyment et al (2013), Australia. E: 120 children / 3 ELC C: 40 children / 1 ELC	Cross-sectional	Age: ELC A = older toddlers, young children; ELC B = young children; ELC C = older toddlers, young children, ELC D = 2-5 year olds Sex: 57%m/ 43% f. SES: the 4 centres differed in terms of SES (Centre A = high SES, B= varied SES, C= low SES, D= medium)	E: three centres all of which contained natural areas (trees, rocks, gardens). Two ELC's also has manufactured elements C: one centre which contained no natural areas	N/A	Play	Descriptives only.
Luchs, & Fikus (2018), Germany. E: 17 children / 1 ELC	Cross-sectional	Age: 5.85 ± 0.49 years Gender: 9m/8f SES not reported.	E: the nature playground has large natural space featuring trees, grass, hills, vegetations, water C: the contemporary playground has traditional play structures such as slides and swings. It has some natural elements, including grass and trees.	N/A	Physical activity	Paired sample t-test
Morrissey et al (2017), Australia. E: 28 children / 1 ELC C: 28 children / same school as E.	Cross-sectional	Age: 4-5 years Gender: 28m/28f SES not reported.	E: ELC contained natural structures such as logs, shrubs, rocks etc. It also contains a few manmade elements. C: a traditional space with standard man-made equipment such as swings and climbing frame. It also had some natural elements like trees but much less than the natural playground.	N/A	Play	Chi-square analyses
Storli et al (2010), Norway. E: 16 children / 1 ELC	Cross-sectional	Age: 3-5 years Gender: 9m/7f SES not reported	Nature - gathering loose nature materials, climbing running. Traditional - children engaged in activities such as cycling, digging, climbing	N/A	Physical activity	t-tests

Torkar & Rejc (2017), Slovenia. E: 25 children / 1 ELC	Cross-sectional	Age: 4 and 5 years old Gender: 16m/9f SES not reported.	E: forest playground which contains a forest patch, river and bushes. The space is approx. 500 m^2 C: Traditional playground which contains fixed equipment such as seesaw, roundabout, slide, climbers and playhouse. There is some nature surrounding the playground (trees, bushes). The space is approx. 500 m^2	N/A	Physical activity	Mann Whitney
Author, year and country	**Study design**	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Follow-up time point**	**Outcome(s)**	**Data analysis**
Types of natural elements						
Ng et al (2020), Australia. E: 159 children / 6 ELC C: 138 children / 5 ELC	Controlled before and after	Age: 2 years 10 months (0.82 SD) Gender: 49%m/51%f SES: No significant differences between intervention and control group reported.	Variable of interest was natural elements. Measured using the modified Environment and Policy Assessment and Observation (EPAO) physical environment domain. This tool assesses the prevalence of PA opportunities in the physical environment. There were 5 subscales: Fixed play equipment' and 'Portable play equipment' from the EPAO, 'Total size of playing area', 'Outdoor play spaces', and 'Natural elements'. A number of items per subscale were scored - 1 if present, 0 if not.	6 months	Physical activity	Mulitvariate linear regression Covariates: age, sex, parental education, accelerometer wear time.
Boldemann et al (2004), Sweden. E: 64 children / 2 ELC	Cross-sectional	Age: 1-6 years Gender:26m/38f SES not reported.	E: ELC 1 had play constructions surrounded by trees but exposed to the sun and ELC 2 had attractive play constructions positioned under a canopy of tree crowns. Average time spent outdoors was 207 min at site ELC 1, and 256 min at site 2.	N/A	UV exposure	t-tests
Boldemann et al (2006), Sweden.	Cross-sectional	Age: 4.5-6.5 years Gender: 114m/85f	ELC environment scores and averages dichotomized to (>2 high, <2 low)	N/A	Physical activity UV exposure	Bivariate analysis; Linear

E: 199 children / 11 ELC		SES not reported.	Outdoor environments were assessed on their play potential. They were scored 1, 2, and 3 with respect to size of outdoor area, overgrown surfaces (trees shrubbery) and integration of play structures or other defined play areas with vegetation.			mixed-models.
Christian et al (2019), Australia. E: 678 children / 48 ELC	Cross-sectional	Age: 3.4 ± 0.8 Sex: 53%m/47%f SES: 32% = low, 34% = medium SES and 34% = high SES.	ELC settings were dichotomized to vegetation < 3m in height or vegetation> 3m in height. High-resolution airborne multispectral 4-band images and Geographic Information System (GIS) was used to identify the location, shape and size of ELC outdoor play spaces. Approximately 31% of centres' outdoor play space had vegetation with 23% (20.5 SD) having <3 m in height and 8% (13.7SD) with >3 m hight.	N/A	Physical activity UV exposure	Multilevel linear regression models. Covariates: age, gender, and ELC SES and size.
deWeger (2017), Australia. E: 274 children / 12 ELC	Cross-sectional	Age: 4.2 years (0.5 SD) Gender: 141m/133f SES not reported.	Variable = natural elements The quality of the outdoor learning environment in the ELC's was assessed for 3 hours per day over 2 days using the POEMS instrument. This is grouped into 5 domains: **Physical environment (13 questions),** Interactions (13 questions), Play and Learning Settings (13 questions), Program (9 questions), and Teacher/Caregiver role (8 questions). Scores are them summed to give a total score	N/A	Physical activity	Hierarchical linear modelling (HLM) Covariates: age, gender, BMI-z score and accelerometer wear time (level 1), outdoor environment quality (level 2)
Gubbels et al (2018), Netherlands.	Cross-sectional	Age: 34.14 months (8.97 SD) Gender: 72m/79f	The SB and PA physical environment of each ELC was assessed using a standardized observation protocol, based on the updated Environment	N/A	Physical activity	Multivariate linear regression analyses

E: 151 children / 22 ELC		SES not reported.	and Policy and Assessment Observation (EPAO). The following natural elements were assessed: large trees (2.5 m or taller), small trees (less than 2.5 m tall), trees that children can climb, shrubs, flowering plants, variation in ground (hills, mounds), grass, rocks large enough to climb, a hill for rolling down or climbing up. A sum score of all the types of natural elements that were present was calculated.			
Maartensson et al (2009), Sweden. E: 198 children / 11 ELC	Cross-sectional	Age: 5.26 (0.56 SD) Gender: 113m / 85f SES not reported	The outdoor settings of each preschool were dichotomized into ‘‘high-score’’ and ‘‘low-score’’ environments in analysis The following were assessed: A. Total outdoor area. 1= small (<2000 m^2), 2= medium (2000–6000 m^2), 3= large (46000 m^2) B. Proportion of the area containing shrubbery, trees or hilly terrain: 1= little/non-existent, 2= <half of the area, 3= >half of the area C. Integration between vegetation, open areas and play structures: 1= no integration. 2= either (a) Play structures adjacent to trees and shrubbery or integrated into areas, or (b) The open spaces are located in between play-areas and not in separate parts of the environment. 3= environments fulfilling both 2a and 2b above. Outdoor environments were scored 1, 2 or 3 along three elements. The three scores of each environment were	N/A	Cognitive	Nested mixed model

			summed up and divided by 3, yielding an average score for each environment ranging from 1 to 3.			
Määttä et al (2019), Finland. E: 864 children / 66 ELC	Cross-sectional	Age: 4 years 4 months (10 SD) Gender: 48% girls SES: 29% had mother with high educational background (at least masters)	Observation instrument was designed for the study and consisted of items from the EPAO. ELC physical environments were assessed, of which, surfaces in the preschool grounds (9 items) and terrain in the playground, related to the natural environment (grass, forest, trees, rocks).	N/A	Physical activity	Multilevel linear regressions models Covariates: age, gender, season, municipality, pre-school group cluster
Määttä et al (2019b), Finland. E: 655 children / 66 ELC	Cross-sectional	Age: 4.7years (0.89 SD) Gender: As above SES: As above	Frequency of nature trips (mean/per week): Teachers completed weekly diary of activities which were categorised into 5 groups (1=outdoors, 2=teacher-led sessions, 3=free play, 4=organised PA lessons and 5=mixed sessions). Daily number of each activity was calculated and summed for the week level and then divided by the number of the days (from 3 to 5) to form the average daily amount of each activity. A questionnaire was then completed to determine activities that are close to the ELC and occur regularly (nature visits). Visits were recorded for mean times per week	N/A	Physical activity	Multilevel linear regressions models. Covariates: age, gender, average attendance at preschool and study season
Olesen et al (2013), Denmark. E: 441 children / 42 ELC	Cross-sectional	Age: 5.8 years Gender: 49.5%m/50.5%f SES not reported.	Researchers collected a range of environmental correlates, of which, vegetation and hilly landscape related to nature	N/A	Physical activity	Univariate analyses and multi-level modelling

						Covariates: Gender, rain, preschool type, afternoon hours, location, indoor area, Playground area, playground time, parent education
Sando (2019), Norway. E: 80 children / 8 ELC	Cross-sectional	Age:3.5 (SD=0.5) Gender: 41m/39f SES not reported.	The places and materials in the playground were categorised into nature, pathways, open area and fixed functional equipment. Nature was coded in four of the institutions and ranged from large forest areas (1500 m^2) to smaller areas with trees and natural surfaces.	N/A	Physical activity Social and emotional	A random intercept multilevel model Covariates: age, gender
Sando & Sandseter (2019), Norway. E: 73 / 8 ELC	Cross sectional (mixed-methods)	Age: 4.2 years (0.7 SD) Gender: 36m/37f SES not reported.	ELC settings featuring nature were coded (places). For objects, these were coded when a child was holding, using or interacting with an object and included: sand, water, mud and nature materials The variables for places and objects describe the percentage of time the child is at a place or in which the object was used during each observation.	N/A	Physical activity	Generalized linear latent and mixed models
Söderström at al (2013), Sweden. E: 172 children / 9 ELC	Cross-sectional	Presented per ELC Age: S1: 4.6 (1.0 SD) S2. 4.1 (0.5 SD) S3: 4.3 (0.7 SD) S4: 4.4 (0.8 SD) S5: 4.7 (0.8 SD) S6: 4.6 (0.9 SD)	Outdoor Play Environment Categories (OPEC) scoring tool was used to assess playgrounds on (i) total outdoor area, (ii) amount of trees, shrubbery and hilly terrain and (iii) integration between vegetation, open areas and play structures, each component with a	N/A	Sleep Harms Weight status Social and emotional	ANOVA and MANOVA Covariates: Age, gender, birth Weight, mother SES.

		S7: 4.3 (0.9 SD) S8: 4.6 (0.6 SD) S9: 4.8 (0.7 SD) Gender: % f S1: 29% S2. 41% S3: 50 % S4: 42% S5: 50% S6: 56% S7: 61% S8: 41% S9: 63%	score range of 1–3 (high score = high quality). The OPEC scores were then dichotomized (low OPEC value< 2, high OPEC value >2)			
Sugiyama et al (2012), Australia. E: 89 children / 10 ELC	Cross-sectional	Age: 4.1 (0.6 SD) Gender: 54%m/46%f SES not reported	Questionnaire assessing characteristics of the ELC's was completed by the centre Director. Outdoor characteristics of relevance were gradient shade, vegetation, surface material (grass).	N/A	Physical activity	Multilevel linear regression Covariate: age, gender and time spent outdoors
Zamani (2013), USA. 36 children / 1 ELC	Cross-sectional (mixed-methods – thesis)	Age: 4-5 years Gender: 21M/15 F SES not reported	Natural zone: wild landscape with non-structured green space (0.40 acres). The natural zone is rich in natural loose elements, such as leaves, twigs, dirt, stones and includes two looped and one straight pathways and boulders. The crawling equipment referred as the "green tube" is the sole manufactured element. This zone also includes three rope settings, tied to the trees. Mixed zone: A widespread mixed outdoor environment of 0.48 acres referred as the "hill". The mixed zone has a moderate, downward slope from its entrance. There is rocking	N/A	Play	Chi square analysis

			equipment, a linear pathway along the hill, a music wall with a stage, a set of six swings, a sand box, a gazebo, a stoned stone-lined swale without water, and two dramatic play settings. There is also a wood which includes a wooden platform, ropes, and musical instruments attached to the trees Manufactured zone: a dramatic play setting (play house), a looped pathway, a composite play structure, a porch, a sand play setting (covered with a shade structure), bike sheds, bikes and scooters, storage (for storing toys and loose material), three gathering settings (benches and tables), a swing pergola, and a basketball loop. This zone also includes a transitional space between the indoors and outdoors. The manufactured zone has a smaller square footage (0.11 acres) compared to the other zones.			
Author, year and country	**Study design**	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Follow-up time point**	**Outcome(s)**	**Data analysis**
Garden-based intervention						
Lillard (2016), USA. E: 55 children / 1 ELC Delay Gratification E: 34 children	Uncontrolled before & after	Age: delay gratification= 4.16 years (9.9 months); Beery = 4.07 years (339.38 days) Gender: 40m/51f (based on students who were assessed) SES not reported	Gardening programme (not clearly described).	6 months	Cognitive	Repeated measures ANOVA

Visual motor integration E: 39 children						
Park et al (2016), South Korea. E: 336 children /12 ELC Science investigation abilities and attitudes= 68 children	Uncontrolled before & after	Age: 5-7 years Gender: 169m/167f SES not reported.	The intervention consisted of horticultural activities that increase children's knowledge of seeds, soil, planting and harvesting etc. The intervention consisted of 24 sessions delivered once per week and lasted an average of 50 minutes per session	Intervention lasted 24 weeks. Outcomes were assessed one week prior to the intervention and one-week post intervention	Cognitive Social and emotional	Paired samples t-test

Abbreviations: E= experimental; C= control; n= number; m=male; f= female; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; SES= socioeconomic status; USD= US Dollars; GLM = General linear modelling.

Table 2. Characteristics of included qualitative studies					
Author, year and country	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Research aims**	**Data collection method**	**Details of analysis**
Nature-based ELC					
Bjørgen (2016), Norway. 24 children / 1 ELC	Age: 3-5 years Gender: 10m/14f SES not reported.	Children played in the ELC outdoor play space for 3 hr/day, and each week would go on trips (1 or 2x) to natural environments. The large outdoor area consists of outdoor toys (buckets, shovels, trucks, balls), swings, sandboxes, climbing racks, natural materials, small trees, a varied surface of grass, sand, asphalt, and small hills. The destination for excursions in diverse natural landscape environment is approximately 300–700m from the centre. One type of natural environment was open fields suitable for tobogganing, running and playing on skis. Another natural environment consisted of woods. Trips were made to the natural environments all year round.	What is the relation between environmental affordances and PA levels among 3–5 year olds?	Observations were made with video recording the different seasons of the year for 20 days, 10 days on trips in a natural environment and 10 days in the centres play space. A total of 50 h of direct observation was conducted. Coding of the physical activity levels of children was assessed and adapted using the Observational System for Recording Physical Activity in Children-Preschool Version (OSRAC-P) manual.	Thematic analysis - the first phases of coding were assessing and identifying the children's level of PA in different play situations. Figures were used as an analytical tool helped to discern patterns, differences and similarities in the data material, which laid foundations for the qualitative analysis of the affordances. Thereafter themes of affordances are identified within the data. The theory of affordances and criteria from the 7Sc were used in the analysis process.
Dowdell et al (2011), Australia.	Age: 2-6 years Gender: 6m/6f	E: Has an emphasis on nature and sustainable education. The space is large and consists of sandpit, fairy garden, play equipment, grass area and vegetable garden.	How are children's play behaviours and social interactions	Play behaviours were recorded using a behaviour mapping schedule. Each child was observed individually and	Once all the observations were made for each child at each centre they were then tallied up.

E: 6 children / 1 ELC C: E: 6 children / 1 ELC	SES not reported.	C: Located in a warehouse this centre has an entirely artificial indoor play area. It consists of a bike track, home corner (playhouse etc), climbing structures, quiet play area, sandpit and obstacle course.	influenced by the opportunities and materials present in their outdoor play environment?	every 10 seconds an observation based on social interaction and play behaviour was recorded.	Play behaviours were then categorised into four different groups: social activities, cognitive activities, physical and motor skill activities and other activities.
Liu (2020), USA Nature interaction: E: 29 children / 1 ELC C: 26/ 1 ELC Restorative experiences: E: 10 children / 1 ELC C: 9 children/ 1 ELC	Age: 4-5 years Gender: 30m/ 25f SES: E: 48,000 US (household income); C: 59,000 (household income) of children attending each centre	E: contains high levels of nature with a variety of perceived affordances. Outdoor time = 1.5 hours/day. 32 types (categories-vegetation (tress, shrubs, flowers, grasses), natural ground surface (wood chips, meadow, multipurpose lawns), natural materials, natural play structures (e.g. wood, stick, water, sand logs, ice, leaves), animals, experiential elements (rain, snow, sky view, light, air) of natural elements and play settings and 4 types of non-nature-based play settings (concrete track, bicycles, concrete hall, concrete sq.) were identified C: low levels of nature and perceived affordances. Outdoor time = 1.5 hours/day. 13 types of natural elements and 11 (vegetation, natural ground, animals) types of non-nature-based play settings (examples include: play structure, playhouse, outdoor kitchen, bicycles) were identified.	How does the designed nature-based outdoor play environment in ELC impact children's interaction with natural elements? How does the designed nature-based outdoor play environment in ELC impact children's restorative experience?	RQ 1. Field observation, behaviour mapping, semi-structured interview with teachers. RQ2. Field observation, structured Interview with children, semi-structured interview with teachers.	Content analysis was used for: children's frequent play locations, types of play behaviors, frequency and diversity of different ways of interaction with natural elements, as well as restorative experience from semi-structured interviews with teacher and structured interview with children. Themes (coding categories) were drawn from the theoretical framework. Specifically, children's types of play behaviors and their ways of interacting with natural elements were coded using function taxonomy of affordance (Heft, 1988; Kyttä, 2002) and Gibson's affordance theory.

