ACUTE EFFECTS OF OUTDOOR VERSUS INDOOR EXERCISE: A SYSTEMATIC REVIEW AND META-ANALYSIS

by

Luke Peddie

B.Kin., The University of British Columbia, 2018

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES

(Kinesiology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

April 2021

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis entitled:

Acute effects of outdoor versus indoor exercise: a systematic review and meta-analysis

submitted by Luke Peddie in partial fulfillment of the requirements for

the degree of Master of Science

in Kinesiology

Examining Committee:

Dr. Eli Puterman, Kinesiology, UBC
Supervisor

Dr. Michael Koehle, Kinesiology, UBC
Supervisory Committee Member

Dr. Guy Faulkner, Kinesiology, UBC
Supervisory Committee Member

Abstract

Physical activity and exposure to nature have each been recognized for their positive effects on health and wellbeing. When taken in tandem, outdoor exercise is proposed to have additive benefits compared to exercising indoors or being inactive outdoors. Previous reviews of green exercise have reported inconclusive findings due to a paucity of high-quality evidence. The present review sought to summarize the body of literature that compares physiological and perceptual differences of a single bout of exercise in outdoor spaces versus indoor spaces.

Following the PRISMA reporting guidelines for systematic reviews, a search was conducted in nine databases for any articles published before November 2019. When studies and outcomes were methodologically uniform, quantitative analyses was completed. Vote counting and harvest plots were used to synthesize the remaining outcomes. Quality of articles was assessed using the Cochrane Risk of Bias Assessment Tool.

The findings of 24 articles (Total N = 757) were examined. Summarized outcomes include objective exercise intensity, perceived exertion, performance, neuroendocrine responses, cardiovascular responses, thermoregulation, enjoyment, intention for future exercise and perceptions of the environment. Meta-analysis was conducted for mean heart rate, perceived exertion, mean speed, time to completion, enjoyment, and future intention for exercise. Significant effects of the environment were detected for meta-analyses of perceived exertion (g = -0.84, 95% CI = [-1.60, -0.09], p = 0.03) and enjoyment (g = 1.24, 95% CI = [0.59, 1.89], $p < 0.001$). Methodological diversity made it impossible to statistically summarize the remaining outcomes, results split between no effect and statistically significant changes, resulting in inconclusive findings. The majority of included studies (54%) were assessed to have a high risk of bias and all other studies (46%) were assessed to have some concerns of risk of bias.

Although no additive physiological benefits were identified as a result of exercising outdoors, exercise still presents well-documented potential for improved health. At equivalent objective intensity, outdoor exercise appears to feel easier and more enjoyable than indoor exercise. Outdoor exercise may be more likely to be repeated and more sustainable, allowing physiological benefits to accrue over repeated bouts.

Lay Summary

Physical activity and exposure to nature are each individually recognized for their benefit to physical and mental health and wellbeing, but it is unclear if exercising outdoors provides any additional benefit. The purpose of this systematic review and meta-analysis was to summarize all research that has examined physical and perceptual outcomes of exercising outdoors compared to indoors. Twenty-four articles were found that met the inclusion criteria, but their low quality and considerable differences in reporting made it challenging to interpret findings. Two outcomes, perceived exertion and enjoyment, were able to be statistically compared. It was concluded that exercise outdoors is at least equally physiologically beneficial as indoor exercise but feels easier and is more enjoyable. Therefore, outdoor exercise is more likely than indoor exercise to result in greater engagement in physical activity and health benefits in the long term.

Preface

This body of research represents work that was conducted while a member of the Fitness Aging and Stress (FAST) Lab under supervisor, Dr. Eli Puterman. I was primarily responsible for the conception, collection and analysis of data, and drafting and revising the manuscript. Graduate students and volunteers of the FAST lab assisted with the collection, cleaning, and management of data. This systematic review and meta-analysis was prospectively registered at PROSPERO on July 24, 2018 (Registration Number: CRD42018100314).

Table of Contents

List of Tables

List of Figures

List of Symbols

#	Denotes calculated values
d	Measure of standardized mean difference
g	Measure of standardized mean difference
I^2	Statistical heterogeneity
n	Sample size
N	Aggregate sample size
p	Level of statistical significance
α	Alpha; *a priori* significance level
η_p^2	Partial eta squared; Measure of effect size
χ^2	Chi squared; Measure of statistical heterogeneity

List of Abbreviations

B	Brown natural environment
BAGE	Belief about green exercise
BH	Basement hallway
bpm	Beats per minute
BSL	Brown natural environment below sea level
BW	Body weight
CB	Counterbalanced
CI	Confidence interval
CL	Clincal setting
CS	Comfortable, self-selected intensity
D	Dance
DS	Dance studio
EIC	Exercise intensity confirmation
F	Female
FC	Fitness centre
FIE	Future intention for exercise
G	Green natural environment
GU	Green-urban mixed environment
H	Hiking
HF	Power in high frequency range
HFn	Power in high frequency range normalized
HG	High exercise experience group (stratified sample)

HR	Heart rate
HRR	Heart rate reserve
I	Indoor condition
I-ES	Indoor condition with external auditory stimuli
I-IS	Indoor condition with internal auditory stimuli
IT	Indoor track
LF	Power in low frequency range
LF/HF	Ratio between power in low frequency and power in high frequency
LFn	Power in low frequency range normalized
LG	Low exercise experience group (stratified sample)
MI	Moderate intensity
M	Male
MS	Moderate, self-selected intensity
Mt	Mountainous natural environment
NC	Not calculated due to insufficient reporting
NM	Near maximal to maximal exercise intensity
NR	Not reported
N-RXT	Non-randomized crossover trial
NS	Not significant
NW	Natural environment by water
O	Outdoor condition
OG	Over-ground
OT	Outdoor track

PACES	Physical activity enjoyment scale
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
Q-RCT	Quasi-randomized comparative trial
R	Running
RCT	Randomized comparative trial
RMSSD	Root mean square of successive differences between successive normal-to-normal beat intervals of heart rate
RPE	Rating of perceived exertion
RXT	Randomized crossover trial
S	Self-selected intensity
SD	Standard deviation
SDNN	Standard deviation of normal-normal beat intervals of heart rate
SE	Standard error
SVM	Self-selected vigorous-maximal intensity
T	Taekwondo
TM	Treadmill
TS	Self-selected training intensity
U	Urban environment
UC	University campus
V	Vigorous intensity
VO_2	Oxygen consumption
VO_2R	Oxygen consumption reserve
VR	Virtual reality or visual stimulus
W	Walking

Acknowledgements

A special thank-you to my supervisor, Dr. Eli Puterman, without whom I could not have completed this research. Thanks, as well, to all the members of the FAST lab who assisted and encouraged me along the way, but particular thanks to Ben Hives for keeping me grounded.

Finally, thank-you to my family and Quincy Brozo for providing me the love and support to continue progressing through this particularly challenging year.

"Things have a way of working out."

Dedicated to my mom and dad for always being my biggest fans and showing me there is always a bright side.

Chapter 1: Introduction

Exercising outdoors has gained significant attention as a method of health promotion in novel research and policy initiatives[1,2]. Such attention is, in part, a result of the ever-growing body of literature suggestive of significant positive relationships between exercise, exposure to nature, and health[3–6]. Investigations into their independent effects suggest that regular physical activity[5] and exposure to rural or urban greenspaces[6] have preventative effects against all-cause mortality and chronic medical conditions. Accumulating evidence further directs attention to the synergistic long-term benefits to physical and mental health and wellbeing when exposure to natural spaces and exercise are completed in tandem, as opposed to when exercising indoors[7–11].

While determining the long-term health impact of outdoor exercise is of particular importance to public health, it is equally important to determine the short term physical and psychological impacts of a single bout of exercise completed outdoors compared to indoors for several reasons. Discovering the fundamental qualities of the environment (e.g., its greenness[12], biodiversity[13,14], or potential for recovery from stress and depleted attention[15–17]) that clarify outdoor exercise's physiological and/or psychological advantage over indoor exercise can offer insight when designing practical exercise programming for improving long-term physical and mental health. Furthermore, by summarizing the literature of the immediate, transient responses to outdoor exercise, the interaction effects of exercise and exposure to nature can be quantified and contextualized to provide paths for future research.

In light of the exponential growth in research since the term "green exercise" was introduced in 2003[18,19] (see Figure 1.1), two systematic reviews have undertaken the task of synthesizing the evidence comparing the health effects of outdoor exercise to indoor exercise in terms of both acute and long term physical and psychological effects[9,10]. While the findings of these reviews were largely inconclusive, significant differences in enjoyment and affective valence between exercise environments were identified. However, low quality evidence and inclusion of virtual reality simulations of outdoor exercise in indoor environments made interpretation difficult. Authors noted equivocal findings for physiological outcomes, which is likely related to the distinct lack of consistent measurements and reporting by studies. The purpose of the current systematic review and meta-analysis, then, is to update the previous review conducted by Lahart and colleagues[10], with a focus on the differential physiological and perceptual effects of acute bouts of exercise performed in outdoor settings compared to ones performed indoors.

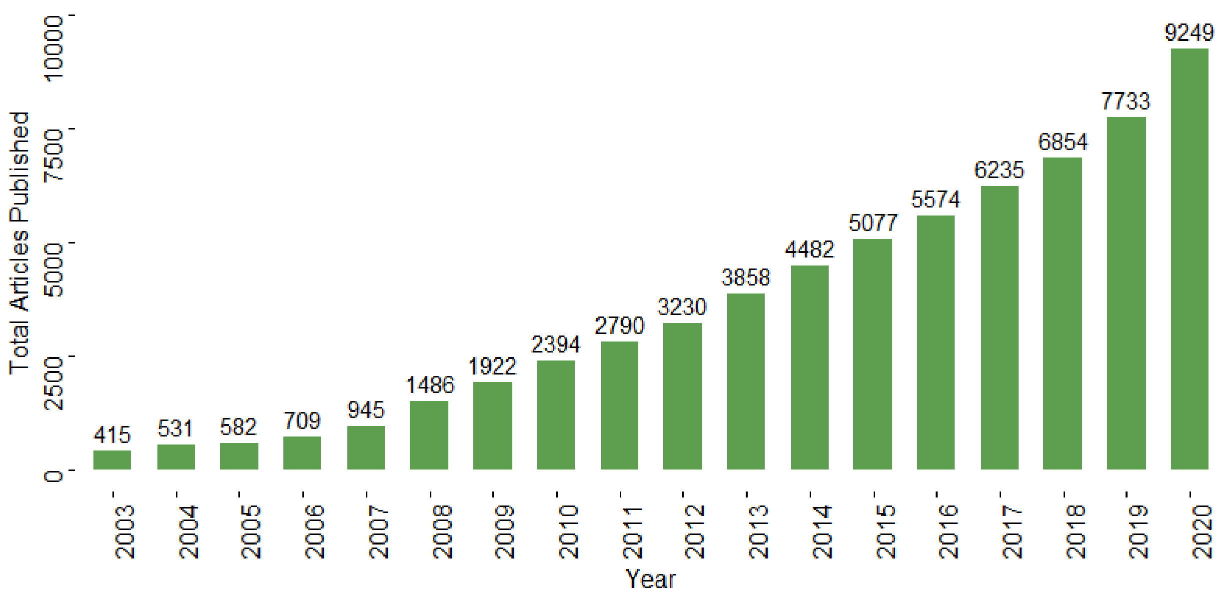


Figure 1.1 Publications indexed in Europe PMC between 2003-2020 that reference "green exercise".

1.1 Benefits of Physical Activity

Physical activity's impact on the health and wellbeing of the population cannot be understated. Regular physical activity has been associated with decreased risk of premature mortality, and primary and secondary prevention of up to 26 chronic conditions, including psychiatric, neurological, metabolic, cardiovascular, and pulmonary diseases, musculoskeletal disorders, and multiple types of cancers[5,20]. For example, individuals with cancer can almost double their chance of survival by becoming physically active after being diagnosed with breast cancer or colorectal cancer[21,22]. In individuals with osteoarthritis, regular exercise has been associated with short-term improvements in pain, function, and quality of life that can extend into the long term[23]. Regular exercise can even be used as treatment with clinically significant effects, which in cases such as depression and hypertension, approach or exceed the clinical effects of pharmacological treatments[5].

The mechanisms by which regular physical activity and exercise behaviours can effectively reduce negative health outcomes are still not fully elucidated. As recently as 2017, there have been calls for more research examining the mechanisms linking regular exercise to health outcomes[20]. One approach to do so is to examine the relationship between exercise and health through an acute, experimental lens. By reviewing studies which utilize an acute experimental design, it is possible to identify which parameters can be affected by a single bout of exercise. If exercise is frequently repeated, such acute, transient changes may result in the permanent adaptations commonly associated with regular exercise[24].

Engaging in a single bout of physical activity can cause cardiovascular, metabolic, neuroendocrine, and immune responses. To meet the demands of exercise, sympathetic activation causes the stimulation of catecholamine release, and increases in heart rate and cardiac output. Upon cessation of a bout of exercise, increased parasympathetic activation results in more rapid heart rate recovery, which is correlated with cardiovascular fitness and is a predictor of lower cardiovascular and all-cause mortality[25]. Similarly, systolic and diastolic blood pressure are significantly reduced following a bout of moderate intensity exercise, an effect that can last for up to 12-16 hours post exercise[24]. This post-exercise hypotensive effect rivals pharmacological interventions[5]. Additional neuroendocrine responses from acute, vigorous-intensity exercise exist, including the release of cortisol and increases in brain derived neurotrophic factor[26]. However, moderate-intensity exercise may actually reduce cortisol concentrations relative to vigorous-intensity exercise and sedentary participants[27]. Acute exercise has also been investigated for its potential to modify the following day's cortisol awakening response, but this research is inconclusive[28]. Metabolic changes can also occur following acute exercise including the improvement of insulin sensitivity, reduction of triglycerides, and increases in high-density lipoprotein cholesterol[24].

Despite the extensive body of literature illustrating physical activity's benefit on health and wellbeing, approximately 80% of Canadians are not meeting the physical activity guidelines, which suggest a minimum of 150 minutes of moderate-vigorous physical activity[29]. Even when met with comprehensive cognitive behavioural treatments including exercise components, about 50% of individuals drop out of new exercise programs within the first six months[30]. Therefore, beyond the immediate physiological effects of an acute bout of exercise, the factors of a single bout of exercise that may predict greater adherence should also be investigated.

How an individual perceives an activity may influence whether they continue the activity. The Theory of Planned Behaviour suggests that one's attitude toward a behaviour (i.e., their positive or negative evaluation of that behaviour) is one factor that impacts an individual's intention to engage in that behaviour over time[31]. In other words, a more positive attitude toward exercise can predict improved future physical activity behaviours[32,33]. One factor of an individual's attitude is their level of enjoyment, which has been identified as predicting exercise adherence in empirical studies[33,34]. In turn, research conducted by Jones and colleagues[35] suggests the enjoyment of a single bout of exercise can be affected by intensity. Intensities lower than an individual's ventilatory threshold is evidenced to be perceived as significantly more enjoyable than intensities above this threshold[36]. Further, perceived intensity and enjoyment of exercise can both be moderated by extrinsic factors, without a significant change in objective physiological challenge. When engaging in attentionally dissociative activities like listening to music, participants reported significant reductions in perceived exertion and greater enjoyment of exercise[35]. In another study, the use of audiovisual stimuli was seen to significantly affect perceived exertion in cyclists when objective intensity was controlled, such that exertion was significantly lower when presented with a pleasant audiovisual stimulus[37]. Therefore, an individual's perception of intensity, enjoyment, and the extrinsic factors of their exercise environment may drive future engagement in activity. Acute studies can further allow researchers to elucidate the mechanisms underlying theories of adherence and engagement by testing the scope and magnitude of immediate responses to exercise and their reactivity to environmental perturbations.

1.2 Healing effect of nature

The 'call of the wild' refers to humans' evolutionary biophilia, which is to say humans have an innate affiliation with nature[38]. Yet, beyond our ancestral roots as hunter-gatherers, there may be some merit to the idea that humans thrive when exposed to the outdoors, particularly higher quality environments like natural spaces[6]. Early research found that surgical in-patients with a view of nature from their rooms had more positive outcomes, including shorter length of hospitalization, fewer analgesic doses, and fewer negative evaluations from nurses, than patients with a view of a brick wall[39]. A recent systematic review and meta-analysis shows that exposure to greenspace was significantly associated with a number of health-related outcomes[6]. While the review did not account for exercise, significant associations included decreased heart rate, diastolic blood pressure, salivary cortisol, high-density lipoprotein cholesterol, and autonomic control[6] (measured via low frequency and high frequency heart rate variability) . The authors also found significantly decreased risk of all-cause mortality, cardiovascular mortality, and increased self-reported health[6].

The mechanistic pathways by which outdoor exposure may improve health are still disputed, but the most popular theories point to cognitive and psychophysiological pathways that are engaged by features of the outdoors[15,17]. The Attention Restoration Theory suggests that restorative natural environments facilitate greater recovery from directed attentional depletion by placing individuals into a more effortless mode of attention[15]. Restorative environments are defined by four properties: (1) ‘being away’ or the novelty of the environmental stimuli; (2) ‘extent’ or the amount of content and structure; (3) ‘fascination’ or the ability to engage effortless attention; (4) ‘compatibility’ or the level of fit between one’s purposes and the environment’s ability to meet them[16]. Features of restorative physical environments draw an individual’s attention away from internal cues, similarly to the mechanism of attentional dissociation engaged when listening to music or viewing a video.

Taking a greater psychophysiological approach, the Stress Recovery Theory focuses on how the physical environment can foster or hamper recovery from stress, defined as the psychological and physiological responses to a situation that challenges or threatens well-being[17]. As the stress response is experienced both psychologically (i.e., coping, anxiety) and physiologically (i.e., autonomic control, cortisol), Ulrich’s Stress Recovery Theory contends that the unthreatening essence of natural scenes allows an unconscious multimodal shift toward a more positive emotional state as well as decreased physiological arousal[17]. Reductions in physiological arousal could be captured through increases in parasympathetic activity relative to sympathetic activity. This effect was demonstrated by Gladwell and colleagues[40] in their finding that views of nature result in significantly greater vagal activity compared to images of built environments. The theoretical perspective of the Stress Recovery Theory is corroborated by a meta-analysis of forest bathing studies which show significant reductions in cortisol in forest bathing groups compared to urban exposure[41]. In addition, a questionnaire study demonstrated that the effect was also detected in subjective perception: visits to a park significantly decreased pre-visit levels of perceived stress[42].

There are three levels of engagement with nature that are proposed to have positive effects on health and well-being: (1) visually in isolation, as when viewing nature through a window; (2) going about activities while in the presence of nearby nature, as in active transport or meeting friends in a park; or (3) active participation and involvement with nature, such as hiking, camping or farming[7]. The particular difference between the second and third levels is both a greater immersion in, and connectedness to, the natural space. Previous studies and reviews have attempted to focus on the effects of engaging in activity while exposed to the natural environment[8–10]. The present review contends that exercise in any outdoor space may provide the possibility for incidental exposure to high quality outdoor spaces. Additionally, previously discussed evidence points to the role that dissociation may play in perception modulation; all outdoor environments where an individual can move through space and interact with their surroundings may provide adequate external stimuli to elicit such changes. Thus, the scope of the present review will extend beyond exercise in natural environments to consider all studies of outdoor exercise in any environment.

1.3 Outdoor Exercise

The health benefits of nature exposure can be linked, in part, to the ability to use outdoor spaces for physical activity. In a systematic review of observational studies, it was found that the majority of visitors to urban green spaces use those spaces for physical activity and may exercise for longer and at higher intensities[43]. In a multi-study analysis published in 2020[11], Rogerson and colleagues identify that the relationships between greenspace and health are partially mediated by the level of physical activities of the population in question, suggesting that the more immediate benefit is not just simply from being exposed to nature, but from engaging in activity outdoors.

A common therapeutic practice for healthy living in Japan, shinrin-yoku or forest bathing, provides some additional context for the health impact of outdoor activity[44]. Shinrin-yoku involves visiting a forest, walking in the environment, mindfully attending to the environment, and breathing in the forest air[41]. In practice, the physical activity component may be coincidental, but an exercise protocol is commonly included in research methodology. While studies of forest bathing do examine relevant outcomes such as autonomic control, cardiovascular health, neuroendocrine function and human immune function, very few use laboratory-based control conditions, making conclusions about the additive effect of forest walks compared to indoor walks difficult to infer.

Similar to studies discussed above which examine the effects of views of nature, shinrin-yoku can indeed modulate autonomic control as well[44]. By measuring heart rate variability before and after walks through the forest, authors showed an increase in parasympathetic activity with a concomitant decrease in sympathetic activity[44]. Such modulation of autonomic control has the potential to moderate many of the acute, transient responses to a single bout of exercise, like heart rate, blood pressure, even neuroendocrine outputs like cortisol[41,44–46].

An acute bout of exercise, as with any physical activity, will increase heart rate in healthy individuals. A highly stable and positive linear relationship exists between exercise intensity and heart rate, whereby greater intensities of exercise will result in greater increases in heart rate.