Maynard et al (2013), Wales, UK. 48 children / 8 ELC	Age: 4-7 years Gender: 24m/24f SES not reported.	Educators introduced child-initiated learning in the outdoor environments. The kinds of activities varied and incorporated free play with natural resources (e.g. ELC A, F and H); growing vegetables (ELC C); (ELC B); and more structured investigations – for example, of snails (ELC D), air/wind (ELC E) and flight (School G). All the teachers had access to a small tarmac yard or grassed area. These were seen by the teachers as 'outdoor classrooms' and used for painting, sand and water play, construction activities etc. The teachers also had access to some additional outdoor space – playing fields, vegetable gardens or common land. 3 ELC settings (A, G and H) had extensive outdoor environments incorporating different types of play equipment or natural features such as a willow tunnel and pond.	To explore these perceived differences as well as teachers' perceptions of 'underachievement'.	Researcher visited teachers three times to undertake individual semi-structured interviews. Interviews were audio recorded and field notes at each interview. Teachers also provided case studies of each student	Interviews were transcribed using Nvivo8. A thematic analysis approach was used where data were analysed in three ways with increasing depth: 1. perceived difficulties of children 2. case studies 3. theoretic issues related to "place and space"
Sandseter (2009), Norway. 29 children from both experimental and control groups E: 1 ELC	Age: 4-5 years Gender: 21f/8m SES not reported.	E: Located in a forest with no fixed play equipment and fencing and children spent most of their time outdoors. C: fixed equipment, such as swings, climbing tower, play hut and a few trees.	To explore affordances for risky play in two different play environments: an ordinary ELC playground and a nature playground.	7 days were spent on each of the ELC playgrounds. Video recordings and field notes of risky play situations were collected based on categories of risky play; a) great heights, b) high speed c) dangerous tools, d) dangerous elements, e) rough-and-tumble play, f)	A content analysis was performed on the data. The analysis was theory-driven. Firstly, each of the play environments' potential affordances for risky play, as categorized by Sandseter (2007), were analysed in relation to the most relevant affordance

C: 1 ELC				where the children can disappear/get lost. Both the children's play and the staff's supervision were observed. The field notes and the video recordings were transcribed into an electronic word file. 12 children in the ordinary preschool and 11 children in the nature and outdoor preschool participated in a one-to-one qualitative interview with the researcher. Each interview was approximately 20-30 minutes and was recorded on audiotape. The interviews were semi-structured, using an interview guide list of questions and issues. The interview guide was based on the six categories of risky play and aimed to explore the types of risky play that the children engaged in within the different play environments and whether the staff constrained or intervened in their actions. Upon completion of the interviews, the audiotapes were professionally transcribed verbatim into an electronic word file.	categories to evaluate their potential affordances for risky play. Secondly, the transcriptions of the video observations, field notes, and interviews were examined to determine the types of risky play children engaged in within different environments. Thirdly, the observations and the interviews were analysed to determine the degree to which children experienced mobility license while engaging in risky play. The transcriptions of the video observations were examined to determine the extent to which, and in which situations, the staff had children under surveillance while they engaged in risky play or was taking initiative to or constrained risky play.

Streelasky (2019), Canada. 15 children / 1 ELC	Age: 5-6 years Gender not reported. SES not reported.	The ELC setting had an outdoor, nature-based focus where children spent afternoons in the forested area. The teacher who was involved in an Outdoor Environmental Leadership Programme engaged the students in an integrated learning approach where key curriculum areas were addressed (e.g. language arts, social studies, science and physical education). Children also had time to freely explore the forest.	What learning experiences do kindergarten children value at school? and what modes are they choosing to express and represent their valued school learning experiences?	Qualitative interpretative approach involving (i) group discussions, (ii) participant observations, (iii) anecdotal notes, (iv) artefact collection and (v) individual semi-structured interviews (children's narratives).	Data were analysed and grouped into themes. Image based analysis was used to develop deeper understanding of children's interests and knowledge. Thematic analysis was used to gain insight into children's practices which followed 6 phases: (i) familiarising oneself with the data and identifying items of potential interest, (ii) generating initial codes, (iii) searching for themes, (iv) reviewing potential themes, (v) defining and naming themes and (vi) reporting the themes.
Author, year and country	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Research aims**	**Data collection method**	**Details of analysis**
Naturalised playgrounds					
Herrington & Studtmann (1998), USA.	Age: 2-6 years Gender:	Pre-modification: Lab A: consisted of a patio area, grass lawn, play structures, swing	What natural materials and conditions of the outdoor	Phase 1: sequence sampling of children during free-play. Children were video-taped interacting	20 hours of videotapes were analysed. During analysis, notes were made. For Phase 1 the

36 children / 1 ELC (2 “labs”)	16m/20f SES not reported.	set, doll house, trees and vegetation. Lab C consisted of a porch area, grass lawn, play areas, swing set, trees and vegetation. Post-modification: Playground were naturalised with increased natural elements: ice sculptures, wind chimes, canopy, chalk, buckets, playhouse, water pay, vegetation and trees were added to the labs. Lab A received more natural elements than lab C but both were more natural post intervention.	environment can contribute to the development of young children ranging from 2 to 6 years old?	with the site for 1 month. Once the modifications were made, data collection began a week later. Data collection involved video-taping, sound recording, and field notes. Videotaping involved following a child for 20 minutes as they moved throughout the yard in free play. Voice recordings of the children were made of one of the two selected children from each Lab. Voice recordings were transcribed into text documents. Field notes (weather, teacher and children present, anecdotal observations etc.) were made daily by researchers. Notes were recorded by researchers on a pre-printed notation sheet that displayed a plan view of both yards. Phase 2: Video documentation and anecdotal notes were employed to record event sampling. Event sampling allowed subjects to be taped if they interacted with the plant	notes were: (1) interaction with an intervention (2) duration of interaction (3) children's behavioural modification made between pre and post intervention (4) children's movement changes made between pre and post intervention. For Phase 2 the criteria were: (1) which children were engaged in the intervention; (2) how many children were engaged (3) the duration and nature of their engagement with the intervention (4) how behavior and paths of movement changed between pre and post intervention. Video clips were selected that illustrated the notes. These clips were put together on one VCR tape using a television and VCR recorder. The conversations of the children participating in Phase 1 were transcribed at 10 second intervals. The anecdotal

				interventions. The specific intervention sites were recorded on a rotating basis. Children were video-taped using the same schedule as in Phase 1 and fieldnotes were made in the same manner as in Phase I	notes were reviewed and complied.
Puhakka et al (2019), Finland. 12-24 children (not clear) / 6 ELC	Age: 3-5 years Gender not reported. SES not reported.	Playground yards were transformed through enhancing the biodiversity by incorporating more greenspace and vegetation. For example, replacing areas covered in gravel with forest floor. Children spent time outdoors every day (0.5–2 h in the morning and in the afternoon) as well as participating in teacher led activities 4-5 days/ week.	Does biodiversity exposure and greening playgrounds affect 3–5 years-old children's physical activity and play, their environmental relationships, and their well-being in the urban environment in Finland.	Educators and child nurses completed interviews and surveys respectively. 49 parents completed surveys. Surveys were completed one month after the playground was modified. Surveys included both structured and open ended questions which related to children's play activities, and enthusiasm. Interviews with parents focussed on children perception of modifications. The educator thematic interviews focused on possible changes in children's play and other activities in the yard, in children's and educators interest in and knowledge of nature, their well-being, attitudes towards outdoor activities, and	Interviews were recorded and transcribed verbatim. Survey and interview data were analysed using qualitative content analysis to identify different affordances. The affordances were then classified into 6 themes which emerged from analysis and coding. How these affordances supported children's relationship with the modified playground were then mapped. Finally, these two elements were brought together to form three perspectives.

				practices and atmosphere in the ELC setting	
Wishart et al (2019), Australia. 75 children / 1 ELC	Age:4-5 years Gender not reported. SES not reported.	The two playgrounds were located on different sides of the building, each extending to the back of the building where a connecting gate was sometimes opened to allow free-flow of children between the two spaces. E: Traditional equipment was replaced with terraces, inclines, logs and rocks designed to afford physical activities and gross motor skills such as climbing and balancing. other elements included: Natural gardens with fruit trees; herb garden and small plants; logs; stepping-stones; log enclosure; small tree forest; sandpit with pebbles and medium-size rocks. C: standard equipment: slide, ladders, swings, climbing frames, sand-pit, surfaces open area. This area also included a grass area, veg garden, trees and shrubs.	Does the naturalised design of the new space provide equivalent actualisable affordances for different types of physical activity to those provided by the more traditional playspace, with its conventional equipment and resources	Behaviour mapping using a time-sampling observation tool. Observations were conducted between 10:30–15:30 during sessions. The two playscapes were divided into zones and children were observed in 3 minute cycles. For each observation, the tool also noted: number of boys and girls (no further count of children was taken); presence of educators; whether play was solitary or group; location and general contextual information. 40 observations in the naturalised space and 42 observations in the traditional space were made.	Behaviour mapping tracked the incidence of different categories of movement across different areas of the two playscapes, to investigate if different categories of movement were more likely to occur in specific areas or in relation to specific features.
Author, year and country	**Age (range or mean ± SD), sex (n or % m/f), SES.**	**Exposure and comparison**	**Research aims**	**Data collection method**	**Details of analysis**
Types of natural elements					

Zamani (2015), USA. 36 children / 1 ELC	Age: 4-5 years Gender: 21M/15 F SES not reported.	See quantitative study characteristics table.	How does an outdoor learning environment with natural features can stimulate children's cognitive play behaviors	1. Photo preference - researcher captured photos based on particular behavior settings or elements of the outdoor environment. The photos represented particular spaces in which children engaged in certain behaviors. The researcher used photo preference to ask children to select their preferred outdoor settings and elements and explain about their play. 2. Drawings from children - The researcher asked children to draw their favourite outdoor play spaces as a means for the researcher to evaluate each setting's cognitive play affordances and the elements children enjoyed. 3. Structured interviews with children - Interview questions aimed toward understanding children's choice of photos, drawings, and opinions of the outdoor learning environment. 4. structured interviews with teachers - to understand the teachers'	1. Used with transcribed child interviews and then coded these into different cognitive play behaviours. The photos were used to understand child's explanations. 2. The analysis of the drawings included three stages. In the first stage, the researcher quantified all 22 drawings by coding their visual features; The drawing codes established the element or behavior setting types depicted in the image; The researcher further evaluated the drawings on the frequency that certain settings or elements appeared 3. Interviews recorded and transcribed and then grouped by themes 4. transcribed and then grouped into themes related to teachers view on curriculum, outdoor learning environment, value of children's play, what children prefer, cognitive play affordances.

				perspectives toward the outdoor environment and children's daily interactions. The interview questions (6) prompted teachers to discuss the play opportunities the different zones provided for children. The following section explains the protocols regarding each of the described methods.	
Abbreviations: E= experimental; C= control; n= number; m=male; f= female; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SES= socioeconomic status; PA= physical activity.					

Appendix D. Quality of included quantitative studies as assessed by the EPHPP tool

Study ID	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawals and drop-outs	Final Grade
Agostini et al (2018)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Barrable et al (2020)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	1 = Strong	N/A	**3 = Weak**
Boldemann et al (2004)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Boldemann et al (2006)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Brussoni et al (2017)	2 = Moderate	2 = Moderate	2 = Moderate	3 = Weak	1 = Strong	1 = Strong	**2 = Moderate**
Carrus (2012)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Choi et al (2014)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	3 = Weak	1 = Strong	**2 = Moderate**
Christian et al (2019)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Cloward Drown & Christensen (2014)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Cooper (2018)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Cordiano et al (2019)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Cosco et al (2014)	1 = Strong	2 = Moderate	1 = Strong	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
deWeger (2017)	2 = Moderate	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Dyment et al (2013)	1 = Strong	3 = Weak	3 = Weak	3 = Weak	1 = Strong	N/A	**3 = Weak**
Elliot et al (2014)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	1 = Strong	1 = Strong	**2 = Moderate**
Ene-Voiculescu &Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001)	3 = Weak	1 = Strong	1 = Strong	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Ernst (2014)	2 = Moderate	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Ernst & Burcak (2019)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**

Ernst et al (2019) & Ernst & Burcak (2019)	3 = Weak	2 = Moderate	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Wojciehowski & Ernst (2018) & Ernst & Burcak (2019)	3 = Weak	2 = Moderate	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Burgess & Ernst (2020)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Zamzow & Ernst (2020) & Ernst & Burcak (2019)	3 = Weak	1 = Strong	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Frenkel et al (2019)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Fyfe-Johnson et al (2019)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	1 = Strong	N/A	**3 = Weak**
Giusti et al (2014)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	3 = Weak	N/A	**3 = Weak**
Gubbels et al (2018)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Lillard (2016)	3 = Weak	2 = Moderate	1 = Strong	3 = Weak	1 = Strong	1 = Strong	**3 = Weak**
Luchs, & Fikus (2013)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Luchs, & Fikus (2018)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Lysklett et al (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Määttä at al (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Määttä et al (2019b)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Maartensson et al (2009)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Meyer et al (2017)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Moen et al (2007)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	3 = Weak	N/A	**3 = Weak**
Morrissey et al (2017)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Müller et al (2017)	3 = Weak	1 = Strong	3 = Weak	3 = Weak	1 = Strong	1 = Strong	**3 = Weak**
Nazaruk & Klim-Klimaszewska (2017)	3 = Weak	2 = Moderate	3 = Weak	3 = Weak	3 = Weak	1 = Strong	**3 = Weak**
Ng et al (2020)	3 = Weak	1 = Strong	1 = Strong	3 = Weak	1 = Strong	1 = Strong	**2 = Moderate**
Olesen et al (2013)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Park et al (2016)	3 = Weak	2 = Moderate	3 = Weak	3 = Weak	1 = Strong	1 = Strong	**3 = Weak**

Rice & Torquati (2013)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Robertson et al (2020)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Sando (2019)	2 = Moderate	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	**3 = Weak**
Sando & Sandseter (2019)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Scholz & Krombholz (2007)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	3 = Weak	N/A	**3 = Weak**
Söderström at al (2013)	2 = Moderate	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Storli et al (2010)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	1 = Strong	N/A	**3 = Weak**
Sugiyama et al (2012)	3 = Weak	3 = Weak	2 = Moderate	3 = Weak	1 = Strong	N/A	**3 = Weak**
Torkar & Rejc (2017)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Weisshaar et al (2006)	2 = Moderate	3 = Weak	2 = Moderate	4 = Weak	1 = Strong	N/A	**3 = Weak**
Wright (2019)	3 = Weak	3 = Weak	3 = Weak	3 = Weak	3 = Weak	N/A	**3 = Weak**
Yılmaz et al (2020)	3 = Weak	2 = Moderate	3 = Weak	3 = Weak	2 = Moderate	1 = Strong	**3 = Weak**
Zamani (2013)	3 = Weak	3 = Weak	1 = Strong	3 = Weak	2 = Moderate	N/A	**3 = Weak**

Appendix E. Findings per eligible study

Quantitative

PHYSICAL

Table 1. Nature-based ELC on physical activity								
Study details (Author, year and country) **Sample size** (n children / n ELC)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Accelerometer								
Nature-based ELC								
Müller et al (2017), Canada. E: 43 children / 1 ELCs C: 45 children / 1 ELCs	Controlled before & after	SB and MVPA ActiGraph GT1M measured for 5 consecutive school days on three separate occasions: Oct (start of school yr), Jan and Apr (end of school yr). Cut points not described	SB (mins/ ELC day)	E: Oct= 167 Jan= 174 C: Oct= 178 Jan= 178	Apr= 151 Apr= 152	Within-group seasonal differences, but no between-group differences. (inferential statistics not provided)	▲	Weak
			MVPA (mins/ ELC day)	E: Oct= 74 Jan= 79 C: Oct = 79 Jan= 79	Apr = 68 Apr= 62	As above.	▲	
Fyfe-Johnson et al (2019), USA.	Controlled cross-sectional	PA and SB ActiGraph GT3X+ accelerometer worn for a minimum if 5	Habitual PA (mins/ day) SB	E: 467 (60 SD) C: 453 (51 SD)	Mean diff: 14.4, (95% CI: -29.1, 58.0)	Children who attended nature-based ELC engaged in more SB, and less light PA and MVPA.	▼	Weak