Modifications of this relationship require physiological adaptation that is the result of regular and frequent activity. However, one's heart rate recovery following an acute bout of exercise may be more susceptible to environmental moderation. Following a light-intensity forest or urban walk[44], significantly decreased pulse rates were found in the forest group compared to the urban group. This finding further support's Ulrich's Stress Recovery Theory, whereby unthreatening natural environments result in increased parasympathetic control and subsequently, a more rapid return to baseline in physiological biomarkers, like heart rate[17].

Reduced sympathetic activity seen following forest walks may have the potential to amplify post-exercise hypotension due to the neural component of sympathetic vascular regulation[47]. In multiple studies that examined changes in blood pressure before and after walks through a forest, compared to walks in a city, significantly greater decreases in systolic and diastolic blood pressures were detected immediately following forest walks[44,48]. One study examined blood pressure three hours following the forest walk and found that significantly lower diastolic blood pressure was still evident in the forest group[48]. An early study of green exercise showed the greatest decrease in mean arterial blood pressure was elicited following exercise while viewing a pleasant rural scene, in comparison to viewing unpleasant rural and urban scenes, suggesting the potential that environment quality has as a moderating factor[7].

1.4 Previous systematic reviews and meta-analyses

The first review to synthesize the body of evidence around green exercise was completed in 2011 by Thompson Coon and colleagues[9]. This review utilized eligibility criteria that included studies that reported on outcomes of physical or mental wellbeing between indoor and outdoor exercise in adults or children. Studies in which the outdoor condition was simulated using projection or a visual reality headset while participants exercised indoors were also included. The authors identified 11 papers that met the study criteria. Despite reporting methodological concerns that made interpretation difficult, the authors identified the potential for outdoor natural environments to positively impact positive and negative emotions, feelings of energy and revitalization, enjoyment of activity, and intention to repeat outdoor activity. Unfortunately, there was a distinct absence of any discussion of the reported physiological effects and ratings of perceived exertion of outdoor exercise, despite being identified as outcomes in a number of the included articles. Further, no quantitative analysis was able to be completed due to the limited available evidence.

Intended as an update due to the expanding body of literature in the field, a more recent review and meta-analysis was completed in 2019[10]. Lahart and colleagues based the search and eligibility criteria from the 2011 systematic review, but only included articles that used experimental or quasi-experimental designs. In this case, studies which used simulated outdoor environments were included, but the analyses were separated to reflect the potential confounding effects of simulation (i.e., comparisons were made between outdoors and indoors, simulated-outdoors and indoors, and outdoors to simulated-outdoors). The greater search yield of 31 papers allowed Lahart and colleagues to meta-analyze outcomes from longitudinal trials, but not from trials which used acute bouts of exercise. Yet, the findings of this most recent review were still largely inconclusive. Quality of evidence and high risk of bias were assessed to still be problematic features of the literature to date. The authors were able to detect a statistical effect that retrospectively assessed perceived exertion was lower following a longitudinal green exercise intervention compared to an indoor exercise intervention, but there were consistent null differences in perceptions of exertion in acute-bout studies of green exercise. Affective valence and enjoyment appeared to be positively influenced by exercising in the natural environment, but all other outcomes were reported to be equivocal across conditions, including measures of emotion and mood, biological markers, exercise intensity, and performance.

These reviews, while reflective of the relative dearth of available high-quality evidence, have some methodological limitations that should be addressed. First, outdoor exercise was operationalized as exercise conducted when exposed to natural spaces. As stated above, positive effects on health have been detected in varying degrees of outdoor environments as well as varying levels of connectedness with the environment when being active in the outdoors[7]. While qualities of an environment (e.g., its greenness[12], perceived biodiversity[13,14], or potential for stress and attentional restoration[15–17]) have been suggested as moderating factors of the synergistic benefits elicited by exercising outdoors, the previously used operational definition of outdoor exercise may preclude detection of the true effect of outdoor exercise as a whole[16]. Related, included studies suffered from a common trend of reporting insufficient information to accurately estimate whether the exercise environment was truly natural. For example, articles usually limit description of the environment to a single line, so highlighting some defining natural features with little explanation is common (e.g., "an outdoor walkway in a calm garden and quiet neighborhood"[49]). Therefore, subgroups should be used in analyses to differentiate well-described, fully natural environments with few urban features from environments with a mixed natural-urban makeup to identify differences in effect sizes based on defining features of the environment.

Similarly, the inclusion of simulated outdoor exercise conducted in indoor spaces in previous reviews is considered by the present review to be problematic. The culmination of the literature suggests that connectedness to an environment is a crucial factor in predicting its salutogenic effects and feelings of 'being away', 'extent', or 'fascination' may not be fostered through simulated outdoor environments to provide the same restoration as true outdoor environments[16]. In the case of shinrin-yoku (i.e., forest bathing), it is believed that the efficacy of the intervention is tied to the complete immersion in a forest as well as the aromatherapy-like effects of breathing in the forest air[41,44,50]. Further justification is provided by an investigation of the unintended confounding effects that sensory occlusion can have on findings[51]. In a study of green exercise, significant greater mood disturbance and changes in heart rate and RPE were identified when auditory or olfactory senses were occluded in comparison to a full sensory experience of simulated green exercise[51]. This suggests that outdoor exercise that is inadequately simulated is not representative of true outdoor exercise and should not be equated.

Finally, consideration should be made to how exercise intensity is prescribed and controlled within a study. Unlike Thompson Coon and colleagues[9], Lahart and colleagues did summarize how included studies operationalized exercise intensity as being objectively (i.e., heart rate) or subjectively defined[10] (i.e., self-selected, rate of perceived exertion). However, when summarizing the findings for ratings of perceived exertion, Lahart and colleagues included studies that used subjective intensity parameters that directed the intensity of the exercise bout as part of the studies' protocols and studies that included ratings of perceived exertion as an outcome[10]. In other words, when participants are given instruction to perform exercise at a constant, normal training effort that "should be perceived as the same for the both [the] indoor and outdoor trial"[52], their perceived exertion should not be considered an outcome but a subjective-control parameter. In the six studies that reportedly found no significant differences in ratings of perceived exertion, four used a subjectively defined intensity to control the exercise protocol and measured perceived exertion at the end of the trial as manipulation checks and not actually as outcomes[10]. Therefore, the findings of consistently null perceived exertion differences reported in the 2019 review could be evidence of appropriate methodological consistency across conditions as opposed to a true effect of the environment[10].

1.5 Present Study

The present review seeks to update the body of literature found in reviews by Thompson Coon and colleagues[9] and Lahart and colleagues[10] as well as address the identified limitations that may affect synthesis and interpretation of findings. Specifically, the purpose of the present review is to (1) summarize the evidence to date which compares the physiological effects of a single bout of exercise conducted outdoors to one conducted indoors; (2) determine whether outdoor exercise provides a difference in perceptions of a single bout of exercise in comparison to indoor exercise (e.g., perceived exertion, enjoyment, future intention for exercise); (3) examine whether the effect sizes of exercise in an outdoor environment are modulated by sub-groups, such as outdoor environment quality (i.e., natural vs. mixed vs. urban) or exercise type and intensity (i.e., walking at light intensity vs. running at vigorous intensity); and (4) evaluate the current literature and determine if it is of sufficient quality and rigour to make recommendations for the utility of exercise in outdoor versus indoor spaces.

Chapter 2: Methods

This systematic review follows the PRISMA (Preferred Reporting Items for Systematic reviews and Meta-Analyses) reporting guidelines for systematic reviews[53]. Where necessary, these reporting standards were supplemented by the Cochrane Handbook for Systematic Reviews of Interventions[54]. The study protocol was registered on PROSPERO on July 24, 2018 (Registration Number: CRD42018100314). Updates to the study protocol included limiting the systematic review to studies which assess physiological changes, perceptions of intensity, and enjoyment resulting from a single bout of exercise (November 11, 2020). Future reviews will follow to summarize general psychological and affective changes that follow acute bouts of outdoor versus indoor exercise, in addition to studies that assess changes following longitudinal exercise interventions with either outdoor or indoor prescriptions.

2.1 Eligibility Criteria

To answer the research questions, the current review conducted a search for any studies that assessed how the exercise environment (*i.e.*, outdoors vs. indoors) affected physiological responses to, or perceptions of, a single bout of acute exercise. Perceptual outcomes included measures of exertion, enjoyment, or other subjective experiences directly related to the bout of exercise. Included studies were required to be peer-reviewed and published in English. Authors were required to measure outcomes during or after a single acute bout of exercise, but no criteria were implemented to limit the type, modality, or intensity of the exercise, as these elements of the exercise protocol were considered for subgroup analyses. The included studies also had to adhere to controlled designs which utilized both an outdoor exercise intervention condition as well as an indoor exercise control condition. No specification was made for the type of indoor exercise control; authors could have used a laboratory, gymnasium, fitness centre, or indoor track, for example. While previous systematic reviews have excluded trials based on outdoor exercise not being performed in a green or mostly natural environment, the present review included any studies which performed outdoor exercise, as the quality of the outdoor environment (i.e., natural, mixed natural and urban, urban) was considered for subgroup analyses. Studies in which the outdoor environment condition was substituted for a simulated condition of outdoor exercise (e.g., virtual reality, projection, sensory stimulation) were excluded from the current review. Articles in which a simulated outdoor condition was used were included in the current review if the simulated outdoor environment served as an indoor control condition to an intervention condition conducted in a true outdoor environment. Authors were required to measure at least one physiological (e.g., heart rate, blood pressure, salivary cortisol) or subjective (e.g., perception of exertion, enjoyment) outcome that was directly related to the bout of exercise. Changes in psychological processes, such as attention, memory, or mood, that may have been altered as a result of the bout of exercise were excluded from the present review and will be examined in a forthcoming review. No criteria were included to limit the population being examined as participant factors were considered for subgroup analyses; sample demographics and characteristics are described in the results of the present review.

2.2 Search Strategy

A literature search was conducted in nine electronic databases (MEDLINE, Embase, PubMed, PsycINFO, Web of Science, CINAHL, SportsDiscus, GreenFile, and CENTRAL) for any peer-reviewed articles published prior to July 2018. To update the search, a search refresh was conducted in November 2019 using the identical procedures to collect any articles published since the original search.

Search terms were selected by members of the research team, in consultation with a subject librarian and following the structure of a previous systematic review on a similar topic[9]. As seen in Table 2.1, the search strategy consisted of four major categories that were used to search titles and abstracts: (1) exercise (exercis* or physical activit* or walk* or physical fit* or run* or athlet*); (2) outdoors (outdoor* or outside* or park* or greenspace* or green space* or bluespace* or blue space* or natural environment* or natur* or forest* or biodivers* or horticultur*); (3) outdoor exercise (green exercis* or green gym* or blue exercise* or blue gym* or ecotherapy*); (4) indoors (indoor* or inside* or laboratory or gym* or home* or buil*). The search was then limited to English language articles. The formal search strategy was supplemented by hand-searching the reference lists of included articles. Two reviewers (LP, FP) conducted the searches and removed duplicates. At least two reviewers (LP, EJB, FP, MN) independently assessed titles and abstracts for eligibility. Full-text articles were reviewed independently for eligibility criteria by at least two reviewers (LP, EJB, SP, MN, FP, MP). The author of the present review (LP) or senior author (EP) were consulted in the event of disagreement between reviewers.

Line	Search Terms
1	green exercis*.ti,ab.
2	green gym*.ti,ab.
3	blue exercis*.ti,ab.
4	blue gym*.ti,ab.
5	ecotherap*.ti,ab.
6	((outdoor* or outside*) and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
7	(park* and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
8	((greenspace* or green space*) and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
9	((bluespace* or blue space*) and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
10	(natural environment* and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
11	(nature and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
12	(forest* and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
13	(biodivers* and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
14	(horticultur* and (exercis* or physical activit* or walk* or physical fit* or run* or athlet*)).ti,ab.
15	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14
16	(indoor* or inside* or laboratory or gym* or home* or buil*).ti,ab.
17	15 and 16
18	limit 17 to English language

Table 2.1 Master search strategy. Note: Master strategy was designed in Ovid MEDLINE(R) 1946 and adapted for other databases. *: indicates truncation to capture multiple suffices of the base search term. Acronyms: ti: Title; ab: abstract.

2.3 Data Extraction

Data were extracted from each article by at least two members of the research team (LP, EJB, SP, MN, FP, MP). Reviewers read each article and extracted data into three categories: (1) general study information, such as country and research setting (e.g., university); (2) study methodology, including study design, sample size, participant characteristics, exercise protocols, and descriptions of the exercise environments; and (3) outcomes and results, which describe the measures and instruments, reported results, level of significance, and overall findings of each study with effect sizes when reported. Data extraction forms were compared by the first author, and any discrepancies prompted review of the relevant studies for correction.

2.4 Risk of Bias

At least two members of the research team (LP, SB, NM) independently assessed each article for quality and risk of bias. Consensus between any incongruent decisions was made during a meeting following the blinded assessments. This process was documented through Covidence systematic review software (Veritas Health Innovation, Melbourne, Australia), an online tool designed to streamline the systematic review process.

The second version of the Cochrane Risk of Bias Assessment Tool was used to assess the bias in all randomized and non-randomized controlled trials[55]. For all crossover trials, the Cochrane Risk of Bias Assessment Tool contains an additional supplementary domain to be assessed. The risk of bias consists of 5 domains, with an additional domain for crossover trials (i.e., Domain S), which are detailed in Table 2.2. For each domain, reviewers responded to a series of signaling questions or prompts (*e.g.*, “Did baseline differences between intervention groups suggest a problem with the randomization process?”) using the following response options: ‘yes’, ‘probably yes’, ‘probably no’, ‘no’, or ‘no information’. These responses were then processed by algorithms which resulted in a risk of bias assessment of ‘high risk’, ‘some concerns’, or ‘low risk’[55]. The overall risk of bias is determined by the highest risk measured within a study, across all domains. To synthesize risk of bias across domains and studies, a bar chart and traffic light plot were created using the r package “robvis”[56].

Domain	Description
1. Risk of Bias arising from randomization process	If properly implemented, randomization should reduce the likelihood of bias from participants’ baseline characteristics to essentially zero. The prompts in this domain assessed whether the allocation sequence generation was random, whether the allocation sequence was effectively concealed, and whether baseline differences between groups existed or suggested error in the randomization process. As this domain was intended to assess bias introduced during randomization, but assumed that trials were all randomized, a risk value of high was assigned for studies that did not utilize randomization.
2. Risk of Bias due to deviation from the intended interventions	The second domain addressed whether deviations from the intended interventions existed or were likely; these deviations may lead to outcomes unrelated to the study question, and therefore, to inaccurate conclusions. The prompts in this domain addressed whether the participants and individuals delivering the interventions were aware of the assigned interventions, whether there were deviations as a result of ineffective blinding, and the likelihood that deviations could have affected the outcome. Additionally, prompts also checked whether appropriate analyses were used to estimate the effect of assignment to intervention, and if not, whether this could potentially impact the reported result.
3. Risk of Bias due to missing outcome data	The third domain assessed the impact that missing data may have had on the reported estimated effect of the intervention. Prompts assessed the magnitude of missing data, whether authors addressed and accounted for missingness, and whether data were missing at random or as a result of their true value.
4. Risk of Bias in measurement of the outcome	The fourth domain examined the degree to which measured outcomes represented their true values. The prompts in this domain questioned the appropriateness of measures, and whether measurements were obtained similarly across conditions or if knowledge of the intervention could have affected experimenters’ approaches to measurement. Outcomes were assessed independently in this domain.
5. Risk of Bias in selection of the reported result	The fifth domain addressed reporting bias that arose as a result of reporting that differed from pre-specified analysis plans. Prompts examined whether the results were selected based on multiple eligible methods or multiple eligible analyses. For crossover trials, this domain also assessed whether results were excluded on the basis of identified carryover effects.
S. Risk of Bias arising from period and carryover effects (Crossover Only)	The effectiveness of a crossover trial relies on the methodology of the trial design, otherwise period and carryover effects can eliminate the possibility of reliable results. Prompts in this domain asked about the allocation to groups, if period effects were accounted for, and if the trial design allowed for carryover effects to dissipate.

Table 2.2 Description of the Cochrane Risk of Bias Assessment Tool domains

2.5 Data Synthesis

Following data extraction, results which reported measures of central tendency and variance were converted to mean ± SD where possible. Standard error was converted according to the Cochrane Handbook by multiplying the standard error by the square root of the sample size[54]. The Cochrane Handbook method was also used to convert 95% confidence intervals; the difference between the upper and lower limit was divided by 3.92 and multiplied by the square root of the sample size[54]. Median and interquartile range were used to estimate mean ± SD using an online calculator found in the literature[57–59]. If no conversion was possible, originally reported values were preserved. If a study examined more than one experimental condition versus one control condition within a study or vice versa, the experimental groups were coded as outdoor or indoor with an additional descriptor (e.g., indoor-treadmill, indoor-track) to convey multiple comparisons within a study. Similarly, when participants were stratified by a baseline characteristic (e.g., gender, level of exercise experience), results retained these distinctions so that distinct pairwise comparisons could be completed. Where summary data were sufficiently reported, Cohen's *d* was calculated for each pairwise comparison. Effect sizes reported by authors were preserved where possible; other measures of effect (e.g., η_p^2) were included in appendices but were not discussed in results. Study findings were then combined with data collected regarding study methodology by coding the relevant study characteristics that were considered for sub-group analyses (*e.g.*, type and intensity of exercise, quality of exercise environment). Findings were then grouped and summarized by outcome for analysis.

If there were more than two comparisons which measured a single outcome in a comparable manner, a quantitative analysis was completed according to the Cochrane Handbook using RevMan Software[54,60]. Continuous outcomes were meta-analyzed using the inverse variance method with random effects to compute mean difference or standardized mean difference (Hedges *g* [95% CI], $\alpha = 0.05$). Only articles which measured outcomes in a uniform manner were compared. For example, scores collected during exercise were not compared with post-exercise scores or change from baseline scores. When comparisons were made using similar measures or scales, effects were estimated used Mean Difference, otherwise effects were expressed as Standardized Mean Difference (i.e., Hedges *g*). Evidence for heterogeneity was assessed for each quantitatively analyzed outcome using χ^2 and $p < 0.10$. To quantify whether variability in effects estimates were a result of heterogeneity as opposed to sampling error, the I^2 statistic was calculated in each meta-analyzed outcome and interpreted according to the following demarcations: 0%-40% "might not be important"; 30%-60% "may represent moderate heterogeneity;" 50%-90% "may represent substantial heterogeneity;" 75%-100% as "considerable heterogeneity"[54].

For outcomes that were only assessed by a single study or in a manner that made meta-analysis impossible, comparisons were synthesized using the method of vote counting to compare the directionality of significant effects and a summary of those findings was included. Accompanying harvest plots were created to visualize the synthesis and summary of findings, so that details about the risk of bias, outdoor environment, exercise protocol, and effect sizes could be graphically represented. In order to address some of the implicit limitations of vote counting using statistical significance, effect sizes were calculated and included in the summary where possible.

Chapter 3: Results

3.1 Search Results

The results of the comprehensive search conducted by the present review are summarized in Figure 3.1. The search strategy described above resulted in a raw yield of 47,066 articles. Other sources, including the articles reviewed by Thompson Coon and colleagues (2011) and Lahart and colleagues (2019), as well as a hand-search of relevant journals and the references of articles reviewed in full-text, yielded an additional 87 articles. Following the removal of any duplicates, 29,700 articles were screened for eligibility in title and abstract. A total of 114 articles were reviewed in full, of which 90 were excluded. 24 articles are included in the present review.