E: 20 children / 1 ELCs C: 13 children (waitlist control or 2-hour nature-based, outdoor enrichment class provided by experimental ELCs		days (inc 1 weekend). Weartime for total PA was 656 (59 SD), C= 667 (59 SD) Pate et al. (2006) cut points	Light MVPA	E: 91.6 (13 SD) C: 102 (10 SD) E: 97.4 (16 SD) C: 113 (24 SD)	-10.1 (95% CI: -19.2, -1.0) -15.5 (95% CI: -31.9, 0.87)			
			Habitual Weekday PA (mins/day) SB Light MVPA	 E: 468 (66 SD) C: 461 (54 SD) E: 93.5 (18 SD) C: 101 (15 SD) E: 97.1 (21 SD) C: 112 (30 SD)	 6.9 (95% CI: -40.1, 54.0) -7.3 (95% CI: -20.1, 5.4) -14.9 (95% CI: -36.3, 6.5)	As above.	▼	
			Habitual Weekend PA (mins/day) SB Light MVPA	 E: 486 (65 SD) C: 453 (51 SD) E: 88.7 (14 SD) C: 103 (15 SD) E: 95.8 (16 SD) C: 113 (22 SD)	 33.0 (95% CI: -14.8, 80.9) -14.2 (95% CI: -25.9, -2.4) -17.7 (95% CI: -33.8, -1.5)	As above.	▼	
			PA (mins/ ELC day – 9.00-13.00) SB Light	 E: 153 (19 SD) C: 166 (13 SD) E: 31.8 (11 SD) C: 32.7 (5 SD)	 -13.5 (95% CI: 63.3, 54.2) -0.9 (95% CI: -2.1, 0.64)	As above, but the differences in light PA and MVPA were much smaller.	▼	

			MVPA	E: 33.2 (15 SD) C: 34.7 (7 SD)	-1.5 (95% CI: -2.8, 1.2)			
			Sedentary bouts (ELC day) Bout, total number Bouts, number per day Bouts, total length Bout, average length	 E: 6.3 (3 SD) C: 6.4 (4 SD) E: 1.9 (1 SD) C: 2.0 (1 SD) E: 88.9 (47 SD) C: 100 (59 SD) E: 12.8 (5 SD) C: 16.1 (3 SD)	 -0.05 (95% CI: -2.9, 2.8) -0.11 (95% CI: -0.94, 0.73) -11.3 (95% CI: -54.4, 31.7) -3.3 (95% CI: -6.7, 0.13)	Children who attended nature-based ELC had similar total bouts and number of bouts per day to the control group. The bout total and average length were also higher in the control group.	▼	
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Naturalised Playground								
Brussoni et al (2017), Canada. E: 48 children / 2 ELC	Uncontrolled before & after (mixed methods)	MVPA ActiGraph GT3X/GT3X+ worn during scheduled outdoor time (20 mins). Pate et al. (2006) cut points	MVPA (mins/ outdoor time)	Not presented.	- 1.32 min, 0.37 SE, p< 0.001	There was a significant decrease in time spent in MVPA from T1 to T2 across ELC's.	▼	Moderate
Luchs, & Fikus (2018), Germany.	Cross-sectional	Gait cycles Microprocessor-based pedometer	Gait cycles/mins at playground	E: 25 (4.99 SD) C: 28.55 (9.60 SD)		No significant difference in mean gait cycles/min between the nature	▼	Weak

E: 17 children / 1 ELC		(StepWatch, Orthocare Innovations, Washington DC, USA) Worn twice for 45 minutes, once on the nature playground and once on the traditional playground.		p = 0.109, d = 0.54)		and traditional playground.		
Storli et al (2010), Norway. E: 16 children / 1 ELC	Cross-sectional	CPM ActiGraph (model not described) Worn for three separate days over 6 months, including 2 days of outdoor activity on the pre-school playground (winter and spring) and one day in nature (spring). Wear time varied between 102–136 minutes Cut points not described,	Mean CPM	E: (spring) 1292 (307 SD) C: (spring) 1261 (426 SD) C: (winter) 1496 (475 SD) (*p*= 0.01)		There is an association between the levels of PA for the natural environment and traditional (spring and winter) playgrounds meaning PA levels are similar across the environments.	▲(spring)	Weak
Torkar & Rejc (2017), Slovenia. E: 25 children / 1 ELC	Cross-sectional	Distance (km) Measured using GPS for 20 mins.	Distance (km)	E: 0.72 (0.49 SD) C: 0.49 (0.19 SD) (p= 0.132, r= 0.21)		There were no significant differences between the forest and traditional playground.	▲	Weak

Study details / Sample size	Study Design	Outcome and measurement	Units	Baseline or one time point (cross-sectional)	Follow-up (if applicable) or mean difference	Summary of Findings	Effect Direction	Quality Rating
Types of natural element								
Ng et al (2020), Australia. E: 159 children / 6 ELC C: 138 children / 5 ELC	Controlled before and after	PA ActiGraph GTX3+ worn during ELC days ELC monitoring days were considered valid based on at least 1 day at ELC with 75% wear time Pate et al. (2006) cut points	Total PA min/ ELC day) MVPA min/ ELC day)	β= 14.46, p< 0.01 β= 10.04, p< 0.01		Natural grassed area was positively associated with Total PA and MVPA. Non-significant time x group interaction for natural elements on Total PA and MVPA (regression coefficients not presented)	▲ N/A	Weak
Boldemann et al (2006), Sweden. E: 199 children / 11 ELC	Cross-sectional	Step counts Yamax Digiwalker SW-200, MLS 2000 pedometer. Wear time not detailed.	Step counts/ min ELC day	High environment = 21.6 (95% CI: 20.6–22.5) Low environment = 17.7 (95% CI: 16.8–18.6) p<0.001		High environment score increased step count	▲	Weak

Christian et al (2019), Australia. E: 678 children / 48 ELC	Cross-sectional	Total PA Actigraph GT3TX+ Valid data included at least 1 day at ELC with 75% wear time. Data was averaged for children who attended more than 1 day during the 7-day monitoring period. Pate et al. (2006) cut points	Total PA (min/ ELC day)	% < 3m vegetation: β<-0.01 (95% CI: -0.22,0.21), p= 0.96) % > 3m vegetation: β = 0.02 (95%CI: -0.28, 0.32), p=0.89		Shade-related variables (vegetation < 3 metres in height and vegetation > 3 metres in height) were not significantly associated with minutes/day of total PA.	▲	Weak
			MVPA (min/ ELC day)	% < 3m vegetation: β = -0.01 (95% CI: -0.18, 0.16), p=0.91 % > 3m vegetation: β = 0.08 (95%CI: -0.16, 0.32), p=0.52		As above for MVPA	▲	
deWeger (2017), Australia. E: 274 children / 12 ELC	Cross-sectional	Total PA and MVPA (min/day at ELC), cpm and step counts Actigraph GT3X+ Accelerometers were worn for one ELC week (range of 1-5 days). Mean wear time was 390 minutes (87.4) or for 6.5 hours (1.5).	Total PA (min/ ELC day)	intercept= 59.5, coefficient= 3.5, 1.8 SE, t= 1.89, p= 0.060		No significant association between setting with natural elements on total PA.	▲	Weak
			MVPA (min/ ELC day)	intercept= 10.3, coefficient= 1.7, 1.2 SE, t= 1.37, p= 0.17		As above for MVPA.	▲	
			Mean CPM / ELC day	intercept= 102000.5, coefficient= 4511.9, 5683.5 SE, t= 0.79, p= 0.43		As above for CPM.	▲	

		Pate et al. (2006) cut points	Step counts / ELC day	intercept= 2889.9, coefficient= 199.5, 89.8 SE, t= 2.22, p= 0.027		There was a positive association between settings with natural elements and step counts.	▲	
Gubbels et al (2018), Netherlands. E: 151 children / 22 ELC	Cross-sectional	SB, MVPA and CPM Actigraph GT3X+ Children were asked to wear the monitor for 7 consecutive days during their waking hours. Minimal wear time per day was 360 minutes and children had to have at least one valid ELC day to be included. Pate et al. (2006) cut points	Habitual SB %	β= −0.31, p < 0.001		Natural elements were significantly and positively associated with a reduction in percent time spent in SB	▲	Weak
			Habitual MVPA %	β= 0.27, p< 0.01		Natural elements were significantly and positively associated with an increased percent time spent in MVPA	▲	
			Habitual Mean CPM	β= 0.21, p< 0.01		Natural elements were significantly and positively associated with increased CPM.	▲	
Määttä et al (2019), Finland. E: 864 children / 66 ELC	Cross-sectional	Total PA Actigraph GT3X Worn for 7 days, 24-hours/day. A minimum wear time of 240 min during preschool hours was set. Evenson et al. (2008) cut points.	Total PA (min/hour in ELC)	Grass: β= 0.31, (95%CI: −0.84 - 1.46) Forest: β= −0.59, (95%CI: −1.87 - 0.69) Trees: β= --0.34, (95%CI: −2.13 - 1.45) Rocks:		There were no significant main or effect for grass, forest, trees or rocks	▲ ▼ ▼ ▲	Weak

				β= 0.01, (95%CI: −1.21 - 1.24)				
Määttä et al (2019b), Finland. E: 655 children / 66 ELC	Cross-sectional	Sedentary Time As above.	Sedentary time (min/hour in ELC)	Frequency of nature trips β= −1.026 (95%CI: −1.804, −0.248), p= 0.010		Frequency of nature trips was associated with children's lower sedentary time.	▲	Weak
Olesen et al (2013), Denmark. E: 441 children / 42 ELC	Cross-sectional	MVPA ActiGraph accelerometer Children wore the monitors for 1 week. Minimum wear time was 3 pre-school days, with at least 3 hours of measurement. Median wear-time was 4 weekdays, 7.15 hours per day. Evenson et al. (2008) cut points.	MVPA (percent/ ELC day)	Vegetation: - 0.7; 95% CI: -1.3 to -0.0, p= 0.04)		The multilevel analysis showed that the daily percentage of MVPA was significantly negatively associated with vegetation	▼	Weak
				Hilly landscape - 0.4; 95% CI: -1.1 to 0.2, p= 0.18.		The multilevel analysis showed that the daily percentage of MVPA was no association with hilly landscape.	▼	
Sugiyama et al (2012), Australia. E: 89 children / 10 ELC	Cross-sectional	MVPA and SB ActiGraph GT1M a Worn for 3 days at ELC. Minimum wear time was 2 days for at least 4 hours during the ELC day. Average wear time was 6 hours 40 minutes per ELC	MVPA (min/ outdoor time)	Mostly natural surface: β= −5.8, (95% CI: −9.9, −1.7), p<0.01		Children attending ELC's with mostly natural surfaces were found to engage in significantly less MVPA compared with ELC with mostly "built" surfaces.	▼	Weak
				More vegetation:		No association.	▼	

		day. Sirard et al. (2005) cut points.		β= −1.2, (95% CI: −5.9, 3.5)				
				Some gradient: β= 1.3, (95%CI: −4.5, 7.0)		As above.	▲	
				Much shade: β= 2.3, (95%CI: −3.5, 8.0)		As above.	▲	
			SB (min/ outdoor time)	Mostly natural surface: β= 8.0, (95% CI: −1.4, 17.4)		Natural surfaces, vegetation, gradient, and shade were not associated with SB.	▼	
				More vegetation: β= 2.3, (95% CI: −7.0, 11.6)			▼	
				Some gradient: β= −2.4, (95% CI: −13.7, 8.9)			▲	
				Much shade: β= −0.9, (95% CI:−12.6, 10.8)			▲	
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Observational								
Nature-based ELC								
Meyer et al (2017), Canada. E: 46 children / 3 ELC C: 35	Controlled cross-sectional	PA and PA types OSRAC-P Sampling Observation System which includes coding for body movements (stationary, slow-easy, moderate, and vigorous	PA frequencies: Stationary Slow-easy Moderate	 E:0.56 (0.15 SD) C: 0.84 (0.02 SD) E:0.30 (0.08 SD)		Children in the nature kindergarten were less stationary and engaged in more slow-easy and moderate physical activity compared to the control ELC.	N/A	Weak

children / 2 ELC		movements) and specific activity types (including climb, crawl, jump/skip, push/pull, rough and tumble, run, sit/squat, stand, throw, walk, and other). 2 students were observed at a time for 30-second intervals (5 sec observation, 25 sec coding). Observations occurred every 30 seconds for a period of 5 minutes which resulted in 20 observations. This was then repeated.	Vigorous	C: 0.16 (0.02 SD) E:0.12 (0.08 SD) C: 0 (0 SD)				
			PA types: (frequencies) Sit/Squat Walk Stand Fine Motor Eat Lie Down Push/Pull Rough & Tumble Run Climb	E: 0.02 (0 SD) C: 0 (0 SD) E: 0.19 (0.13 SD) C: 0.53 (0.09 SD) E: 0.17 (0.02 SD) C: 0.06 (0.01 SD) E: 0.14 (0.08 SD) C: 0.16 (0 SD) E: 0.14 (0.06 SD) C: 0.12 (0.09) E: 0.08 (0.03 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C:			N/A	

			Jump Throw Crawl Balance Other	E: 0 (0 SD) C: 0 (0 SD) E: 0.04 (0.02 SD) C: 0 (0 SD) E: 0.10 (0.07 SD) C: 0 (0 SD) E: 0 (0 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: 0 (0 SD) E: 0.01 (0.01 SD) C: 0.01 (0.01 SD) E: 0.05 (0.04 SD) C: 0.01 (0.01 SD) E: 0.05 (0.02 SD) C: 0.10 (0 SD)				
Wright (2019), USA. 48 children / 2 ELC	Cross-sectional	PA Children were observed and recorded over 2 school years. A randomised time sampling protocol was used with 10	overall frequency / relative frequency (% each type of activity was out of total instances of all PA)			“manipulation” was the most frequent PA type observed. balance, run, sit stand and squat were less frequent.	N/A	Weak

		min intervals at five zones. A sub-sample of the recordings was taken and coded at the 0:00, 1:00 and 2:00 mark for 20-second intervals. An adapted version of (OSRAC-P) was used to code the PA types.	Balance: Climb: Dig/Rake: Jump/Skip: Lie Down: Manipulation: Push/Pull: Resistive: Run: Sit: Stand: Squat: Throw: Walk:	34 / 7% 22 / 5% 19 / 4% 29 / 6% 9 / 2% 107 / 23% 21 / 4% 28 / 6% 34 / 7% 33 / 7% 38 / 8% 44 / 9% 16 / 3% 16 / 3%				
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Naturalised Playground								
Cosco et al (2014), USA. E: not clear / 27 ELC	Uncontrolled before & after	PA Children's Activity Rating Scale (CARS) CARS allows trained observers to record children's PA on a five-point scale: 1) stationary or motionless, 2) stationary with limb or trunk movements, 3) slow-easy, 4) moderate, and 5) fast.	PA		Unstandardised (standardised effects) 0.113 (0.067), p= 0.001	At post-intervention there was an effect on children's PA.	▲	Weak
			Non sedentary PA		0.202 (1.22), p= 0.001	As above for non-sedentary PA.	▲	
			MVPA		0.061 (1.063), Non-sig	Non-significant	▲	

Study details / Sample size	Study Design	Outcome and measurement	Units	Baseline or one time point (cross-sectional)	Follow-up (if applicable) or mean difference	Summary of Findings	Effect Direction	Quality Rating
Types of natural elements								
Sando (2019), Norway. E: 80 children / 8 ELC	Cross-sectional	PA Observational System for Recording PA in Children-Preschool (OSRAC-P) PA is coded from 1 (stationary) to 5 (fast movement). 2 children were filmed per day. The 1st for 2 minutes followed by a 6-minute break, then the 2nd child. Filming alternated between each child until 6 video observations of each child were recorded. 480 video clips in the outdoor environment constituted a full sample. There was a total of 471 video clips in the final analysis.	PA (1-5)	3.2 (0.9 SD), (regression coefficient= 0.004)		Nature was not a statistically significant predictor of PA.	▲	Weak