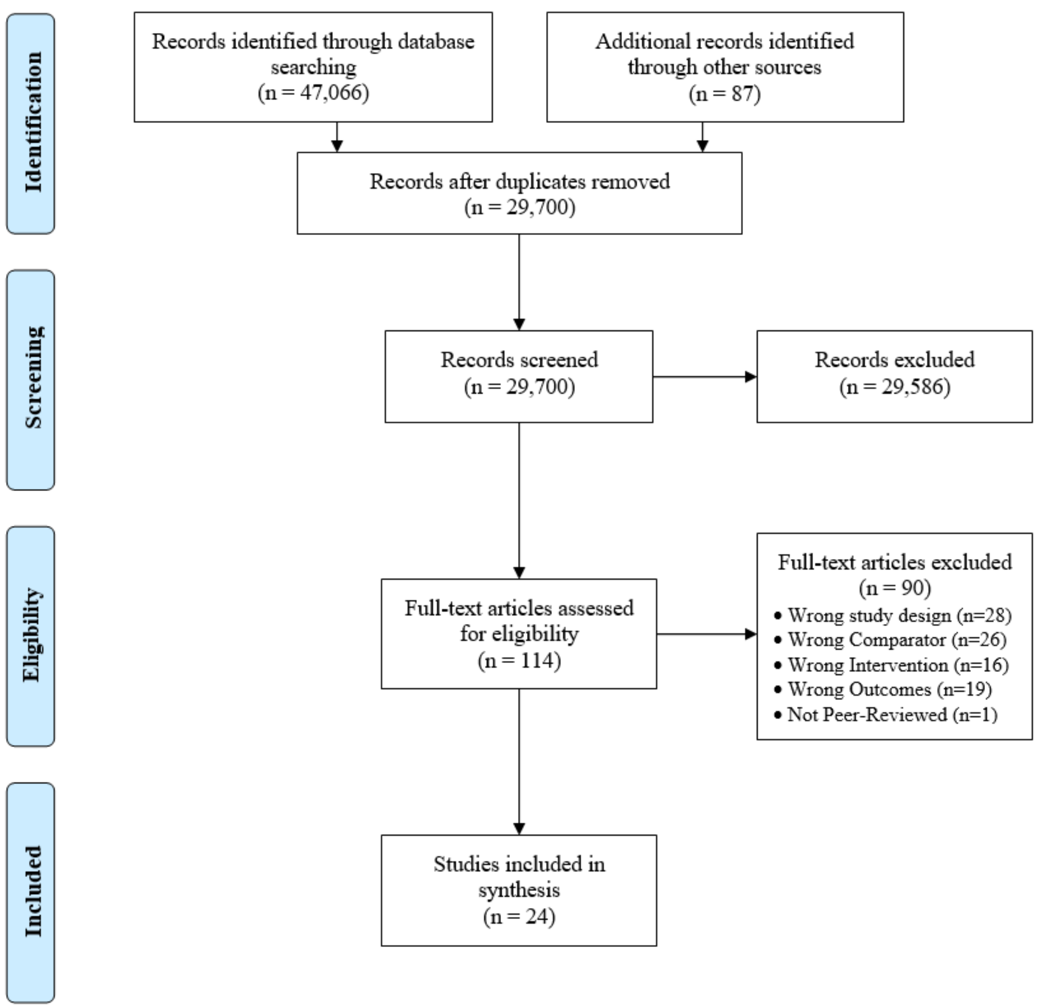


Figure 3.1 Search yield flow diagram

The primary reason for exclusion at the full-text level was due to inappropriate study designs that did not capture the effects of a single bout of acute exercise. Other reasons for exclusion included incorrect control conditions (e.g., outdoor control, sedentary control) or experimental intervention conditions (e.g., sedentary exposure to nature, virtually simulated outdoor exercise), lack of relevant outcomes, and one article that was found to not be peer-reviewed. The final result was a total of 24 articles that met the eligibility criteria of the present study. Two separate articles included distinct findings of a single trial and were assessed separately [61,62]. Some studies included more than one independent variable (e.g., gender, lower body function) or defined the exercise environment by more than two levels (e.g., outdoor, indoor treadmill, indoor track), which resulted in the reporting of multiple sets of findings that were assessed separately as distinct pairwise comparisons. Three articles stratified their included populations[49,63,64], two used more than one outdoor intervention condition[65,66], and three included more than one indoor control condition [49,67,68].

3.2 Study Design

To meet the eligibility requirements of the present review, each of the included studies was required to follow an experimental comparative or repeated-measures design with outcomes assessed during or following a single bout of exercise. Most studies utilized a repeated-measures, crossover design[49,52,70–78,61–63,65–69] (n = 19, 79%); studies either used randomization to assign participants to a counterbalanced order[49,52,61,62,65,72,75,76,78] (n = 9, 38%), counterbalanced without randomization[63,67,68,71] (n = 4, 17%), or neither counterbalanced nor randomized[66,69,70,73,74,77] (n = 6, 25%). Five studies utilized a comparative design[64,79–82] (n = 5, 21%); these studies either used randomization[64,79,80,82] (n = 4, 17%) or quasi-randomization[81] (n = 1, 4%) to assign participants to conditions.

The greatest proportion of studies were completed in Europe (*n* = 11, 46%), with studies being conducted in Poland[79], Norway[69,70], Sweden[49], Italy[71], UK[75,78,80], Austria[61,62], and Spain[76]. Eight studies (33%) were completed in North America – all of them originated in the USA[52,64,66,68,72,73,77,82]. Studies were also completed in Asia (*n* = 2, 8%), including both South Korea[81] and Taiwan[63], South America (*n* = 2, 8%), specifically Brazil[65,74], and Australia[67] (*n* = 1, 4%). An overview of the defining characteristics of the study designs of included articles is set out below, with a more detailed summary of specific studies and their study designs presented in 4.4Appendix A

3.2.1 Participants

The 24 included articles consisted of 23 unique participant samples. In total, 757 participants were included in the 23 samples. Sample sizes ranged from five participants to 112 participants, with a mean ± SD of 32.9 ± 26.9. Ten studies (42%) utilized small sample sizes of less than 20 participants[52,65–67,69,73,74,76,77,81], and only two studies (8%) utilized sample sizes greater than 80 participants[64,82]. More than half of the studies (*n*=13, 54%) reported conducting a power calculation to determine sample size[52,61,80,81,62,65,66,71,72,76,78,79], only six of which (25%) reported *a priori* sample size calculations[52,61,62,72,78,80]. Two of these 13 studies (15%) conducted analysis on fewer than required participants to meet a statistical power of greater than 80% for their anticipated effect size[62,80].

The aggregate sample were predominantly female (57%), resulting in 434 female participants and 323 male participants (43%). Eight studies (33%) included a majority of female participants[64–66,68–70,75,79] and three (13%) included completely female samples[72,77,82], whereas seven studies (29%) utilized a sample of at least 50% male participants[49,61,62,76,78,80,81] and five

(21%) utilized completely male samples[52,63,67,73,74]. A single study (4%) utilized an evenly distributed male and female sample[71].

The weighted mean ± SD age of the included participants was 29.36 ± 1.65 years, excluding the three studies that did not adequately report mean and standard deviations[64,67,73]. The majority of distinct samples (*n* =13, 57%) included participants with a mean age of between 19 and 30[63,66,80–82,67,68,70–72,74,76,79], while five (22%) had an average participant age between 30 and 50[52,61,62,69,75,78], and three studies (13%) had an average participant age over 50[49,65,77].

Within the 23 unique participant samples, the majority of the participants were described as healthy[69–71,75,77] (*n* = 5, 22%), active[61–63,68,72,76] (*n* = 5, 22%), or trained athletes[52,67,73,74,78,79,81] (*n* = 7, 30%). A small portion of samples were recruited from specific populations, like patients who had recently recovered from a stroke (*n* = 1, 4%)[49], postmenopausal women (*n* = 1, 4%)[77], or adults over 60 (*n* = 1, 4%)[65]. Samples also included university students (*n* = 6, 26%)[63,64,72,76,80,82], students and employees of a university (*n* = 1, 4%)[70], and employees in the workplace (*n* = 1, 4%)[69]. A single study[66] (4%) did not describe their sample in adequate detail to report here. Four (17%) studies stratified their samples for analysis following recruitment, on the basis of baseline characteristics such as gender (*n* = 2, 9%)[64,71], clinical walking speed (*n* = 1, 4%)[49], and exercise experience (*n* = 1, 4%)[63]. For the purpose of analysis in the present review, each was treated as a unique comparison where summary data was provided in enough detail.

3.2.2 Exercise Environment

As stated in the eligibility criteria strategy, all studies were required to include an outdoor exercise protocol as the intervention condition and an indoor exercise protocol as the control condition. A small portion of studies collected data from more than one intervention condition[65,66] (e.g., beach and outdoor track; *n* = 2, 8%) or more than one control condition[49,67,68] (e.g., treadmill and indoor track; *n* = 2, 8%). In these cases, each pairwise comparison between the unique intervention and control conditions (e.g., green space vs treadmill, blue space vs treadmill) was treated as unique. Therefore, within the 24 studies, measurements were conducted in 28 outdoor exercise environments and 26 indoor exercise environments.

The 28 outdoor exercise environments were characterized as either mostly natural with very few urban features[52,61,80,81,62,65,66,69,75,77–79] (i.e., green, blue, brown, mountainous; *n* = 14, 50%), mostly urban[66,68] (*n* = 2, 7%) or a mix of natural and urban features[49,63,76,82,64,65,67,70–74] (i.e., green/urban, blue/urban, university campus, outdoor track; *n* = 12, 43%). The 26 indoor exercise environments were primarily laboratory spaces[52,64,75–77,80,66–73] (*n* = 14, 54%), but other indoor environments included traditional exercise spaces[49,61,62,65,68,78,79,81,82] (i.e., fitness centres, gymnasiums, studios, indoor tracks; *n* = 9, 35%) and a health-care clinic[49] (*n* = 1, 4%). Two studies did not describe the indoor environment in adequate detail to report here[63,74]. More detailed descriptions and intra-study comparisons of exercise environments are available in 4.4Appendix A

3.2.3 Exercise Protocol

Exercise protocols can vary in multiple ways, including time spent exercising, type of exercise (e.g., running vs. cycling), modality of exercise (e.g., treadmill vs. over-ground), and intensity of exercise (e.g., required intensity within a certain range vs. self-selected). One study used multiple modalities of indoor exercise by including both a treadmill in a laboratory and running in an indoor track environment as control conditions[68]. These were counted as separate protocols for the purpose of this review and pairwise comparisons were completed relative to the intervention condition that was completed in an outdoor environment. Eight studies (33%) implemented exercise protocols that were identical across conditions within each respective study[49,65,66,75,76,79–81]. The majority of the remainder (*n* = 13, 54%) included exercise protocols that were matched for exercise type and intensity and only differed in modality[52,63,77,78,82,64,68–74]. Two studies (8%) used a distance-based, mountain hiking protocol for the outdoor condition with an indoor protocol that utilized time-based, inclined treadmill walking[61,62]. One study (4%) employed a distance-based running outdoor protocol in which intensity was constant paired with a time-based, treadmill running indoor protocol in which intensity was variable[67]. More complete descriptions and intra-study comparisons of exercise protocols are available in 4.4Appendix A .

3.2.3.1 Time spent exercising

The majority of studies (*n* = 15, 63%) measured the length of protocol by time; all 15 protocols ranged from 10 minutes to 60 minutes[63,64,77,79–82,65,66,69–72,74,75]. Four studies (17%) used distance-based protocols for both the outdoor and indoor exercise, ranging from 5 km to 40 km[52,68,73,78]. Three studies (13%) used a 12 km distance-based protocol when participants exercised outdoors and time-based indoor protocols that were approximately matched for the duration of the outdoor exercise[61,62,67] (i.e., 45 minute run; 90 minute walk). One study (4%) utilized an incremental and constant power protocol to exhaustion, so duration was highly variable[76]. Finally, one study (4%) examined two clinically relevant protocols using both time (6-minute walk test) and distance (30-meter walk test).