Sando & Sandseter (2019), Norway. E: 73 / 8 ELC	Cross sectional (mixed-methods)	PA and wellbeing (combined outcome) Wellbeing - Leuven Wellbing Scale measures wellbeing on a scale 1 (extremely low) -5 (extremely high). A score of 1 is when children exhibit high levels of discomfort (whining, screaming, sadness) and 5 is clear signs of happiness, relaxed and lively. Physical activity: see above, OSRAC-P which codes PA from 1 (stationary) to 5 (fast-movement)	PA and wellbeing	Nature: No association Sand: b =−0.027, (95% CI =−0.043–0.011), p= 0.001. Nature materials: b =−0.008, (95% CI =−0.015–0.001), p = 0.028. Water: no association Mud: no association		Nature is not associated with observations with high wellbeing and PA.	▼	Weak

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); PA= physical activity; MVPA= moderate to vigorous PA; SB= sedentary behaviour; CPM= counts per minute; Yr= Year; min = minutes; SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance (p<0.05)
▼: negative health impact and statistical significance (p<0.05)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 2. Nature-based ELC on motor skills								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled Before & After	Body function, gross motor skills and fine motor skills Kuno Beller Developmental Tables completed by educators which assesses development in 8 developmental areas: Body Function, Awareness of the Surrounding Environment, Social and Emotional Development, Play, Language, Cognitive Development, Gross and Fine Motor Skills.	Body Function	T1 (Jan 2014) E:11.02 (0.81 SD) C:10.15 (1.03 SD)	T4 (May 2015) 12.81 (0.71 SD) 12.39 (1.24 SD) p= 0.010; ηp^2= 0.27	There was a significant time x group interaction on children's body function. There were no significant differences between groups at T4.	▲	Weak
			Gross Motor Skills	E:11.79 (1.01 SD) C:10.87 (0.91 SD)	13.32 (0.80 SD) 12.96 (1.07 SD) p= 0.021; ηp2= 0.24	As above.	▲	
			Fine Motor Skills	E:10.86 (0.76 SD) C:10.01 (1.34 SD)	12.73 (0.88 SD) 12.56 (1.28 SD) p= 0.000; ηp2= 0.15.	As above.	▲	

Ene-Voiculescu & Ene-Voiculescu (2015), Fjortoft (2004), Fjortoft (2001), Norway. E: = 46 children / 1 kindergarten C: 29 children, / 2 kindergartens	Controlled Before & After	Motor fitness The EUROFIT Physical Fitness Test which consists of: flamingo balance test (standing on 1 foot - balancing); plate tapping (tapping of 2 plates alternatively- speed of limb movement); sit and reach (flexibility); standing broad jump (jumping for distance from a standing start – explosive strength); sit-ups (max n of sit-ups in 30 secs); bent arm hang (from a bar- functional strength); shuttle run (running and turning, shuttle - speed and agility) Beam walking to test dynamic balance and Indian skip (clapping right knee with left hand and vice versa - coordination), which were added.	Flamingo balance test / n of instabilities in 30 secs	E: 4.7 (0.8 SE) C: 4.0 (0.6 SE)	E: 1.5 (0.3 SE), p<0.001 C: 3.3 (0.7 SE)	At post-test, there were significant differences in the intervention group compared to the control group in the Flamingo balance test (p< 0.001).	▲	Weak
			Plate tapping / time in secs for 50 taps	E: 35.0 (1.9 SE) C: 29.9 (1.1 SE)	E: 28.1 (1.2 SE), p<0.001 C: 27.4 (2.6 SE)	No significant differences at post-test.	▼	
			Sit and reach / cm	E: 24.9 (0.8 SE) C: 25.3 (1.0 SE)	E: 24.4 (0.8 SE) C: 25.5 (0.9 SE)	As above.	▼	
			Standing broad jump / cm	E: 102.8 (2.9 SE) C: 103.1 (4.3 SE)	E: 113.1 (3.6 SE), p<0.001 C: 111.3 (3.8 SE), p<0.01	As above.	▲	
			Sit-ups / reps.30 secs	E: 5.3 (0.6 SE) C: 5.9 (0.8 SE)	E: 6.5 (0.6 SE) p<0.01 C: 7.0 (1.1 SE)	As above.	▼	
			Bent arm hang / sec	E: 2.6 (0.4 SE) C: 2.6 (0.6 SE)	C: 7.0 (1.0 SE), p<0.001 C: 5.4 (1.1 SE), p<0.001	As above.	▲	
			Beam walking / sec	E: 11.4 (1.4 SE) C: 7.7 (0.8)	E: 7.5 (0.7 SE), p<0.01 C: 7.2 (1.1 SD)	As above.	▼	
			Indian skip / reps.30 secs	E: 21.8 (2.2 SE)	E: 43.6 (1.9 SE), p<0.001	At post-test, there were significant differences in the	▲	

				C: 27.8 (2.4 SE)	C: 37.2 (1.8 SE), p<0.001	intervention group compared to the control group in the Indian skip co-ordination test (p< 0.01).		
			Shuttle run run/sec	E: 31.9 (0.7 SE) C: 30.7 (0.8 SE)	E: 29.7 (0.5 SE), p<.01 C: 30.3 (0.7 SE)	No significant differences at post-test.	▼	
Müller et al (2017), Canada. E: 43 children / 1 nature-kindergarten C: 45 children / 1 traditional kindergarten	Controlled before & after	Perceived physical competence, and locomotor and object control skills. Subscale of the Pictorial Scale of Perceived Competence and Social Acceptance for Young Children (six items) - children were asked to indicate who they are more like based on two descriptions of children (one competent and one not). Each item was scored on a four-point scale, where 4 indicates a high degree of perceived competence and 1 indicates a low score. TGMD-2 - assesses 6 locomotor and 6 object control skills.	Perceived Physical Competence	E: 18.72 (0.47 SE) C: 18.58 (0.44 SE)	E: 19.03 (0.48 SE) C: 19.47 (0.44 SE) p= 0.45, η2= 0.01	At post-test there was a small and non-significant effect	▼	Weak
			Locomotor skills	E: 24.68 (1.01 SE) C: 24.61 (0.94 SE)	E: 28.03 (0.82 SE) C: 25.72 (0.80 SE) p= 0.03, η2= 0.06	At post-test there was a moderate and significant effect	▲	
			Object control skills	E: 21.71 (0.98 SE) C: 23.05 (0.91 SE)	E: 23.97 (0.89 SE) C: 23.05 (0.91 SE) p= 0.15, η2= 0.03	At post-test there was a small and non-significant effect	▲	

		Scored either 1 or 0 depending on whether component was performed correctly.						
Lysklett et al (2019), Norway. E: 43 children / 4 preschools C: 49 children / 4 preschools	Controlled cross sectional	Motor competence Assessed using the Movement Assessment Battery (MABC-2). The test includes 8 subtests divided into 3 categories: 1) manual dexterity (posting coins, threading beads and drawing a line into a trail), 2) ball skills (catching beanbag and rolling ball into goal), and 3) static and dynamic balance (one-leg balance, walking heel raised and jumping over cord). Children are scored from 0-5. The total score sums the eight tests with a score of 0 the best and 40 the poorest.	Manual dexterity Ball Static and dynamic balance Total	E: 3.72 (2.99 SD) C: 3.29 (2.67 SD) E: 2.60 (2.34 SD) C: 2.41 (1.67 SD) E: 1.08 (1.71 SD) C: 0.94 (1.58 SD) E: 7.41 (4.91 SD) C: 6.64 (3.72 SD)	Mean difference 0.43 (95% CI: −0.74–1.59), p= 0.498 0.20 (95% CI: −0.64–1.03), p= 0.641 0.14 (95% CI: −0.53–0.82), p= 0.678 0.76 (95% CI: −1.03–2.56), p= 0.399	No significant differences in scores between the nature and traditional preschools for total and subtest scores.	▼	Weak
		The assessment for fitness consisted of 9 subtests: standing broad jump, Jumping on two feet, Jumping on one foot, Throwing a tennis	Standing broad jump (cm)	E: 94.78 (14.07 SD) C: 97.63 (15.59 SD)	Mean difference −2.86 (95% CI: −9.26–3.55), p= 0.378	Children attending the traditional preschools performed better in the shuttle run, reduced Cooper test and the total score	▼	

		ball (m), Putting a medicine ball, Climbing wall bars, Shuttle run, 20 m sprint, Reduced Cooper test. A total test score was calculated and transformed into z-scores (standardized scores).	Jumping on two feet (s)	E: 6.16 (3.58 SD) C: 5.18 (1.61 SD)	0.98 (95% CI: −0.22–2.18), p= 0.108	compared to the nature playground. The rest were non-significant.	▼	
			Jumping on one foot (s)	E: 5.48 (2.19 SD) C: 4.85 (1.19 SD)	0.63 (95% CI: −0.22–1.49), p= 0.144		▼	
			Throwing a tennis ball (m)	E: 6.00 (2.17 SD) C: 6.21 (1.88 SD)	−0.21 (95% CI: −1.06–0.64), p= 0.623		▼	
			Putting a medicine ball (m)	E: 1.88 (0.49 SD) C: 1.96 (0.43 SD)	−0.08 (95% CI: −0.27–0.11), p= 0.379		▼	
			Climbing wall bars (s)	E: 32.32 (14.60 SD) C: 31.21 (11.38 SD)	1.11 (95% CI:−4.37–6.59), p= 0.688		▼	
			Shuttle run (s)	E: 31.40 (3.96 SD) C: 30.00 (2.45 SD).	1.40, 95% CI: 0.05–2.74, p= 0.043		▼	
			20 m sprint (s)	E: 5.66 (0.48 SD) C: 5.53 (0.57 SD)	0.13 (95% CI: 0.13 – −0.08), p= 0.232		▼	
			Reduced Cooper test (m)	E: 740.09 (120.44 SD) C: 817.56 (105.32 SD)	77.47, 95% CI: −124.22–−30.71, p= 0.001),		▼	

			Total test score (z)	C: −0.12 (0.65 SD) E: 0.17 (0.57 SD)	0.29, 95% CI: −0.55– −0.04, p= 0.025		▼	
Scholz & Krombholz (2007), Germany E: 45 children / 10 forest kindergartens C: Rural = 42 children / 2 ELC; Urban = 42 children / 2 ELC	Controlled cross-sectional	Fundamental movement skills (test not described) Consisted of the following domains: balancing forward (balance); balancing backward (balance); jumping left and right; (coordination, speed); long jump; (coordination, speed); jumping forwards on one leg (coordination, endurance); hanging on pull up bar (strength endurance); shuttle run (speed, coordination)	Balancing forward (n of correct steps)	E:22.5 (1.7 SD) C (R): 20.5 (3.5 SD) C (U): 19.4 (3.6 SD) p<0.000		There was a significant higher performance in forest nurseries vs conventional rural and urban nurseries for balancing forwards and backwards, hanging on pull up bar, jumping left/right, shuttle run and one-leg jump forward on left.	▲	Weak
			Balancing backward (n of correct steps)	E: 51.5 (10.1 SD) C (R): 39.9 (10.9 SD) C (U): 35.5 (14.3 SD) p<0.000			▲	
			Jumping left and right (n of jumps)	E: 29.9 (6.0 SD) C (R): 31.1 (7.3 SD) C (U): 27.0 (7.1 SD) p=0.012			►	
			Long jump (distance in cm)	E: 94.0 (16.1 SD) C (R): 102.4 (18.4 SD) C (U): 94.0 (18.7 SD)			►	
			Hanging on pull up bar (time in seconds - max 30 sec)	E: 25.6 (6.2 SD) C (R): 20.7 (7.7 SD) C (U): 19.7 (7.0 SD) p<0.000			▲	
				E: 9.6 (1.2 SD)			▼	

			Shuttle run (time in seconds) Jumping forwards on one leg (n of jumps on each leg – max 20)	C (R): 9.1 (0.8 SD) C (U): 10.2 (1.5) p<0.000 Right: E: 17.5 (4.4 SD) C (R): 17.2 (4.9 SD) C (U): 16.0 (6.0 SD) Left: E: 17.8 (4.5) C (R): 16.8 (5.3), C (U): 14.1 (6.8) p=0.007			▲ ▲	
Ernst (2014), USA. E: 46 educators	Cross-sectional	Physical development Questionnaire (not described) on importance of natural outdoor settings on children's cognitive, social, and physical development and their appreciation for the environment. Responses were provided on a five-point scale, ranging from one (strongly disagree) to five (strongly agree)	Physical development (1-5)	4.39 (1.31 SD), r= 0.05		Educators agreed that experiences in natural settings were important for children's physical development. There was no association between frequency of nature experiences and belief regarding importance of outdoor settings for physical development.	▲	Weak

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals; cm= centimetres; sec= seconds; R= rural; U= urban

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 3. Nature-based ELC on weight status								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Types of natural elements								
Söderström at al (2013), Sweden. E: 172 children / 9 ELCs	Cross-sectional	BMI Weight = digital scale, height = measuring tape	BMI	Low OPEC Overweight= 16% Normal weight= 82% High OPEC Overweight= 7% Normal weight= 87% p= - 0.07		Outdoor environment quality was not significantly associated with BMI or waist.	▲	Weak
		Waist Measuring tape	Waist (cm)	Low OPEC: 52.6 (3.5 SD) High OPEC: 52.2 (3.5 SD) p= 0.25			▲	

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; BMI= body mass index; cm= centimetres; OPEC= outdoor Play Environmental Categories

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 4. Nature-based ELC on Sleep								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Choi et al (2014), South Korea. E: 18 children / 1 ELC C: 19 children / ELC	Controlled Before & After study	Sleep Parents competed the CSHQ which consists of 33 items with a 3 point scale, "usually (5–7 times a week)", "sometimes (2–4 times a week)", and "rarely (0–1 time a week)". This questionnaire consists of 8 domains: bedtime resistance, sleep onset delay, sleep duration, sleep anxiety, night wakings, parasomnia, sleep-disordered breathing, and daytime sleepiness. These domain scores are accumulated for a total CSHQ Score.	Total score of CSHQ Total sleep time (hours) Bedtime resistance Sleep onset delay Sleep duration Sleep anxiety	E: 51.6 ± 8.2 C: 55.6 ± 6.6 E: 10.5 ± 1.1 C: 10.7 ± 1.1 E: 11.8 ± 2.6 C: 12.7 ± 2.5 E: 1.3 ± 0.6 C: 1.2 ± 0.5 E: 3.7 ± 1.1 C: 4.1 ± 1.4 E: 7.1 ± 2.0	E: 47.7 ± 5.7, p= 0.02 C: 55.8 ± 6.5, p= 0.92 Between group: p < 0.01 E: 10.5 ± 1.0, p= 0.68 C: 10.4 ± 0.9, p= 0.21 E: 11.3 ± 2.4, p= 0.34 C: 12.8 ± 2.2, p= 0.98 E: 1.2 ± 0.4, p= 0.08 C: 1.4 ± 0.7, p= 0.36 E: 3.3 ± 0.6, p= 0.13 C: 3.7 ± 1.3, p= 0.37 E: 6.5 ± 2.0,	After post-test, the CSHQ total score, sleep disordered breathing and daytime sleepiness were significantly lower in children from the forest kindergarten program compared with the regular kindergarten program. There was no significant difference in total sleep time or other sub-scales.	▲ ▲ ▲ ▲ ▲ ▲	Moderate

		Total sleep time was also reported.		C: 7.4 ± 1.8	p= 0.28 C: 7.5 ± 1.5, p= 0.84			
			Night wakings	E: 3.6 ± 0.8 C: 3.6 ± 0.8	E: 3.5 ± 0.4, p= 0.71 C: 3.6 ± 1.0, p= 0.99		▲	
			Parasomnia	E: 9.2 ± 2.0 C: 10.0 ± 1.8	E: 8.6 ± 1.5, p= 0.11 C: 9.3 ± 1.9, p= 0.12		▲	
			Sleep disordered breathing	E: 3.3 ± 0.6 C:3.4 ± 0.8	E: 3.1 ± 0.5, p= 0.16 C: 3.7 ± 1.0, p= 0.10 Between group: p = 0.04		▲	
			Daytime sleepiness	E: 11.6 ± 2.5 C: 13.3 ± 2.9	E: 9.8 ± 1.0, p= 0.02 C: 13.7 ± 3.5, p= 0.52 Between group: p < 0.01		▲	
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Types of natural elements								
Söderström at al (2013), Sweden. E: 172 children / 9 ELC	Cross-sectional	Sleep A sleep diary was completed for one week by the children's parents. Parents recorded the time the children	Mean sleep time (minutes)	Low OPEC (n= 103): 642 (32 SD) High OPEC (n= 66): 658 (44 SD)		Outdoor environment quality was significantly associated with night sleep	▲	Weak

		woke up and the time they went to sleep. Sleep time was calculated as a mean of the seven days.		p= 0.03				

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); CSHQ= Children's Sleep Habits Questionnaire; OPEC= outdoor Play Environmental Categories

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 5. Nature-based ELC on UV Exposure								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Types of natural elements								
Boldemann et al (2004), Sweden. E: 64 children / 2 ELC	Cross-sectional	UV Exposure Measured using a Dosimeter (Biosense VioSpor blue line, type III 0.8–33 MED). Each child wore 2 Dosimeters attached to each shoulder using safety pins. They were worn during the school day.	UV exposure per day (JCIE/m^2)	Site 1: 222 JCIE/m^2, 15.3 % (95% CI 14.3–17.5, $p<0.05$) Site 2: 175 JCIE/m^2, 13.3 % (95% CI 9.9–14.6, $p<0.05$)		The was a statistically significant difference in UVR exposure between site 1 and site 2.	▲	Weak
Boldemann et al (2006), Sweden. E: 199 children / 11 ELC	Cross-sectional	UV Exposure Measured using a Polysulphone dosimeter (Diffey, 1984; Herlihy et al., 1994) The Dosimeter was pinned to the right shoulder and worn during school hours.	UV Exposure (J/m^2)	Low environment: ELC 3: 160 (95%CI:130–190) ELC 4: 241 (95%CI:200–281) ELC 6: 156 (95%CI:115–196) ELC 7: 83 (95%CI: 67–98) ELC 8: 269	Daily UV exposures ranged between 74 and 292 J/m	Outdoor environment quality was significantly associated with UV Exposure.	▲	Weak

				(95%CI:214–324) ELC 10: 243 (95%CI:217–268) High environment: ELC 1: 104 (95%CI: 95–113) ELC 2: 129 (95%CI:104–154) ELC 5: 289 (95%CI:230–348) ELC 9: 292 (95%CI:232–351) ELC 11: 196 95%CI: 177–215)				
Christian et al (2019), Australia. E: 678 children / 48 ELC	Cross-sectional	UV Exposure Measured using a Polysulphone film mounted cardboard holders (UV badge) The UV badge was attached to the child's left shoulder and worn each day whilst at ELC for up to 3 days.	UV exposure (J/m^2) per average day of ELC.	% <3 m vegetation: β = -2.26 (95%CI -3.03, -1.49); p <0.01 % >3m vegetation: β = 0.91 (95%CI -12.46, 14.28), p= 0.89		ELC centre vegetation was significantly negatively associated with children's UVR exposure. For every 1% increase in centre vegetation, children's UVR exposure decreased by 2.3 J/m2 per day at ELC (p <0.01).	▲	Weak

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings

▼: negative health impact
▲: positive health impact and statistical significance (p<0.05)
▼: negative health impact and statistical significance (p<0.05)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 6. Nature-based ELC on harms								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Frenkel et al (2019), USA. E: 71 children / 5 ELC C: 70 children / 4 ELC	Controlled cross-sectional	Illness and injury Educators completed a standardised weekly illness and injury tracking log developed for this study. An illness episode was when a child was absent for at least 1 day due to illness (fever, respiratory, stomach, other).	Illness total Fever Respiratory Stomach Other	E: 1.49 C: 1.62 (age adjusted IRR: 0.93, 95% CI: 0.64, 1.34). E: 0.25 C: 0.47 E: 0.92 C: 1.01 E: 0.29 C: 0.37 E: 0.18 C: 0.07		No significant difference in the incidence of total illness between nature ELC and traditional ELC	▲	Weak
		An injury was counted if it required first-aid attention from teachers	Total injury	E: boys= 0.94 girls= 1.87 C: boys= 0.96 girls= 0.34		No significant difference in minor injury was found between boys at nature and traditional ELC. Girls at nature ELC had a significantly higher incidence of minor	▲ (boys) ▼ (girls)	

				boys: (age-adjusted IRR: 1.46, 95% CI: 0.59, 3.6) Girls: (age-adjusted IRR: 5.91, 95% CI: 1.98, 17.7).		injury compared with girls at traditional ELC.		
			Open wound/cut	E: boys= 0.60 girls= 1.31 C: boys= 0.48 girls= 0.23				
			Sprain	E: boys= 0 girls= 0 C: boys= 0 girls= 0				
			Child Bite	E: boys= 0.17 girls= 0 C: boys= 0 girls= 0				
			Other	E: boys= 0.17 girls= 0.56 C: boys= 0.48 girls= 0.11				
Moen et al (2007), Norway.	Controlled cross-sectional	Sickness absenteeism	Sickness absenteeism	estimate =		No statistically significant difference in sickness	▼	Weak

E: 267 children / 37 ELC C: 264 children / 32 ELC		Parent noted daily reports of sickness absenteeism Absenteeism refers to the ratio of the total number of sickness absenteeism days to the sum of the number of sickness absenteeism days and the number of days the child was attending the day care centre during the study period.		- 0.0083, SE= 0.1830, t= 20.045, p> 0.05		absenteeism between the outdoor ELC and regular day ELC.		
Weisshaar et al (2006) E: 506 children / 25 ELC C: 1201 children / 28 ELC	Controlled cross-sectional	Tick bites and borreliosis Self- report questionnaire. Presence of at least 1 tick bite (yes/no). Presence of borreliosis (yes/no)	Tick bite % (presence – yes/no) Risk	Yes: E: 73.2% C: 26.6% No: E: 26.8% C: 73.4% p=0.0001 Adj OR= 6.74, 95% CI: 5.29–8.60		Children attending forest kindergartens reported a significantly higher prevalence of tick bites compared to the traditional kindergartens. Attending a forest kindergarten was a risk factor for having at least one tick bite when adjusting for age, sex, skin inspection and recommended vaccination.	▼	Weak
			Borreliosis % (presence – yes/no)	Yes: E: 2.0% C:0.4%		As above	▼	

			Risk	No: E: 98.0% C: 99.6% (p= 0.004) Adj OR= 4.61, 95% CI: 1.50–14.17				
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Types of natural elements								
Söderström at al (2013), Sweden. E: 172 children / 9 ELC	Cross-sectional	Symptoms (illness) The sum of days with symptoms of illness (runny nose, cough, fever, respiratory problems/asthma, itchy skin, diarrhoea, stomach ache, ear pain, body ache, sticky eyes, any medicine taken and days where parents had worries for their child). High score = less healthy.		p= 0.12 (descriptive statistics not presented)		Outdoor environment quality was not significantly associated with symptoms	N/A	Weak

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance (p<0.05)

▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

COGNITIVE

Table 7. Nature-based ELC on cognitive outcomes								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled Before & After	See Table 2.	Language	T1 (Jan 2014) E:11.01 (1.30 SD) C:9.83 (1.53 SD)	T4 (May 2015) 12.88 (1.03 SD) 12.74 (1.24 SD) p= 0.000; ηp^2= 0.42	There was a significant time x group interaction on children's language. There were no significant differences between groups at T4.	▲	Weak
			Cognitive development	E:10.94 (0.89 SD) C:9.63 (1.35 SD)	12.49 (0.95 SD) 12.58 (1.31 SD) p= 0.000; ηp2= 0.51.	As above	▼	
Cooper (2018), United Kingdom (England). E: 13 children	Controlled before & after study	Communication Assessed using FOCUS-34 (Focus on the Outcomes of Communication Under Six) which evaluates communication development.	Communication (median and range)	E: 206 (73) C: 214 (93)	206 (73), Z=2.49 p=0.0013 214 (93), Z=2.85 p=0.004 U=54.5 p=0.694	No significant between-group differences at T2	▼	Weak

C: 11 children Children from the same school		FOCUS -34 is divided into 2 sections (34 items in total) and scored on a 7-point Likert scale.						
		The Devereux Early Childhood Assessment for Pre-schoolers, Second Edition (DECA-P2) consists of 38 items on a 5-point likert scale. The assessment measures protective factors and screen for behavioural concerns. The protective factors are divided into 3 subscales: initiative **self-regulation** and attachment/ relationships which form an overall measure of social and emotional wellbeing when combined. Parent and teachers completed the form and they were asked to reflect on the child's behaviour for the previous 2 weeks.	Self-regulation (median and range)	E: 24 (22) C: 23 (19)	25 (20); Z=1.48 p=0.138 24 (18); Z=1.63 p=0.102 **U=56.0 p=0.767**	No statistically significant between-group differences at T2 for self-regulation, initiative	▲	

Cordiano et al (2019), USA. E: 12 children / 1 ELC class. C: 14 children / 1 class. Children from the same school.	Controlled before & after study	Kindergarten readiness Tool assessed letter number recognition, sorting and classifying information, counting, rhyming, and recognizing one's name. The skills were rated by the teachers as "Never," "Sometimes," "Often," or "Always".	Kindergarten readiness	T1 - baseline E:19.09 (3.86 SD) C:23.42 (3.44 SD)	T3 - endpoint 24.72 (2.87 SD) 26.79 (1.71 SD) Within group: η^2p= 0 .10 (small effect), p>0.05 Between group: F= 4.05, η2p= 0.16, p> 0.05.	Non-significant and moderate effect for between group differences.	▼	Weak
Ernst & Burcak (2019), USA E: 34 children / 2 ELC C: 43 children / 2 ELC	Controlled Before & After study	Curiosity Curiosity Drawer Box task - There are a total of 12 possible points (1 point per drawer) for each of these three dependent measures (toys out, toys explored, toys engaged with further), with higher numerical scores indicating higher levels of the respective forms of curiosity. If a child returns to a drawer or toy after having already opened that drawer or interacted with that toy, they do not receive additional points.	Toys Taken Out: Toys Explored: Toys Engaged With:	E: 8.38 (3.39 SD) C: 7.81 (4.19 SD) E: 6.44 (3.09 SD) C: 3.50 (2.71 SD) E: 4.15 (2.60 SD) C: 4.23 (2.89 SD)	Adj post-test (mean and SE) 9.61 (0.46 SE) 8.85 (0.40 SE) p = 0.21, ηp2 = 0.02 6.05 (0.66 SE) 6.24 (0.57 SE) p = 0.83 ηp2 < 0.01 7.61 (0.48 SE) 5.92 (0.42 SE) p = 0.01 ηp2 = 0.09	At post-test, there were no significant differences between the nature and non-nature groups for toys taken out or toys explored, toys engaged with was significant. (controlled for pre-test, age, gender, and prior participation)	▲ ▼ ▲	Weak

Burgess & Ernst (2020). E: 84 children / 4 ELC C: 24 children / 2 ELC		Learning behaviours Preschool learning behaviours scale which consists of 24 items with 3 dimensions: competence motivation; attention/ persistence and attitudes. Teachers score on a 3-point Likert scale (doesn't apply, sometimes, apply, most often applies)	Adj means (SE) Competence motivation Attention/ persistence Attitudes Total	 E:16.73 (0.45 SE) C:19.53 (0.83 SE) E:13.18 (0.37 SE) C:+ SE) E:11.11 (0.28 SE) C:11.77 (0.39 SE) E:36.53 (0.83 SE) C:41.77 (1.51 SE)	 E:20.41 (0.33 SE) C:18.66 (0.65 SE) p=0.02, n2=0.05 E:16.66 (0.30 SE) C:16.13 (0.59 SE) p=0.41, n2=0.01 E:12.74 (0.22 SE) C:12.22 (0.42 SE) p=0.27, n2=0.01 E:44.16 (0.68 SE) C:41.76 (1.34 SE) p=0.12, n2=0.02	At post-test, the nature ELC had significantly higher competence motivation compared to the non-nature ELC. (adjusted for pre-test levels, age, gender, prior participation, and part v. full-time participation)	▲ ▲ ▲ ▲	
Zamzow & Ernst (2020). E: 78 / 4 ELC C: 44 children / 2 ELC	Controlled Before & After study	Executive functions Minnesota Executive Function Scale (MEFS) - conducted using an App, children perform a game like activity where they sort cards to boxes. This games changes commands to assess cognitive flexibility, inhibitory control, and	 Executive functions	 E:41.78 (14.89 SD) C:38.54 (14.40 SD)	Adj post-test (mean and SE) 50.86 (1.29 SE) 49.72 (1.73 SE) p= 0.60, ηp2 < 0.01	No significant differences between the nature and non-nature groups when controlling for pre-test, age, gender, and prior participation.	▲	

		working memory and provides an executive function total score.						
Wojciehowski & Ernst (2018). E: 75 children / 4 ELC Ernst et al (2019). E: 78 children / 4 ELC	Uncontrolled Before & After study	Creative thinking Thinking Creatively in Action and Movement (TCAM) consists of four activities that measure fluency, originality, and imagination.	Fluency Originality Imagination	E: 89.89 (17.76 SD) E: 96.13 (20.16 SD) E: 89.85 (17.68 SD)	104.76 (28.35 SD), p < 0.001 113.61 (36.58 SD), p< 0.001 99.99 (18.42 SD), p< 0.001	Significant improvements in fluency, originality, and imagination in the nature preschool from baseline to follow-up.	▲	
		Resilience Devereux Early Childhood Assessment for Preschoolers, Second Edition (DECAP2) - Parents and teachers evaluate 27 positive behaviors, which form 3 subscales: initiative, **self-regulation**, and attachment. Three subscales were converted to standard scores (T-scores) with a mean of 50 and SD of 10.	Teacher: Self-regulation:	E:54.49 (6.00 SD)	56.78 (8.05 SD), p= 0.01	Significant improvements in self-regulation scores in the nature preschool from baseline to follow-up.	▲	
			Parent Self-regulation:	E:49.31 (7.98 SD)	53.35 (9.34 SD), p= 0.01	Significant improvements in self-regulation in the nature preschool from baseline to follow-up.	▲	
Müller et al (2017), Canada.	Controlled before & after study	Executive functions Working memory: the boxes task is a touch-screen	Working memory	E:25.38 (1.25 SE) C:26.69 (1.18 SE)	E:20.85 (1.91 SE) C:24.84 (1.87 SE)	At post-test there was a small and non-significant effect for working memory and	▲	Weak

E: 43 children / 1 nature-kindergarten C: 45 children / 1 traditional kindergarten		operated, self-ordered search task designed to measure working memory. Attention: Continuous Performance Test (CPT)- a computer based task that requires children to respond to stimuli by touching an animal on the touchscreen and to refrain from responding to a number of other stimuli types. The task lasted 5 minutes and included 200 stimulus of which 29 were targets. The number of correctly identified targets was used as performance indicator of directed attention. Inhibition: The Head-Shoulders-Knees-Toes task (HSKT) - a task that involved children listening to commands and performing the opposite (e.g. touching head when researcher instructed them to touch their feet). Children were	Attention Inhibition	E:22.67 (0.92 SE) C:23.87 (0.86 SE) E:28.96 (3.24 SE) C:27.83 (3.16 SE)	p= 0.19, η2= 0.02) 23.70 (1.01 SE) 24.98 (0.94 SE) p= 0.51, η2= 0.01 34.73 (2.34 SE) 33.44 (2.29 SE) p= 0.76, η2= 0.00	attention. No effect for inhibition.	▼ ▲	

		given a score out of 40.						
		Social Skills Rating Scale (SSRS) completed by parents and teachers. This assesses the following social skills: cooperation, assertiveness, social responsibility and **self-control** and items assessing psychological health (internalising and externalising behaviour). Questionnaires were completed by teachers and parents. They were asked to indicate how often a behavior occurred (never, sometimes, very often).	Teacher Self-control Parent Self-control	 E:16.12 (0.56 SE) C:14.71 (0.55 SE) E:14.75 (0.54 SE) C:14.68 (0.70 SE)	 18.10 (0.56 SE) 13.52 (0.55 SE) p= 0.00, η2= 0.32 15.78 (0.53 SE) 15.00 (0.69 SE) p= 0.29, η2= 0.02	 At post-test there was a large and significant effect. At post-test there was a small and non-significant effect.	 ▲ ▲	
Fyfe-Johnson et al (2019), USA. E: 20 children / 1 ELC C: 13 children (waitlist control or 2-hour nature-	Controlled cross-sectional	Child behaviour SDQ: 25-items consisting of 5 domains: emotional problems, conduct problems, **hyperactivity/ inattention**, peer relationship problems, and prosocial behavior. Parents rated their child on a scale of 0	Hyperactivity/ inattention	E: 2.74 (2.27 SD), C: 3.58 (2.27 SD)	Mean diff -0.88 (95% CI: -2.71, 0.94)	Children in the nature ELC did not differ compared to the control.	▲	Weak

based, outdoor enrichment class provided by experimental ELC		to 2 per question (0=not true; 1=somewhat true; 2=certainly true). Overall score was calculated (sum of all domain scores except prosocial behavior; overall score range: 0-40). Prosocial was scored separately.						
Ernst (2014), USA. E: 46 educators	Cross-sectional	Cognitive development See Table 2.	Cognitive development (1-5)	4.33 (1.30 SD), r= 0.05		There was no association between frequency of nature experiences and belief regarding importance of outdoor settings for cognitive development.	▲	Weak
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Naturalised Playgrounds								
Carrus (2012), Italy. E: 16 children / 1 ELC	Cross-sectional	Visual spatial task (indicator of children's direct attention) Children were asked to colour or to glue paper on to a drawing provided. Performances were evaluated by two independent coders.	Visual spatial task	No inferential stats provided.		Children exposed to free play in external green spaces exhibited a higher accuracy in the performance of the visual-spatial tasks compared to the control.	N/A	Weak