3.2.3.2 Type and modality of exercise

All but two articles (*n* = 22, 92%) utilized protocols for both indoor and outdoor activities that were matched by exercise type[49,52,71–80,63,81,82,64–70]. The remaining two articles (8%) had participants hike during the outdoor condition and walk on an inclined treadmill for the indoor condition[61,62]. Walking was the most frequent type of exercise used in the included articles[49,61,82,62,64–66,70–72,77] (*n* = 11, 46%). Other exercise types were running[63,67,68,73,74,78] (*n* = 6, 25%), cycling[52,69,75,76,80] (*n* = 5, 21%), taekwondo[81] (*n* = 1, 4%), and dancing[79] (*n* = 1, 4%). Given the difference in muscular activity when engaging in stationary exercise (e.g., running on a treadmill or using a cycle ergometer) compared to travelling overground, type of exercise must be contextualized with the selected modality of exercise. In 19 studies (79%), participants engaged in activities in which they travelled overground in the outdoor condition[49,52,69–74,77,78,82,61–68], as opposed to using an ergometer[75,76,80] (*n* = 3, 13%) or engaging in activities like dance[79] (*n* = 1, 4%) or taekwondo[81] (*n* = 1, 4%). Of these 19 studies, 16 (84%) utilized a treadmill or cycle ergometer to simulate exercise in the indoor condition[52,61,72–74,77,78,82,62–64,67–71] and one article used both an overground and treadmill indoor control protocol[68].

3.2.3.3 Intensity of exercise

As there is a positive dose-response identified between intensity and elicited health benefit, it is important to identify how authors accounted for intensity in their respective protocols[83,84]. Intensity of an exercise protocol can be prescribed based on either objective parameters (e.g., heart rate reserve) or feelings of exertion[83]. Authors of the reviewed studies prescribed intensity of the exercise protocol using one of two approaches: (1) *objective intensity*: researchers instructed participants to exercise at a pre-specified level of intensity based on objective physiological parameters[64,67,69,73,75,80,82] (e.g., percentage of heart rate reserve, percentage of maximal oxygen consumption; *n* = 7, 29%) or (2) *subjective intensity*: researchers instructed participants to exercise at a level of intensity based on subjective parameters[49,52,72,74,76–79,81,61–63,65,66,68,70,71] (i.e., using values of the Borg rating of perceived exertion scale or descriptions of exertion like "brisk without overspending"[61,62]; *n* = 17, 71%). Regardless of the intensity-control approach, all studies designed exercise protocols that were matched for intensity across outdoor and indoor environments.

Where possible, intensity levels of exercise protocols have been classified using evidence-based relative exercise intensity zones[83], as seen in Table 3.1. Nine studies (38%) did not provide enough information to appropriately classify exercise intensity into these intensity zones as participants were asked to exercise at a self-selected[65,66,68,71,79] or comfortable self-selected intensity[70,77]. Two of these nine (8%) prescribed intensity using the description of an approximate normal training effort in a sample that consisted of athletes[52,81]. Most of the remaining studies' protocols were classified as a moderate level of intensity, either using objective parameters[64,75,80,82] (e.g., 50% heart rate reserve; n = 4, 17%) or subjective parameters[61–63,72] (e.g., Rating of Perceived Exertion (RPE) on Borg 6-20 scale of 15, moderate self-selected intensity; n = 4, 17%). Other studies conducted exercise protocols at a vigorous level of intensity, via objective parameters[67,69,73] (n = 3, 13%) or subjective parameters[74,76] (n = 2, 8%), or at a variable exertion associated with vigorous to near maximal intensity[78] (n = 1, 4%). Finally, one study (4%) used a near-maximal to maximal intensity during a clinically relevant exercise protocol in addition to a test at self-selected subjective intensity during another exercise protocol[49].

Two studies of the seven (29%) that controlled the exercise protocol by using objective intensity parameters reported participants' actual intensity levels (percentage of heart rate reserve or mean oxygen consumption), which can be used to confirm that exercise intensity was adequately controlled between environments[69,73]. As the harvest plot in Figure 3.2 illustrates, there were no significant physiological differences between the outdoor experimental condition and indoor control condition in either study.

Intensity	Perceived Exertion (Rating on Borg 6-20 RPE Scale)	%HRmax	%HRR	%VO2Max	%VO2R
Very light	< Very Light (RPE < 9)	<57	<30	<37	<30
Light	Very light to fairly light (RPE 9-11)	57-63	30-39	37-45	30-39
Moderate	Fairly light to somewhat hard (RPE 12-13)	64-76	40-59	46-63	40-59
Vigorous	Somewhat hard to very hard (RPE 14-17)	77-95	60-89	64-90	60-89
Near-maximal to maximal	≥Very hard (RPE ≥ 18)	≥96	≥90	≥91	≥90

Table 3.1 Zones of relative intensity for cardiorespiratory exercise. Note: Table adapted from the American College of Sports Medicine[83]. Acronyms: RPE: Rating of Perceived Exertion; HR: Heart Rate; HRR: Heart Rate Reserve; VO2Max: Maximal oxygen consumption; VO2R: Oxygen consumption reserve.

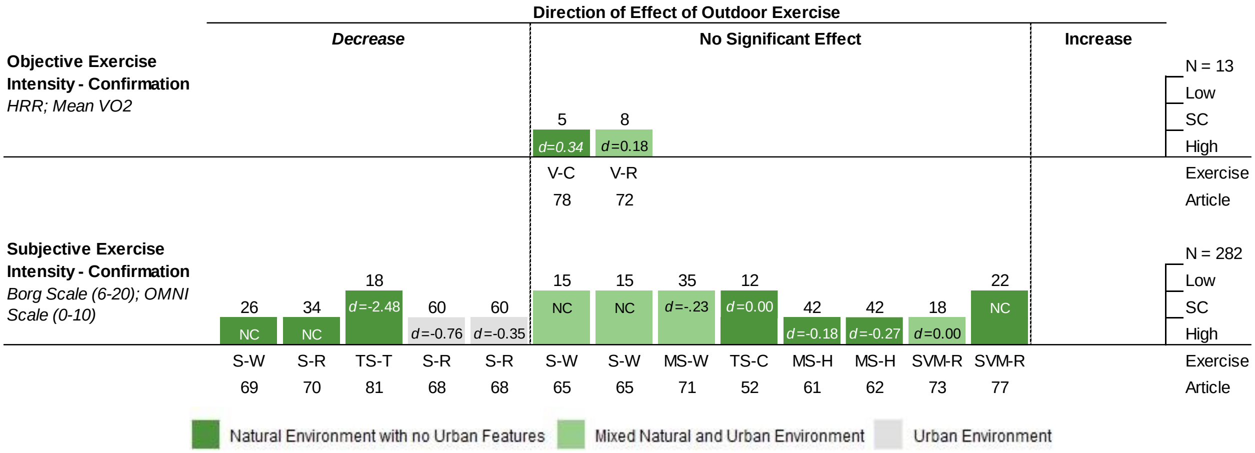


Figure 3.2 Harvest plot of direction of effect in objective intensity and subjective intensity of study protocol between outdoor and indoor environments. Note: The number above each column indicates the total sample size for the study. The height of each column represents overall risk of bias of the study, with taller columns representing less risk of bias. Within the column, the Cohen's *d* represents effect size. Under each column, exercise intensity and type is indicated below each column and separated by a hyphen with the article reference number below. Acronyms: HRR: Heart Rate Reserve; VO2: Oxygen consumption; V: Vigorous intensity; C: Cycling; R: Running; NC: Not calculated due to insufficient reporting; S: Self-selected intensity; W: Walking; TS: Self-selected training intensity; MS: Moderate self-selected intensity; MI: Moderate intensity; H: Hiking; SVM: Self-selected vigorous to maximal intensity.

Alternately, 11 of the 17 studies (65%) that used subjective intensity parameters to control the intensity of the exercise protocol also reported participants' actual subjective intensity levels during or at the end of the bout[52,61,81,62,65,68,70–72,74,78]. Figure 3.2 illustrates that five of the eleven studies (thirteen comparisons as two studies each included two pairwise comparisons due to multiple outdoor intervention conditions or multiple indoor control conditions) found significant differences between the indoor and outdoor environments[68,70,71,81]. These five comparisons resulted in lower levels of subjective intensity in the outdoor condition compared to the indoor condition. This suggests that researchers inadequately controlled for participants' subjective intensity across conditions, which could confound subsequent findings as the intensity of an exercise protocol has been positively associated with greater physiological effects of exercise[83,84].

3.3 Risk of Bias

All studies were assessed using the same tool, regardless of study design, and thus can be summarized together. Studies which employed a comparative design were not assessed and designated as “no information” for Domain S: Risk of bias arising from period and carryover effects. Figure 3.3 provides a summary of the risk of bias assessments by domain. A more detailed summary, including reviewer decisions by article and domain, is available in 4.4Appendix B . Overall risk of bias of respective studies has also been incorporated into the harvest plots as the height of the columns.

Overall, there were no studies that were assessed to have a low risk of bias. More than half of the studies were assessed to have a high risk of bias[49,61,73,74,77,62,63,66–71] (n = 13, 54%) and the remaining 11 studies[52,64,82,65,72,75,76,78–81] (46%) were assessed to have some concerns. Since the overall rating of risk of bias is assigned to be the highest level of risk assessed within an article, more detail can be gleaned from looking at each domain in detail.

Risk of bias in the randomization process was assessed for all articles in Domain 1, which included those that did not include randomization in the study design and were thus assigned a high risk of bias for this domain[63,66–71,73,74,77] (n = 10, 46%). Of these 10 articles, six articles (25%) used a crossover design[66,69,70,73,74,77] and four articles (17%) employed a counterbalanced crossover design[63,67,68,71]. Four articles (17%) were assessed to have some concerns about the

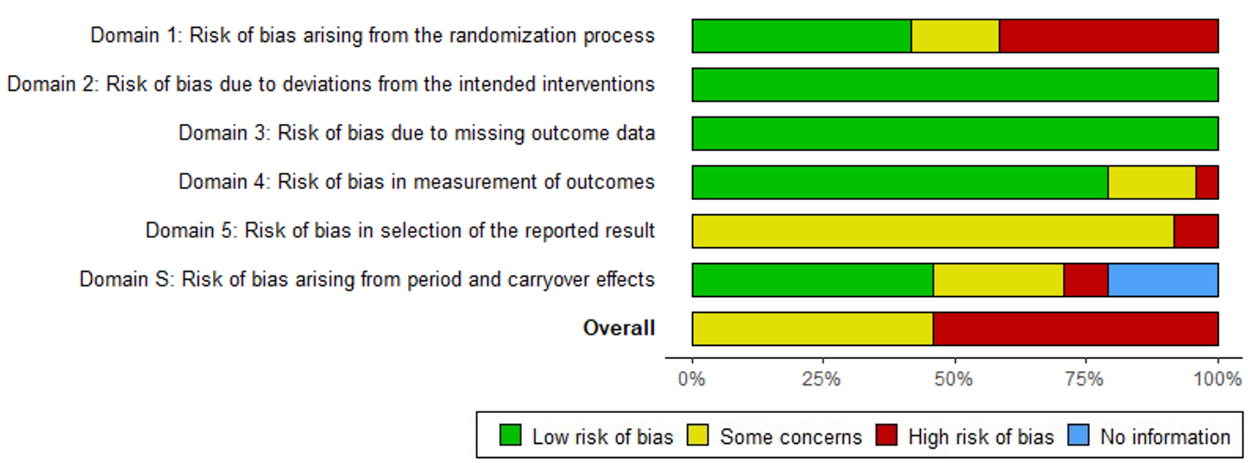


Figure 3.3 Summary of risk of bias assessments

randomization process[52,76,81,82]. Of these four articles, three (13%) were designated as ‘some concerns’ due to not providing any information about allocation concealment[52,76,82] and one (4%) used a quasi-random comparative design that randomized based on baseline concentration scores[81]. The remaining 10 articles (42%) were assessed to have low risk of bias that arose from the randomization process[49,61,62,64,65,72,75,78–80].