Study details / Sample size	Study Design	Outcome and measurement	Units	Baseline or one time point (cross-sectional)	Follow-up (if applicable) or mean difference	Summary of Findings	Effect Direction	Quality Rating
Types of natural elements								
Martensson et al (2009), Sweden. E: 198 children / 11 ELC	Cross-sectional	Attention The Early Childhood Attention Deficit Disorders Evaluation Scale (ECADDES, School) consists of 2 domains: inattention (32 items) and hyperactivity/ impulsivity (24 items) which are rated by two members of staff who observe the children in their daily routines. Each item is rated from 0-4 (0= child does not engage in the behavior at all, 1= behavior occurs one to several times per month, 2= behavior occurs one to several times per week, 3= behavior occurs one to several times per day, and 4= behavior occurs one to several times per hour) with a lower score indicating a lower occurrence. Rating are summed per child and raw scores	Hyperactivity/ impulsivity Inattention	OPEC: Low Score= 1.59; High Score= 1.23, F= (-) 4.25, p= 0.069 OPEC: Low Score= 1.87; High Score= 1.46, F= (-) 7.38, p<.05		OPEC was significantly related to inattention dimension only:	▲ ▲	Weak

		converted into standard scores taking into account sex and age.						
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Garden-based intervention								
Park et al (2016), South Korea. E: 336 children /12 ELC Science investigation abilities and attitudes= 68 children	Uncontrolled before & after	Scientific attitudes The Scientific Attitude Survey revised by Lee (2000) was used. This consists of 27 questions on a five-point likert scale (strongly agree - strongly disagree) with 9 subcategories: curiosity, volunteerism and activeness, forthrightness, objectivity, openness, criticism, objectivity, cooperation, and patience. Teachers completed this questionnaire based on their daily observations. Higher scores indicate better scientific attitude.	Scientific attitudes (1-5) Curiosity Activeness Forthrightness Objectivity Openness Criticism Judgement reservation Cooperation Patience	 3.17 ± 0.98 3.13 ± 0.95 3.31 ± 0.77 3.07 ± 0.72 2.98 ± 0.64 2.79 ± 0.69 2.72 ± 0.74 3.13 ± 0.67 2.57 ± 0.77	 4.11 ± 0.67, p=0.000 4.10 ± 0.65, p=0.000 4.07 ± 0.54, p=0.000 3.88 ± 0.69, p=0.000 3.55 ± 0.58, p=0.000 3.46 ± 0.59, p=0.000 3.42 ± 0.70, p=0.000 3.94 ± 0.65, p=0.000 3.77 ± 0.89, p=0.000	There were significant improvements in Science attitudes subcategories from baseline to follow-up.	▲	Weak

		Scientific investigations ability of younger children questionnaire revised by Lee (2000) was used. This consists of 21 questions on a five-point likert scale (strongly agree - strongly disagree) with 5 subcategories: prediction, observation, classification, measurement, and discussion. A higher score indicates better investigation ability.	Scientific investigation abilities (1-5) Prediction Observation Classification Measurement Discussion	 3.11 ± 0.83 3.34 ± 0.92 3.25 ± 0.93 2.88 ± 0.97 3.04 ± 0.85	 3.54 ± 0.63, p=0.002 3.99 ± 0.67, p=0.000 3.93 ± 0.66, p=0.000 3.70 ± 0.68, p=0.000 3.55 ± 0.81, p=0.001	As above.	▲	
Lillard (2016), USA. E: 55 children / 1 ELC Delay Gratification E: 34 children Visual motor integration E: 39 children	Uncontrolled before & after	Delay Gratification Participants were assessed individually. The researcher followed a script which involved the child receiving a treat if they waited for the researcher to complete a task. If they wanted the treat immediately, they could ring a bell for the researcher to come back but would get a smaller treat. Measurement was in seconds from when they rang the bell, or	Delay Gratification (seconds)	426.15	676.18, Non-sig	There was not a significant improvement from baseline to follow-up	▲	Weak

		they reached 15 minutes.						
		Visual Motor Integration Assessed using the Beery-Buktenica Developmental Test of Visual-Motor Integration 5th Edition (short form). This was a short pencil and paper test in which participants copy a sequence of shapes. Raw scores ranged from 0-20 and were transformed to standardized scores. Standard scores I have a mean of 100 (15 SD). Scores are age specific.	Visual Motor Integration (scores)	98.62	100.37, non-sig	As above		

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals; OPEC= Outdoor Play Environment Categories.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Social, Emotional and Environmental

Table 8. Nature-based ELC on social and emotional outcomes								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled Before & After study	See Table 2.	Social and emotional development	T1 (Jan 2014) E:11.18 (1.09 SD) C:10.24 (1.14 SD)	T4 (May 2015) 12.96 (0.94 SD) 12.86 (0.94 SD) p= 0.000; ηp2= 0.38	There was significant time x group interaction on children's social and emotional development. There were no significant differences between groups at T4.	▲	Weak
Cooper (2018), United Kingdom (England). E: 13 children C: 11 children	Controlled before & after study	The Devereux Early Childhood Assessment for Pre-schoolers, Second Edition (DECA-P2) consists of 38 items on a 5-point likert scale. The assessment measures protective factors and screen for behavioural concerns. The protective factors are	Attachment / relationships Self-regulation	E: 23 (13) C: 25 (15) Presented in cognitive domain E: 21 (14)	27 (11); Z=2.82 p=0.005 31 (17); Z=2.61 p=0.009 **U=32.0 p=0.058** Presented in cognitive domain 26 (13); Z=2.41	No statistically significant between-group differences at T2 for attachment/ relationships, initiative, and social and emotional wellbeing	▼ ▼	Weak

Children from the same school		divided into 3 subscales: initiative, self-regulation and attachment/ relationships which form an overall measure of social and emotional wellbeing when combined. Parent and teachers completed the form and they were asked to reflect on the child's behaviour for the previous 2 weeks.	Initiative Social and emotional wellbeing (median and range)	 C: 20 (12) E: 69 (40) C: 71 (39)	p=0.016 29 (16); 2.63 p=0.009 **U=40.5 p=0.187** 76 (32); Z=2.49 p=0.013 83 (48); Z=2.49 p=0.013 **U=42.0 p=0.224**		▼	
Cordiano et al (2019), USA. E: 12 children / 1 ELC class. C: 14 children / 1 class. Children from the same school.	Controlled before & after	Preschool and Kindergarten Behavior Scales, Second Edition (PKBS-2) is a 76-item behavior rating instrument which assesses social skills and behavioural problems. The Social Skills scale assess the dimensions of Social Cooperation, Social Interaction, and Social Independence. The Problem Behavior scale assesses the dimensions of Externalizing Problems and	Social skills Teacher Parent	T1 - baseline E: 101.92 (11.69 SD) C: 110.07 (7.41 SD) E: 102.20 (15.51 SD); C: 104.00 (7.29 SD)	T3 - endpoint 106.21 (13.34 SD) 112.96 (6.29 SD) Within-group: p= non-sig, η2p= 0.01 Between group: F=1.98, η2p= 0.08, p> 0.05 108.40 (12.67 SD) 128.73 (64.96 SD) Within-group: p= non-sig, η2p= 0 .08 Between group:	Small effect for between group Small effect for between group	▼ ▼	Weak

		Internalizing Problems			F= 0.87, η2p= 0.05, p> 0.05			
			Behavioural problems Teacher Parent	 E: 91.58 (9.14 SD) C: 82.46 (6.39 SD) E: 97.00 (21.12 SD) C: 101.10 (13.16 SD)	 89.96 (12.26 SD) 83.93 (5.03 SD) Within-group: p= non-sig, η2p= 0.01 Between group: F=4.81, η2p= 0.17, p<0.05 92.67 (16.52 SD) 95.20 (9.94 SD); Within-group: p= non-sig, η2p= 0.21 Between group: F= 0.15, η2p= 0.01, p>0.05	 Moderate effect for between group No effect for between group	 ▼ ▲	
Müller et al (2017), Canada. E: 43 children / 1 nature-kindergarten C: 45 children / 1 traditional	Controlled before & after study	Social Skills Rating Scale (SSRS) completed by parents and teachers. This assesses the following social skills: cooperation, assertiveness, social responsibility and self-control and items assessing	Teachers Assertiveness Cooperation	 E:17.15 (0.57 SE) C:12.40 (0.55 SE) E:17.14 (0.52 SE) C:15.00 (0.49 SE)	 19.16 (0.47 SE) 12.86 (0.45 SE) p= 0.00, η2= 0.34 18.63 (0.45 SE) 15.25 (0.43 SE) p= 0.00 η2= 0.20	 At post-test there was a large and significant effect. As above.	 ▲ ▲	Weak

kindergarten		psychological health (internalising and externalising behaviour). Questionnaires were completed by teachers and parents. They were asked to indicate how often a behavior occurred (never, sometimes, very often).	Self-control Externalizing Behavior: Internalizing Behavior	Presented in cognitive domain. E: 2.63 (0.48 SE) C: 1.91 (0.47 SE) E: 0.96 (0.16 SE) C: 0.36 (0.15 SE)	Presented in cognitive domain. 2.05 (0.43 SE) 1.98 (0.41 SE) p= 0.11, η2= 0.03 0.20 (0.11 SE) 0.41 (0.10 SE) p= 0.04, η2= 0.05	At post-test there was a small and non-significant effect. At post-test there was a small and significant effect.	▼ ▲ ▲	
			Parent Assertiveness Social Responsibility Cooperation Self-control Externalizing Behavior: Internalizing Behavior	 E:15.27 (0.43 SE) C:15.31 (0.62 SE) E:11.58 (0.48 SE) C:10.50 (0.67 SE) E:12.76 (0.37 SE) C:12.00 (0.52 SE) Presented in cognitive domain. E: 3.67 (0.38 SE) C: 3.79 (0.50 SE)	 16.24 (0.42 SE) 14.75 (0.60 SE) p= 0.01, η2= 0.13 13.10 (0.44 SE) 11.06 (0.61 SE) p= 0.03, η2= 0.11 13.18 (0.36 SE) 11.75 (0.52 SE) p= 0.06, η2= 0.08 Presented in cognitive domain. 3.06 (0.36 SE) 3.63 (0.47 SE) p= 0.25, η2= 0.03 0.94 (0.17 SE)	At post-test there was a moderate and significant effect. As above. At post-test there was a moderate but non-significant effect. As above. At post-test there	▲ ▲ ▲ ▲ ▼	

				E: 1.17 (0.17 SE) C: 0.79 (0.23 SE)	0.90 (0.23 SE) p= 0.68, η2= 0.00	was a non-significant effect.		
Ernst et al (2019) & Ernst & Burcak (2019), USA E: 78 children / 4 ELC	Uncontrolled Before & After study	Resilience Devereux Early Childhood Assessment for Preschoolers, Second Edition (DECAP2) - Parents and teachers evaluate 27 positive behaviors, which form 3 subscales: initiative, self-regulation, and attachment. Three subscales were converted to standard scores (T-scores) with a mean of 50 and SD of 10.	Teacher: Total protective factors Initiative: Self-regulation: Attachment:	 E:54.54 (5.95 SD) E:52.74 (7.98 SD) presented in cognitive domain. E:55.26 (6.91 SD)	 57.71 (7.87 SD), p=0.01 56.93 (8.55 SD), p= 0.01 presented in cognitive domain. 57.21 (7.45 SD)	 Significant improvements in total protective factors and initiative in the nature preschool from baseline to follow-up. No significant improvements in attachment scores.	 ▲ ▲ ▲	Weak
			Parent Total protective factors Initiative Self-regulation: Attachment:	 E:50.21 (7.62 SD) E:49.84 (8.45 SD) presented in cognitive domain. E:51.64 (7.24 SD)	 53.13 (8.81 SD), p = 0.01 53.63 (8.17 SD), p= 0.01 presented in cognitive domain. 51.39 (9.93 SD)	 Significant improvements in in the total protective factors, and initiative in the nature preschool from baseline to follow-up. No significant improvements in attachment scores.	 ▲ ▲ ▲	

Fyfe-Johnson et al (2019), USA. E: 20 children / 1 ELC C: 13 children (waitlist control or 2-hour nature-based, outdoor enrichment class provided by experimental ELC	Controlled cross-sectional	Child behaviour SDQ: 25-items consisting of 5 domains: emotional problems, conduct problems, hyperactivity/ inattention, peer relationship problems, and prosocial behavior. Parents rated their child on a scale of 0 to 2 per question (0=not true; 1=somewhat true; 2=certainly true). Overall score was calculated (sum of all domain scores except prosocial behavior; overall score range: 0-40). Prosocial was scored separately.	Overall Score Emotional problems Conduct problems Hyperactivity/ inattention Peer relationship problems Prosocial behavior	E: 6.55 (4.35 SD) C: 7.51 (4.23 SD) E: 1.20 (1.67 SD) C: 1.00 (0.95 SD) E: 1.63 (1.54 SD) C: 1.83 (1.59 SD) presented in cognitive domain. E: 1.05 (0.94 SD), C: 1.08 (1.24 SD) E: 8.15 (1.57 SD), C: 7.83 (1.59 SD)	Mean diff -0.95 (95% CI: -4.39, 2.49) 0.2 (95% CI: -0.82, 1.22) -0.23 (95% CI: -1.49, 1.03) presented in cognitive domain. -0.03 (95% CI: -0.95, 0.88) 0.32 (95% CI: -0.95, 1.59)	Children in the nature ELC did not differ in behavioural scores compared to the control.	▲ ▼ ▲ ▲ ▲	Weak
Ernst (2014), USA. E: 46 educators	Cross-sectional	See table 2.	Social development (1-5)	4.43 (1.31 SD), r= 0.05		There was no association between frequency of nature experiences and belief regarding importance of outdoor settings for social development.		Weak

Study details / Sample size	Study Design	Outcome and measurement	Units	Baseline or one time point (cross-sectional)	Follow-up (if applicable) or mean difference	Summary of Findings	Effect Direction	Quality Rating
Naturalised Playground								
Brussoni et al (2017), Canada. E: 48 children / 2 childcare centres	Uncontrolled before & after (mixed methods)	Sociometric status was determined by rating how "dominant or influential" and "popular" each child is with peers	Sociometric: Dominance Acceptance		 Centre A= 3.42 Centre B= 2.70 Centre A= 3.44 Centre B= 3.25	Mean sociometric scores remained stable over time.	 ▲ ▲	Moderate
		Strengths and difficulties questionnaire (SDQ)-25 items that measure emotional symptoms, conduct problems, hyperactivity, peer relationships, and prosocial behaviour.	Strengths and difficulties (median)	2.3	2.0; z= -2.10, p= 0.036	There was a significant decrease in the SDQ peer problems scale. No other scores differed significantly (not reported).	▲	
		Preschool social behaviours skill (PSBS-T) - 19 items assessing relational aggression, overt aggression, depressed affect.	Social behaviour (median)	6.0	3.0 z= -2.24, p= 0.03	There was significant decrease in The PSBS depression score. No other scores differed significantly (not reported).	▲	
Cosco et al (2014), USA. E: not clear / 27 centres.	Uncontrolled Before & After study	Social interactions Observational behaviour mapping was conducted. Location of children, gender, PA level, social interactions	Custodial (i.e tying shoe laces, offering water) teacher-child interaction		-0.156, B=-0.095), p< 0.05	At follow- up, observations highlighted significantly less custodial teacher-child interactions, more negative teacher-child	▼	Weak

		(alone, pair, group), teacher interactions (not present, positive, custodial, negative) were recorded by observers and entered into a handheld computer.	Negative teacher-child interaction No teacher present Positive teacher-child interaction Child is alone Child is with one other child Child is in group		0.030, B= - 034, p< 0.05 0.082, B= - 0.002, non-sig - 0.064, B= - 0.088, p< 0.05 - 0.195, B= not estimated - 0.034, B= - 0.031, p< 0.05 - 0.168, B= - 0.113, p< 0.05	interactions, less positive teacher-child interactions and less children with another child or in a group:	▼ ▼ ▼ ▼ ▼	
Carrus (2012), Italy. E: 16 children / 1 ELC	Cross-sectional	Social interactions Frequency of small group play, self-organised play, direct interventions by educators, boredom feelings episodes were observed. Trained observers recorded and coded these on a six-step scale, ranging from 0= never to 5 = always.	small group play self-organised play direct interventions by educators boredom feelings episodes	t (9)= 2.36; p= 0.02) t (9)= 2.36; p= 0.03 t (9) = -1.42; p = 0.09 t (9) = -1.48; p= 0.09		There was a significantly higher frequency of small group play and self-organised play in the external green space compared to the internal space. There was not a significantly lower frequency of direct interventions by educators and of boredom feelings episodes	▲ ▲ ▲ ▲	Weak
		Stress	Dispute-resolution	F (1, 9) = 7,63; p= 0.022; eta square = 0.46		There was a significant 2- way interaction for	▲	