For both Domain 2 (risk of bias due to deviations from the intended intervention) and Domain 3 (risk of bias due to missing outcome data), all studies (n = 24, 100%) were assessed to have a low risk of bias. The signaling questions in Domain 2 begin by inquiring about whether participants and researchers were blinded – a feat which would be nearly impossible to achieve in a study of this nature. However, no deviations from the intervention were reported; authors consistently reported similar experimental conditions between the outdoor condition and indoor condition.

The majority of studies (n = 17, 71%) were assessed to have a low risk of bias for arising from the measurement of their evaluated outcomes[52,61,74–82,62–64,66,67,69–71] (Domain 4). Four studies (17%) had at least one outcome that resulted in some concerns about the risk of bias due to how they measured time or speed[49,65,68,72], and only one study (4%) registered a high risk of bias due to measuring oxygen consumption using two different methods[73] (i.e., open circuit spirometry when indoors and Douglas Bag method when outdoors).

All but two articles (n = 22, 92%) omitted a pre-specified analysis plan and were therefore determined to have ‘some concerns’ about the risk of bias from selection of the reported result[49,52,71–80,63,81,82,64–70] (Domain 5). However, the two articles (8%) that did include a pre-specified analysis plan as supplementary material with their publication did not conduct analysis in accordance with the pre-registered protocol and were therefore deemed to have a high risk of bias[61,62].

Only 19 articles were assessed for risk of bias arising from period and carryover effects (Domain S), as the remaining five articles followed a comparative design as opposed to a crossover design[49,52,70–78,61–63,65–69]. Most studies (*n* = 11, 58%) were assessed a low risk of bias for this domain[52,61,78,62,63,65,68,71,72,75,76]. There were 'some concerns' about the risk of bias in five studies[66,69,73,74,77] (26%) due to order effects not being considered methodologically or statistically, and in one study[67] (5%) due to the possibility that there may have been carryover effects to methodology (i.e., four 12 km runs separated by 24 hours). Two studies[49,70] (11%) were assessed a high risk of bias due to providing insufficient washout periods where carryover effects were likely.

3.4 Reported Findings

The 24 studies included in this systematic review reported on outcomes from at least one of the following nine groupings: (1) objective exercise intensity[52,61–63,70–74,79] (*n* = 10, 42%), (2) perceived exertion[67,69,73,75,80] (*n* = 5, 21%), (3) performance[49,52,80,65,68,70,71,74,76–78] (*n* = 11, 46%), (4) neuroendocrine response[62,67,73] (*n* = 3, 13%), (5) cardiovascular response[61,62,66,67] (*n* = 4, 17%), (6) thermoregulation[52] (*n* = 1, 4%), (7) enjoyment[64,67,69,70,72,75,82] (*n* = 7, 29%), (8) future exercise intentions[72,75] (*n* = 2, 8%), and (9) perception about the environment[80,81] (*n* = 2, 8%). More detailed descriptions of the reported findings, including means ± SD and effect sizes, are available in 4.4Appendix C .

3.4.1 Objective Exercise Intensity

Participants' objective exercise intensity was measured as an outcome in 10 studies[52,61–63,70–74,79]. Three studies assessed objective intensity using two measures[70,71,74]. In total, 14 unique comparisons were made between conditions. Objective exercise intensity was assessed using six different assessment techniques, including average heart rate[52,61,62,70,73,74] (n = 6, 47%), percentage of participants' heart rate reserve[63,71] (n = 3, 20%), maximal heart rate achieved during physical activity[70,74] (n = 2, 13%), percentage of participants' maximal heart rate[72] (n = 1, 7%), percentage of VO2 reserve[71] (n = 1, 7%), and time spent engaging in moderate-vigorous physical activity[79] (n = 1, 7%). With one exception that used relative oxygen consumption to control exercise protocol[73], all comparisons of objective exercise intensity as the outcome were completed in studies that used subjective intensity to control exercise protocol.

Five of six studies included summary data for average heart rate with sufficient homogeneity to conduct a quantitative analysis[52,61,62,70,73,74] (see Figure 3.4). In the pooled analysis, an outdoor exercise environment did not significantly affect average heart rate (Mean difference = 3.11, 95% CI = [-2.68, 8.90], p = 0.29). Significant heterogeneity was detected in this analysis (I^2 = 76%, χ^2 = 16.78, df = 4, p = 0.002). The only subgroup analysis that could be completed was by exercise environment type. In the pooled analysis of studies using natural outdoor environments

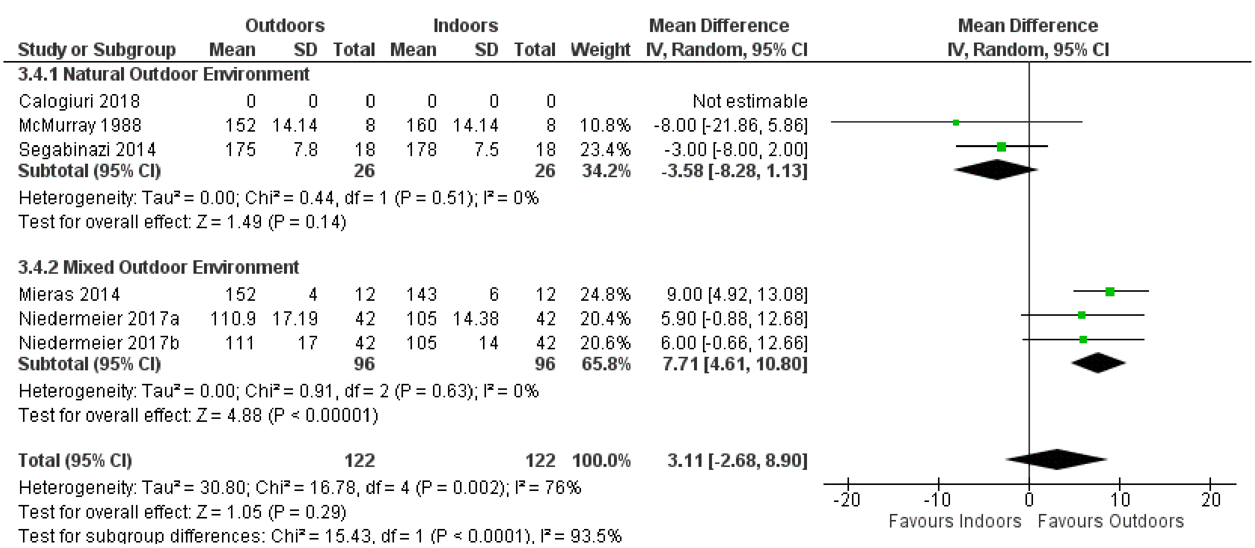


Figure 3.4 Forest plot of mean difference in average heart rate between outdoor and indoor environments. Subgroups are based on quality of outdoor environment.

(see Figure 3.4.1), the exercise environment did not significantly affect average heart rate during exercise (Mean difference = -3.58, 95% CI = [-8.28, 1.13], $p = 0.14$) and no heterogeneity was detected ($I^2 = 0\%$). In the pooled analysis of studies using mixed outdoor environments (see Figure 3.4.2), the exercise environment significantly raised average heart rate during exercise (Mean difference = -7.71, 95% CI = [4.61, 10.80], $p < 0.001$) and no heterogeneity was detected ($I^2 = 0\%$). The subgroup analysis detected significant difference in effects between subgroups ($I^2 = 93.5\%$, $\chi^2 = 15.43$, df = 1, $p < 0.001$).

Comparisons for all other objective intensity measurements are depicted in the harvest plot in Figure 3.5. Four of eight comparisons included in the harvest plot (50%) detected significantly higher objective exercise in the outdoor environments[63,74,79], two comparisons (25%) detected significantly higher objective intensity in the indoor environment[71], and two comparisons (25%) detected no significant effect[70,72]. However, both findings that indoor environments elicited higher intensity exercise and one of the findings that showed no effect were from a study that was unable to properly control for subjective intensity as described above in section 3.2.3.3[71]. In this study, participants' subjective intensity was greater in the indoor condition and may explain the elevated objective exercise intensity[71].

Each of the studies included in these syntheses used subjective parameters to control for exercise intensity. Therefore, when comparisons demonstrate significant effects that cannot be explained by poor control of subjective intensity, it can be reasoned that the exercise environment may be causing a dissociation between subjective and objective exercise intensities. The decoupling of these variables could result in greater physiological benefits of exercise at equivalent subjective intensity due to the positive dose-response of exercise intensity[83,84].