		Frequency of dispute-resolution interventions by educators, crying episodes and capacity of being quickly comforted in case of crying were observed. Trained observers recorded and coded these on a six-step scale, ranging from 0= never to 5 = always.	interventions by educators Crying episodes Capacity of being quickly comforted in case of crying	F (1, 9) = 4,46; p= 0.064; eta square = 0.33 F (1, 9) = 9,17; p = 0.014; eta square = 0.50		frequency of dispute resolution interventions by educators and capacity of being quickly comforted in case of crying, but not frequency of crying episodes.	▲ ▲	
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Types of natural elements								
Sando (2019), Norway. E: 80 children / 8 ELC	Cross-sectional	Emotional wellbeing Leuven Well-Being Scale which assesses children's emotional wellbeing. This is an observational assessment where children are scored on a scale from 1 to 5. 1= clear signs of discomfort (screaming, anger, sadness) and 5= happy, relaxed.	Emotional Wellbeing (1-5)	Well-being 3.6 (0.6 SD), (regression coefficient = 0.004, p=< 0.05)		Nature was a statistically significant predictor of emotional wellbeing	▲	Weak

Söderström at al (2013), Sweden. E: 172 children / 9 ELC	Cross-sectional	Stress The Salivette®kit (Sarstedt, Numbrecht, Germany). Children were asked to chew a swab for 1 min once in the mid-morning (AM cortisol, 9–10 am) and again the afternoon (PM cortisol, 1 –2 pm). The difference between PM cortisol and AM cortisol was calculated. A positive value implied a rise in PM cortisol level suggesting increased stress.	Stress (PM = AM cortisol)	Low OPEC: -0.4 (1.3 SD) High OPEC: -4.4 (1.9 SD) p= 0.03		Outdoor environment quality a significantly association with stress.	▲	Weak
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Garden-based intervention								
Park et al (2016), South Korea. E: 336 children /12 ELC Prosocial behaviour: 133 children Emotional intelligence: 135 children	Uncontrolled before & after	The revised prosocial behavior questionnaire by Lee (1996) was used. This consists of 20 questions on 4 subscales: helping, sharing, cooperation and kindness. Answers are given on a three-point likert scale (agree, neutral, disagree. Teachers completed this questionnaire based	Emotional intelligence (1-5): Utilization of emotions Recognition and consideration of others' emotions Recognition and	3.35 ± 0.83 3.36 ± 0.59 3.86 ± 0.73	4.01 ± 0.88, p=0.000 3.79 ± 0.68, p=0.000	Emotional intelligence: There was significant improvements in emotional intelligence subcategories from baseline to follow-up	▲	Weak

		on their daily observations. Higher scores indicate a more positive behaviour.	expression of own emotions Emotional regulation and impulse control Relationships with teachers Relationships with peers	 3.62 ± 0.65 3.77 ± 0.90 3.73 ± 0.92	4.30 ± 0.63, p=0.000 4.11 ± 0.81, p=0.000 4.19 ± 0.71, p=0.000 4.09 ± 0.84, p=0.000			
		The emotional intelligence questionnaire consisted of 50 questions on a five-point likert scale (strongly agree - strongly disagree) which was completed by teachers. Higher scores indicate a more positive behaviour.	Prosocial behaviour (1-3). Helping Sharing Cooperation Kindness	 2.37 ± 0.46 2.53 ± 0.41 2.42 ± 0.43 2.30 ± 0.38	 2.57 ± 0.43, p = 0.000 2.66 ± 0.36, p= 0.001 2.66 ± 0.38, p= 0.000 2.55 ± 0.40, p= 0.000	Prosocial behaviour: There was significant improvements in prosocial behaviour subcategories from baseline to follow-up.	▲	

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance (p<0.05)
▼: negative health impact and statistical significance (p<0.05)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association.

Table 9. Nature-based ELC on nature connectedness

Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled Before & After study	See Table 2.	Awareness of surrounding environment	T1 (Jan 2014) E:11.35 (1.22 SD) C:10.07 (1.80 SD)	T4 (May 2015) 13.20 (0.66 SD) 12.86 (1.09 SD) p= 0.004, ηp2= 0.30.	There was a significant time x group interaction on children's awareness of surrounding environment. There were no significant differences between groups at T4.	▲	Weak
Elliot et al (2014), Canada. E: 21 children / 1 ELC	Controlled Before & After (mixed-methods)	Nature relatedness and environmentally responsible behavior An activity where children played against the interviewer. 11 choices were presented (4 nature	Nature Relatedness (out of 8) Environmentally	E: 6.43 (1.25 SD) C: 6.05 (1.05 SD) E:10.57 (0.93 SD)	6.62 (0.97 SD) 5.82 (1.50 SD), p < 0.05 10.71 (1.06 SD)	At post-test, there was a significant difference in nature relatedness scores between the groups. At post-test, there was no significant	▲ ▼	Moderate

C: 22 children / 2 ELC		and 6 environmental behaviour) and the child chose between 2 options. Children received a score of 2 for choosing the more nature-oriented action or environmentally responsible option, and 1 for choosing the alternative option. The max score for nature relatedness was 8 and 12 for environmental behavior.	responsible behavior (out of 12)	C:10.59 (1.14 SD)	10.73 (0.83 SD), p< 0.40	between group differences.		
Müller et al (2017), Canada. E: 43 children / 1 nature-kindergarten C: 45 children / 1 traditional kindergarten	Controlled before & after	As above.	Nature Relatedness (out of 8) Environmentally responsible behavior (out of 12)	E: 6.37 (0.17 SE) C: 5.82 (0.16 SE) E:10.49 (0.18 SE) C:10.29 (0.17 SE)	6.52 (0.18 SE) 6.14 (0.17 SE) p= 0.22, η2= 0.02 10.49 (0.18 SE) 10.51 (0.17 SE) p= 0.83, η2= 0.00	At post-test there was a small and non-significant effect At post-test there was no significant effect	▲ ▼	Weak
Nazaruk & Klim-Klimaszewska (2017), Poland. E: 90 children (50	Uncontrolled before & after	Knowledge and skills of nature Pre-test: A standard card test consisting of 6 illustrated worksheets with tasks for children to	Knowledge and skills of nature categorised into the following: pre-test:	City Low= 12% Average= 56% High= 32% Rural Low= 0%	City Low= 0% Average= 28% High= 72% Rural Low= 0% Average= 20%	Children scored higher at post-test compared to pre-test.	▲	Weak

urban / 40 rural)		complete. Teachers explained and conducted the test. Children's performance was rated on a scale of 1 to 3 (1= nature skills have not been mastered, 3= nature skills have been fully mastered). Children could score a max of 18 points. Post-test: Observation and a picture test consisting of 10 illustrated worksheet cards with tasks for children. A similar scoring to pre-test was used and the children could get a max of 30 points.	Low (0-9) Average (10-14) High (15-18) Post-test: Low (0-15) Average (16-23) High (24-30)	Average= 50% High= 50% p = 0.3	High= 80% p = 0.8093			
Yilmaz et al (2020), Turkey. 40 children / 1 ELC	Uncontrolled before & after	Biophilia Adapted tool originally developed by Rice and Torquati (2013) below.	Biophilia Scores (out of 11)	19.78, 1.510 (SD), 0.239 (SE)	20.33, 1.309 (SD), 0.207 (SE) Mean diff: -0.55, 1.584 SD, 0.251 SE (95% CI: -1.057, -0.043), p= 0.034	There was a significant difference in the Biophilia scores from pre-test to post-test.	▲	Weak
Barrable et al (2020), UK (England, Scotland, Wales).	Controlled cross-sectional	Connectedness to nature The connectedness to Nature Index for	Total CNI score	E: 4.22 (0.47 SD) C: 3.92 (0.60 SD)		Children attending nature nurseries scored higher for	▲	Weak

E: 141 /12 ELC C: 110 children / 6 ELC		Parents of Preschool Children (CNI-PPC) consists of 16-items and responses are given on a five-item Likert scale ranging from “strongly disagree” to “strongly agree”. It consists of 4 dimensions: enjoyment of nature, empathy for nature, responsibility toward nature and awareness of nature.	Enjoyment of nature Empathy for nature Responsibility toward nature Awareness of nature	E: 4.41 (0.54 SD) C: 4.05 (0.67 SD) ($\beta = 0.59$, $p = 2.61 \times 10^{-15}$) E: 3.78 (0.71 SD) C: 3.63 (0.80 SD) E: 3.96 (0.68 SD) C: 3.85 (0.71 SD) ($\beta = 0.76$, $p = 2 \times 10^{-16}$) E: 4.45 (0.53 SD) C: 3.98 (0.67 SD)		enjoyment and responsibility	▲ ▲ ▲ ▲	
Giusti et al (2014), Sweden. E: 11 children / 2 ELC C: 16 children / 5 ELC	Controlled cross-sectional	Children's affinity with biosphere The teacher presented children with image-based tasks (games) in which they had to select an image based on set questions. This assesses emotional and cognitive affinity to nature.	Emotional Affinity with the Biosphere Cognitive Affinity with the Biosphere	E: 0.792 (0.121 SD) C: 0.665 (0.154 SD), p= 0.031, d= 0.916 E: 0.771 (0.134 SD) C: 0.660 (0.133 SD), p= 0.045, d= 0.845		Children with nature-rich routines score significantly higher than children with nature-deficit routines. As above.	▲ ▲	Weak

Rice & Torquati (2013), USA. E: 68 children / 6 ELC C: 46 children /4 ELC	Controlled cross-sectional	Biophilia Interview consisting of 11-items which assess preference for being outdoors, enjoyment of sensorial aspects of nature, curiosity about nature, and interacting with nature. Biophilic responses were scored 1 and non-biophilic responses were scored 0.	Biophilia Scores (out of 11)	E: 7.7 (2.3 SD) C: 7.7 (2.4 SD), p= 0.94		There was no significant difference between the nature and non-nature groups	►	Weak
Ernst (2014), USA. E: 46 educators	Cross-sectional	Development of environmental appreciation See table 2.	Environmental appreciation (1-5) Belief regarding difficulty in using natural outdoor settings Belief regarding one's relationship with nature	4.43 (1.31 SD) r= 0.83, p ≤ 0.05 b= 0.71, SE= 0.08, B= 0.83, p<.001 r= 0.31, p ≤ 0.05 b= 0.25, SE= 0.21, B= 0.11, p= 0.25		There was an association between frequency of nature experiences and belief regarding difficulty in using natural outdoor settings and belief regarding one's relationship with nature Belief regarding difficulty in using natural outdoor settings was a significant predictor of use of natural outdoor settings with their preschool students, belief regarding one's	▲ ▲	Weak

						relationship with nature was not.		

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Table 10. Nature-based ELC on play behaviour								
Study details (Author, year and country) **Sample size** (n of children / n ELC settings for exp and con)	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Nature-based ELC								
Agostini et al (2018), Italy. E: 41 children / 7 teachers / 1 school C: 52 children / 13 teachers / 1 school	Controlled Before & After study	Play See Table 2.	Play (mean and SD)	T1 (Jan 2014) E:11.26 (1.08 SD) C: 9.89 (1.22 SD)	T4 (May 2015) 13.15 (0.99 SD) 12.78 (1.14 SD) p= 0.00; ηp2= 0.41	There was a significant time x group interaction on children's play. There were no significant differences between groups at T4.	▲	Weak
Cordiano et al (2019), USA. E: 12 children / 1 ELC class. C: 14 children / 1 class. Children from the same school.	Controlled before & after study	Play Interaction, Play Disruption, and Play Disconnection Assessed using the Penn Interactive Peer Play Scale (PIPPS), which is a 32-item behaviour rating instrument assessing aspects of children's peer play behaviors. Pretend Play rating consisted of 5	Teacher Play interaction Pretend play	T1 - baseline E:49.46 (6.99 SD) C:54.96 (2.64 SD) E:15.18 (1.66 SD) C:18.21 (2.12 SD)	T3 - endpoint 54.69 (5.07 SD) 55.82 (2.76 SD) Within group: p<0.01, η^2p= 0.26 Between group: (F=2.70, η2p= 0.11, p>0.05) 23.45 (2.12 SD) 18.86 (3.35 SD) Within group: p<0.01 η2p= 0.29	Small effect for between group No effect for between group	▼ ▲	Weak

		questions on a 5 point likert scale to assess children's imagination in play, use of make-believe, enjoyment of play, amount of emotion expressed in play, and use of make-believe in dramatic play.			Between group: F=0.00, η2p= 0.00, p>0.05			
			Play disruption	E:50.38 (5.96 SD) C:43.69 (6.43 SD)	47.71 (7.26 SD) 38.31 (5.53 SD) Within group: non-sig, η2p= 0.06 Between group: F=17.64, η2p= 0.45, p<0.001	Large effect for between group	▼	
			Play disconnection	E:52.13 (7.34 SD) C:43.71 (5.63 SD)	45.75 (9.28 SD) 40.14 (4.69 SD) Within group non-sig, η2p= 0.08 Between group: F=14.59, η2p= 0.39, p<0.01	Large effect for between group	▼	
			Parent Play interaction	E:46.90 (6.72 SD) C:48.00 (7.00 SD)	51.30 (7.46 SD) 51.22 (9.91 SD) non-sig, η2p= 0 .07	There were non-significant and small effects for between group and school x time across all four play types.	▲	
			Pretend play	E:20.90 (3.54 SD) C:21.80 (3.58 SD)	21.50 (3.24 SD) 22.00 (4.03 SD) non-sig, η2p= 0 .00		▼	
			Play disruption	E:49.11 (9.21 SD) C:50.00 (3.81 SD)	44.89 (8.25 SD) 44.00 (7.50 SD) non-sig, η2p= 0 .02		▼	

			Play disconnection	E:49.63 (11.20 SD) C:50.33 (8.54 SD)	48.38 (10.04 SD) 46.11 (9.32 SD) non-sig, η2p= 0 .03		▼	
Burgess & Ernst (2020), USA. E: 84 children / 4 ELC C: 24 children / 2 ELC	Controlled Before & After study	Play behaviours The Penn Interactive Peer Play Scale consists of 32 items with 3 dimensions: play interaction, play disruption and play disconnection Teachers and parents indicate frequency of behaviours on a 4-point Likert scale (never, seldom, often, always)	Adj means (SE) Teacher: Play interaction	E: 23.44(0.31 SE) C:17.75 (0.37 SE)	E:28.82 (0.32 SE) C:26.13 (0.63 SE) p<.001, η2= 0.12	At post-test children in the nature ELC had significantly higher play interaction scores and lower play disruption and disconnection scores compare to the non-nature ELC. (adjusted for pretest levels, age, gender, prior participation, and part v. full-time participation)	▲	Weak
			Play disruption	E:28.11 (0.67 SE) C:25.19 (1.69 SE)	E:20.06 (0.48 SE) C:25.22 (0.95 SE) p <001, η2=0.19		▲	
			Play disconnection	E:19.40 (0.53 SE) C:15.88 (1.47 SE)	E:12.44 (0.32 SE) C:15.17 (0.65 SE) p<.001, η2= 0.12		▲	
			Parent: Play interaction	E:25.77 (0.30 SE) C:25.33 (0.75 SE)	E:27.15 (0.28 SE) C:26.92 (0.58 SE) p= 0.72, η2<.01	No significant differences between the nature and non-nature ELC at post-test.	▲	
			Play disruption	E:29.82 (0.45 SE) C:28.47 (1.20 SE)	E:27.85 (0.45 SE) C:28.45 (0.94 SE) p= 0.57, η2< .01		▲	
			Play disconnection	E:17.75 (0.37 SE)	E:16.06 (0.33 SE)		▼	

				C:18.27 (1.27 SE)	C:16.03 (0.69 SE) p= 0.97, η2<.001			
Robertson et al (2020), Australia. E: 15 children / 1 ELC C: 15 children / 1 ELC	Controlled cross-sectional	Sociodramatic play Smilansky Scale for the Evaluation of Dramatic and Socio Dramatic play (SSEDSP). Observation of each child (6x5 minute intervals) and scored: 0=characteristic is not present 1=characteristic is present but to a limited degree 2=characteristic is present to a moderate degree 3=characteristic is present consistently and in many situations during the child's play Total score was calculated using sum of each 5 min interval (score could be 0 - 18) and represented overall complexity of play	Sociodramatic play: Role play Make believe with objects Actions and situations Persistence Interaction	E: 6.35 (1.96 SD) C: 2.04 (2.65 SD) t (28) = 5.07, p= 0.00) E: 1.04 C: 0.34 SD= 0.16, p= 0.00, eta squared= 0.39 E: 0.92 C: 0.31 SD= 0.14, p= 0.00, eta squared= 0.42 E: 0.99 C: 0.34 SD=0.14, p= 0.00, eta squared= 0.44 E: 1.11 C: 0.27 SD= 0.16, p= 0.00, eta squared= 0.50 E: 1.20 C: 0.34 SD= 0.14, p= 0.00, eta squared= 0.56	Mean diff= 0.86, (95% CI: - 2.04–6.35, eta squared = 0.47).	There was a significant difference between the sociodramatic play of children in nature ELC compared to the control The magnitude of the differences in the means was large. There were also significant differences in characteristic of Socio Dramatic Play.	▲ ▲ ▲ ▲ ▲	Weak

			Verbal communication	E: 1.20 C: 0.34 SD= 0.15, p= 0.00 eta squared= 0.53			▲	
Study details / Sample size	**Study Design**	**Outcome and measurement**	**Units**	**Baseline or one time point (cross-sectional)**	**Follow-up (if applicable) or mean difference**	**Summary of Findings**	**Effect Direction**	**Quality Rating**
Naturalised Playground								
Brussoni et al (2017), Canada. E: 48 children / 2 childcare centres Play: 16 children (sub-sample)	Uncontrolled before & after (mixed methods)	Play behaviours Each child was observed twice over 30 min of outdoor play at baseline and follow-up by two researchers. Observations were coded as follows: prosocial behaviours (co-operative play, social conversation), antisocial behaviours (physical and verbal aggression, object possessiveness, rejected bids for engagement), lack of engagement in play (onlooking, unoccupied), channel surfing (transitioning frequently between activities), child teacher interactions (teacher initiated, child-initiated, interruption by	Play: Prosocial behaviours Antisocial behaviours Lack of engagement in play Channel surfing, Child teacher interactions Play with natural materials Risky play Gender-segregated play		OR: 2.81, (95% CI: 1.17-6.91), p< 0.05 OR: 1.40, (95% CI 0.47-4.13) OR: 0.52, (95% CI: 0.24-1.14) No change. OR: 1.30, (95% CI: 0.65-2.57) OR: 7.29, (95%CI: 1.53-38.09), p< 0.05 OR: 1.11, (95% CI: 0.55-2.27) No change.	There were a significant intervention effects for play with natural materials and prosocial behaviour. There were no significant intervention effects for the remaining play types. Channel surfing and gender segregated play did not change.	▲ ▲ ▲ ► ▲ ▲ ▲ ► ▲ ►	Moderate

		teacher), play with natural materials (natural loose materials, natural play elements), risky play (rough and tumble, height, mastery, unstable, speed, risk of getting lost), and gender-segregated play.	Solitary play		OR: 1.13, (95% CI 0.60-2.15).		▲	
Cloward Drown et al (2014), USA. E: 24 children / 1 ELC (observed in 2 different playgrounds, natural vs manufactured)	Controlled cross-sectional	Dramatic Play Smilansky Scale (modified) was used to code children's dramatic play. The scale uses 5 behaviors and persistence of a play episode to indicate dramatic play: imitative role-play, make-believe with objects, make-believe with actions and situations, interaction, verbal communication and persistence of play episode	Dramatic Play (%) Playground type (natural vs manufactured) Play props (natural, manufactured, none)	E: 12% C: 10% Pearson x2 = (3, 1006) = 12.19, p = 0.007) Pearson x2 = (6, 802) = 23.09, p= 0.001)		Playground type and type of dramatic play were found to be significantly related with the natural playground affording more dramatic play than the manufactured playground. A significant relationship was found between play prop use and dramatic play Natural play props were not used frequently or highly associated with dramatic play.	▲ ▲ ▲	Weak
		Social Play MildredParten's (1932) stages of play were used to describe social interaction and maturity of play:	Social Play (%) Playground type (natural vs manufactured)	Pearson x2 = (3, 751), 5.07, p= 0.167		There was no relationship between playground type and type of social play indicating both playgrounds provided similar affordances for social		

		unoccupied play, solitary play, onlooker play, parallel play, associative play, cooperative play. Child's play was observed in 30-second intervals for ten-minute period. Observers recorded a child's location at the start of each 30-second interval and or the remainder of 30-second interval, the play types, persitance and location (natural, manufactured, none).	Play props (natural, manufactured, none)	No association		play.		
Luchs, & Fikus (2013), Germany. E: 38 children / 1 ELC C: 21 children / 1 ELC	Controlled cross-sectional	Play episodes and frequency Observation - information on place, duration, social category of play and narrative was collected. The play episodes were then coded afterwards: **-play with**: functional play and constructional play. -**play as**: well-known meaning and displays a different object within the child's play and imagination,	Number of play episodes Duration of play episodes 0-5mins 6-10 mins 11-15mins 16-20mins:	E: 3.05 ± 1.71 C: 5.57 ± 1.47. E: 36% C: 58% E: 32% C: 35% E: 12% C: 7% E: 8% C: 0%/ E: 5% C: 0%		During the 30 minutes observed, there were significantly different number of play episodes between the natural and contemporary playgrounds.		Weak

		orientation on role-models, not only copying but also developing their own play while realizing their own ideas, wishes and needs **-play for**: play with rules, organizing activities of several players **- others** **- combination**	21-25mins	E: 8% C: 0%				
			26-30mins					
			Frequency of play categories Play with	E: 1.45 ±1.37 C: 3.14 ±1.68 p= 0.000		Children in the contemporary playground engaged in significantly higher play episode categories. Combination was non-significant	▼	
			Play as	E: 0.53 ±0.83 C: 0.62 ±0.97 p= 0.701			▼	
			Play for	E: 0.13 ±0.41 C: 0.52 ±0.68 p= 0.023			▼	
			Other	E: 0.24 ±0.49 C: 0.67 ±0.73 p= 0.022			▼	
			Combination	E: 0.71 ±0.8 C: 0.62 ±0.8 p= 0.677			▼	
			Combination Patterns of play categories (%) Play with	E: 44.66 ±35.67 C: 56.18 ±27.45 p= 0.204		Play for and combination play were significantly different.	▼	
			Play as	E:18.92 ±27.87		Combination play which was preferred	▼	

				C: 11.78 ±23.28 p= 0.324		by children in the nature playground.		
			Play for	E: 3.23 ±10.46 C: 9.93 ±13.45 p= 0.056			▼	
			Other	E: 6.3 ±13.34 C: 11.45 ±12.31 p= 0.151			▼	
			Combination	E: 26.9 ±32.71 C: 10.66 ±15.0 p= 0.012			▲	
Dyment et al (2013), Australia. E: 120 children / 3 ELC C: 40 children / 1 ELC	Cross-sectional	Play types System for Observing Play and Leisure Activity in Youth (SOPLAY) was used to collect data on play types across various playground areas. The categories of play types were functional, constructive, symbolic, self-focused/looking on and talking.	Play types in natural areas Functional (physical play activities)	E: ELC A= 24.0 ELC C= 58.3 ELC D= 52.2 C: ELC B= N/A		Functional play was the most popular type of play in natural areas in the experimental schools. Symbolic play was infrequent and only observed in one experimental ELC.		Weak
			Constructive (building play activities)	E: ELC A= 14.7 ELC C= 19.2 ELC D= 13.0 C: ELC B= N/A				
			Symbolic (creative/ imaginative play)	E: ELC A= 8.0 ELC C= 0 ELC D= 0 C: ELC B= N/A				

Morrissey et al (2017), Australia. E: 28 children / 1 ELC C: 28 children / same school as E.	Cross-sectional	Sociodramatic play episodes Observation (2 independent researchers) using the Dramatic Play Data Collection Tool. The following play behaviours were coded: - Play themes or roles were identified as present or absent in the episode: fantasy, domestic, occupational, conventional superhero or other. - Frequencies of object substitutions - Frequencies of imaginative transformations - Frequencies of explicit metacommunications used to plan and organise play Additional contextual information was also collected	Fantasy Domestic Occupational Superhero Other Relationship between sociodramatic play variables and context. Object substitutions Explicit metacommunication Imaginative transformations	E: 10 / C: 4 E: 8 / C: 15 E: 1 / C: 3 E: 2 / C: 0 E: 0 / C: 2 $\chi 2 = 21.71$, $p < 0.001$ $\chi 2 = 10.04$, $p < 0.01$ $\chi 2 = 6.63$, $p < 0.05$		There were significant associations between object substitutions, explicit metacommunication and imaginative transformations and the yard type (natural versus traditional). Children from the natural playground engaged in longer episodes of sociodramatic play episodes compared to children from the traditional playground and were more likely to engage in object substitutions, explicit metacommunication and imaginative transformations.	▲ ▲ ▲	Weak

Study details / Sample size	Study Design	Outcome and measurement	Units	Baseline or one time point (cross-sectional)	Follow-up (if applicable) or mean difference	Summary of Findings	Effect Direction	Quality Rating
Types of natural elements								
Zamani (2013), USA. 36 children / 1 ELC	Cross-sectional (mixed-methods – thesis)	Cognitive Play Behaviour mapping - assesses individual cognitive play in the different zones. Children are observed for 7 days in 12 observation sessions during recess (11.30am and 4.15pm - lasted 45 minutes). The researcher scanned each zone and repeated for 4 rounds per recess. Childs location, gender, ethnicity, behaviour setting type, physical elements, cognitive play behaviour and teacher interactions were recorded. Each child was observed for 10 seconds and recorded for 20.	% time in play categories Functional Constructive Exploratory Dramatic Games with rules Functional Constructive Exploratory Dramatic	**Natural:** Within = 30.7; withinCog= 27.5 Within = 8.1; withinCog= 47.2 Within = 12.8; withinCog= 45 Within = 37.1; withinCog= 40.2 Within = 3.1; withinCog= 3.1 x= 281.70, 4*** **Mixed:** Within = 35.2; withinCog= 35.2 Within = 4.5; withinCog= 29.1 Within = 10.9; withinCog= 42.7 Within = 26.8; withinCog= 32.5 Within = 13.9; withinCog= 62.1		All zones mainly afforded functional play opportunities. The natural zone afforded higher levels of dramatic, exploratory and constructive play compared to the other zones.	N/A	Weak

			Games with rules Functional Constructive Exploratory Dramatic Games with rules	x= 201.46, 9*** **Manufactured:** Within = 44.2; withinCog= 37.3 Within = 4.3; withinCog= 23.6 Within = 3.7; withinCog= 12.3 Within = 26.7; withinCog= 27.3 Within = 6.8; withinCog= 25.7 x= 224.86 3***				

Abbreviations: E= experimental; C= control; n= number; ELC = early learning and childcare (includes preschools, day care, kindergarten etc.); SD= standard deviation; SE= standard error; CI= confidence intervals.

Effect direction explained:
▲: positive health impact
►: no change/ conflicting findings
▼: negative health impact
▲: positive health impact and statistical significance ($p<0.05$)
▼: negative health impact and statistical significance ($p<0.05$)
No arrow: no inferential statistics reported

Controlled before & after studies – difference between experimental and control group at follow-up (unless stated) or difference in change between experimental or control group. Uncontrolled before & after studies – change since baseline (unless stated). Controlled cross sectional – difference between experimental and control (unless stated). Cross-sectional – positive, negative or no association

Qualitative

Table 12. Findings from eligible qualitative studies			
Theme	**Sub-theme**	**Studies**	**Quotes**
Natural settings provide more affordances compared to traditional settings	Natural settings **enable** children to **diversify** their **play** (inc. imaginative, spontaneous, risky, manipulative, cognitive, exploratory and active play)	Dowdell et al (2011); Herrington & Studtmann (1998); Liu (2020); Puhakka et al (2019); Sandseter (2009); Wishart et al (2019); Zamani (2015).	*"The children also invent themselves; when they have stimulus for their eyes, children invent it [activity] without your help. And it should be like this; some part should be like this. But you need to have stimulus. It's not enough to have a brown yard and a climbing frame. So, it [green yard] added somehow; they definitely had good games. They pretended that they had a campfire, they got the stones as sand pretended that they were on a trip. And their imagination was in use there, and when children use their brains, natural tiredness arises, and it did them good, a lot of good. Then rest comes naturally, and you have a good appetite and we're in the positive cycle. So they could use their imagination, and we encouraged them. We didn't prohibit them, we just advised them not to rip anything." (Puhakka et al, 2019).*
	Natural settings **enable** children to engage in **high intensity physical activity**	Bjørgen (2016); Puhakka et al (2019).	*"High physical-motor levels are created, the children jump down and run back up. They talk, shout and laugh. Three of the girls jump together and try to land in differing ways. They hold hands and try to jump together from the small knoll. There is laughter. They are eager and enduring. The small knoll has many opportunities for variation, in height and width, which invite challenges suitable for each child's resources. The children have visual, verbal and physical contact with each other. The top of the knoll provides an overview. Some find it scary the first time they try, but together they challenge each other, supporting and encouraging each other. The children decide how much they will participate and how they jump, and how they wish to solve the challenges offered by the knoll" (Bjørgen, 2016).*
	Natural settings **afford** children with **higher levels of risk** compared to traditional settings	Sandseter (2009); Streelasky (2019).	*I like playing in the fallen logs and trees on the playground; it is so much fun, but a bit scary too! I like the big pile of sticks and logs that we made – it is for another fort that is going to be really high off the ground." (Streelasky, 2019)*

	Natural settings **afford** more variation (the space and elements) to support children to **use and increase their imagination and creativity**	Liu (2020); Streelasky (2019); Zamani (2015).	*"I like being outside with my friends. We make shelters and we make up different games, like getting trapped on an island, or being on a boat and making our escape! I like doing science outside too – like different experiments, especially when the sun is out." (Streelasky, 2019).*
	Natural settings **enable peers and teachers to interact differently**	Bjørgen (2016); Dowdell et al (2011); Liu (2020); Streelasky (2019).	*"The children are shouting 'X… can't you catch us? Please catch us, try to catch us …'. The staffs join the situation and run after the children. The children are shouting 'Catch me … can't catch me' … There is excitement and the staff are running after the children, catching them and holding them before releasing them. The staffs have high energy, the children focus on the adults, avoiding being caught. The adults show empathy, holding and hugging the child when it is caught. The game is exciting and creates enthusiasm. A high level of physical activity is created, by climbing up, sliding down, running around and hiding in the tower to escape capture by the adults. They run at high speed and the children's body language shows that they are very much engaged in the game" (Bjørgen, 2016)*
	Natural settings **increase child-initiated learning** and students perceiving themselves as capable learners compared to traditional settings	Dowdell et al (2011); Maynard et al (2013), Zamani (2015).	*"[CogG] has poor concentration, sees herself as the baby, finds it difficult to sit and listen to story. She is extremely lacking in confidence … shy … she won't look at you indoors. With child-led learning she is totally engrossed and remains on task. Outside is the best learning environment for her … she remains on task. When outside she will come over and say 'I like this' and 'I like doing that', 'this is my favourite place'." (Maynard et al, 2013).*
	Children have **increased contact with nature** enabling them to increase their knowledge of nature	Dowdell et al (2011); Liu (2020); Puhakka et al (2019).	*"Especially about the forest floor mat, I remember that our children kept asking, 'what is it' and 'what's growing there', and explored it very carefully; they were almost lying on their stomachs there. Especially the older ones, and they had a lot of questions about it." (Puhakka et al, 2019).*
Natural and traditional settings	**Movement types and intensity** similar across	Wishart et al (2019).	Not available.

provide similar affordances	natural and traditional spaces		
	Frequency of risky play is similar in both natural and traditional settings	Sandseter (2009)	Not available.
Children's preferences of setting types	**Natural environment** is more diverse and engaging and **preferred** by children compared to traditional settings	Bjørgen (2016); Streelasky (2019).	*"I like going outside and playing! I like playing with my friends, Sydney and Megan. We play hide and seek on the playground and hide in the forest in the logs and trees. I like outside because it's so fun and I really like to play. Sometimes I play with my sister too; I like all the colours outside and all the space." (Streelasky, 2019).*
	Mixed areas (combining both natural with traditional elements) are **preferred** by children	Zamani (2015).	Not available.
Restorative effect of nature		Liu (2020); Puhakka et al (2019),	*"Now it's become very difficult to finish playing. They would rather continue, and those who need to take a nap, they've had a nice, long time outdoors and nice games so they fall asleep more easily, and it affects their energy in the afternoon. Some children have very long days here. They come in the morning and stay until five o'clock; they seem to be somehow energetic and lively in the yard. This is new for us. The contrast to the previous yard is so great that the effects can be seen here very quickly." (Puhakka et al, 2019).*