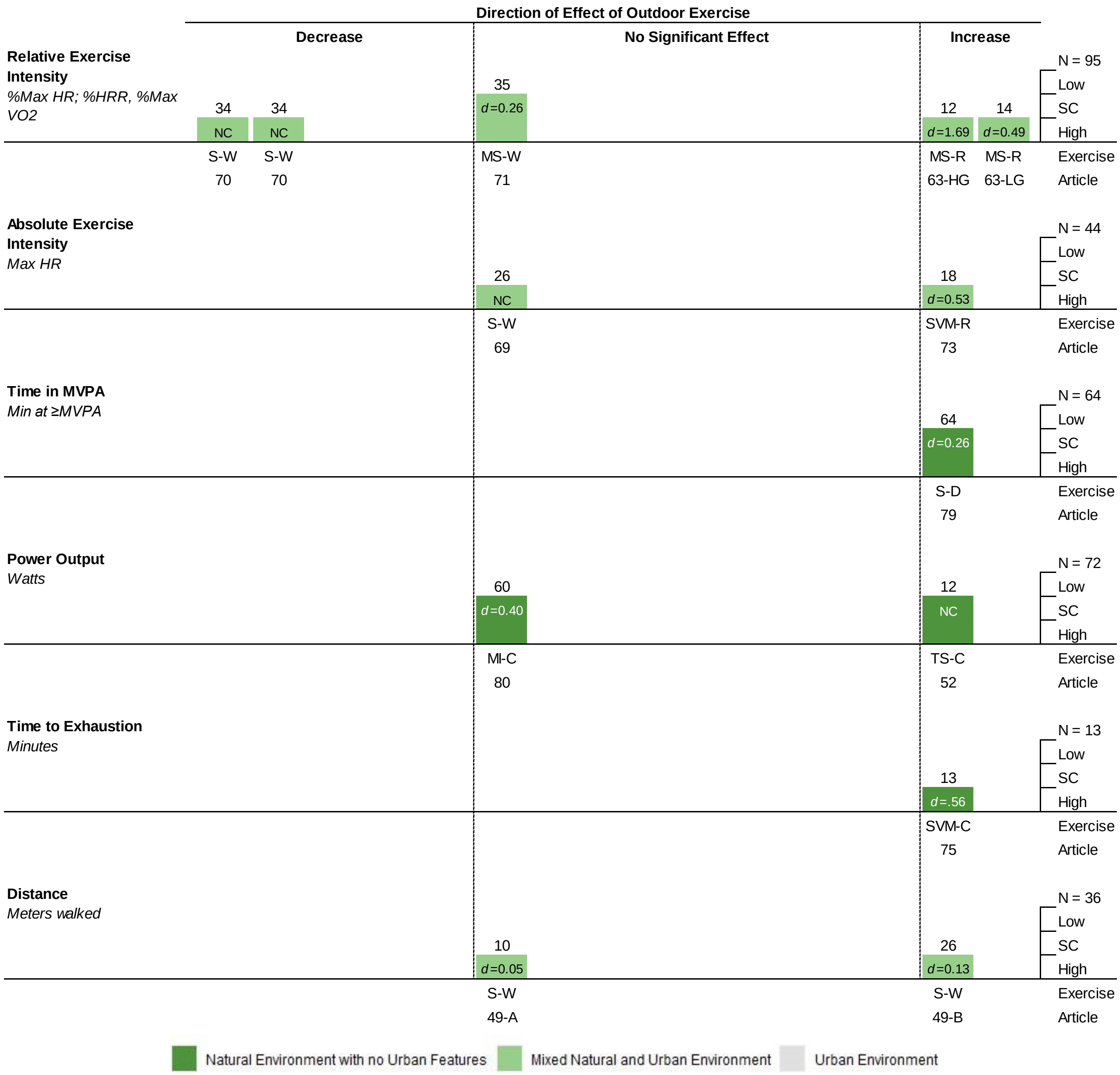


Figure 3.5 Harvest plot of direction of effect in objective exercise intensity (relative and absolute exercise intensity, time in MVPA) and performance (power output, time to exhaustion, distance) between outdoor and indoor environments. Note: The number above each column indicates the total sample size for the study. The height of each column represents overall risk of bias of the study, with taller columns representing less risk of bias. Within the column, the Cohen's *d* represents effect size. Under each bar, exercise intensity and type is indicated below each column and separated by a hyphen with the article reference number below. Acronyms: HR: Heart Rate, HRR: Heart Rate Reserve; VO2: Oxygen consumption; NC: Not calculated due to insufficient reporting; S: Self-selected intensity; W: Walking; MS: Moderate self-selected intensity; R: Running; HG: High exercise experience group (stratified sample); LG: Low exercise experience group (stratified sample); SVM: Self-selected vigorous to maximal intensity; D: Dancing; MI: Moderate intensity; C: Cycling; TS: Self-selected training intensity; VM: Vigorous to maximal intensity.

3.4.2 Perceived Exertion

Perceived exertion was measured as an outcome in 5 studies[67,69,73,75,80], using two measures: the Borg 14-point RPE scale[69,73,75,80] (*n* = 4, 83%) and the Borg 9-point RPE scale[67] (*n* = 1, 17%). There was sufficient homogeneity to conduct a quantitative analysis (see Figure 3.6). One study conducted two indoor control conditions using external or internal auditory stimuli while exercising on a treadmill, and these have been included as separate comparisons. In the pooled analysis, an outdoor exercise environment was found to significantly reduce perceived exertion (Hedges *g* = -0.84, 95% CI = [-1.60, -0.09], *p* = 0.03). Significant heterogeneity was detected in this analysis (I^2 = 75%, χ^2 = 19.81, df = 5, *p* = 0.001). The only subgroup analysis that could be completed was by environment type. In the pooled analysis of studies using natural outdoor environments (see Figure 3.6.1), the exercise environment did not significantly affect perceived exertion during exercise (Hedges *g* = -0.41, 95% CI = [-0.90, 0.08], *p* = 0.10) and no significant heterogeneity was detected (I^2 = 18%). In the pooled analysis of studies using mixed outdoor environments (see Figure 3.6.2), the exercise environment did not significantly affect

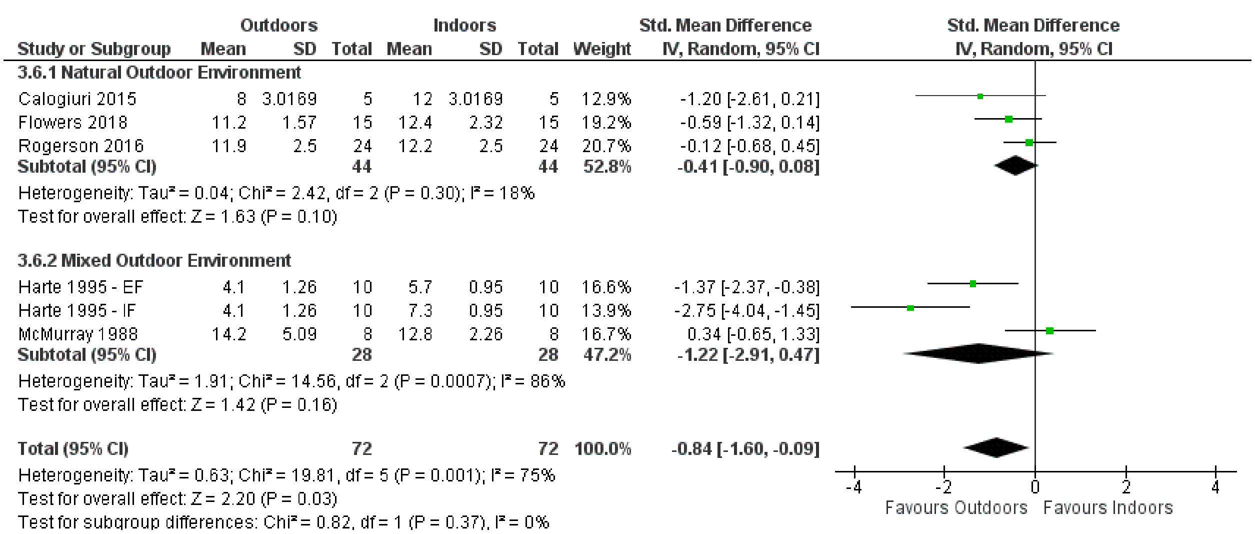


Figure 3.6 Forest plot of standardized mean difference in perceived exertion between outdoor and indoor environments. Subgroups are based on quality of outdoor environment.

perceived exertion (Hedges g = -1.22, 95% CI = [-2.91, 0.47], p = 0.16) and significant heterogeneity was detected (I^2 = 86%, χ^2 = 14.56, df = 2, p = 0.0007). The subgroup analysis detected no significant difference in effects between subgroups (I^2 = 0%, χ^2 = 0.82, df = 1, p = 0.37).

Each of the studies included in the meta-analysis in Figure 3.6 used an objective parameter to control the intensity of the exercise protocol across conditions. Therefore, the detection of a negative overall effect suggests the reciprocal of the relationship described in section 3.4.1 above: at the same level of objective intensity, participants reported that the bout of exercise felt easier. While this may be the case, three of five studies did not report participants' actual levels of objective intensity[67,75,80]. In the two studies that did, both were found to have non-significant differences but the directions of effects trended toward greater objective intensity outdoors[69,73].

3.4.3 Performance

Exercise performance was measured as an outcome by 11 studies[49,52,80,65,68,70,71,74,76–78] over 23 unique comparisons, using 5 types of measures: average speed[49,65,70,71,74,77] (n = 14, 61%), power output[52,80] (n = 2, 9%), time to completion[52,68,78] (n = 3, 17%), time to exhaustion[76] (n = 1, 4%), and distance completed[49] (n = 2, 9%). There was only sufficient homogeneity to conduct a quantitative analysis on average speed (see Figure 3.7) and time to completion (see Figure 3.8).

Summary data for average speed was adequately reported to calculate an effect size in four of six studies[49,65,70,71,74,77] (one study provided two comparisons due to two outdoor intervention conditions[65], one study provided eight comparisons due to two stratified sample groups, two indoor control conditions, and two exercise protocols[49]). In the pooled analysis (see Figure 3.7), the exercise environment did not significantly affect average speed (Hedges g = 0.09, 95% CI = [-0.11, 0.29], p = 0.87). No significant heterogeneity was detected in this analysis (I^2 = 6%). The only subgroup analysis that could be completed was by exercise environment type. In the pooled analysis of studies using natural outdoor environments (see Figure 3.7.1), the exercise environment did not significantly affect average speed during exercise (Hedges g = 0.55, 95% CI = [-0.51, 1.61], p = 0.14) and significant heterogeneity was detected (I^2 = 78%, χ^2 = 4.57, df = 1, p = 0.03). In the pooled analysis of studies using mixed outdoor environments (see Figure 3.7.2), the exercise environment did not significantly affect average speed (Hedges g = -0.00, 95% CI = [-0.20, 0.21], p = 0.97) and no heterogeneity was detected (I^2 = 0%). The subgroup analysis detected no significant difference in effects between subgroups (I^2 = 0%, χ^2 = 0.97, df = 1, p = 0.33).

Summary data for time to completion was reported in sufficient detail to allow for calculation in all three studies[52,68,78] (one study provided two comparisons due to two indoor control conditions[68]). In the pooled analysis (see Figure 3.8), the exercise environment did not significantly affect time to completion (Hedges g = -0.82, 95% CI = [-1.66, 0.02], p = 0.05).

	Outdoors			Indoors				Std. Mean Difference
Study or Subgroup	Mean	SD	Total	Mean	SD	Total	Weight	IV, Random, 95% CI
3.7.1 Natural Outdoor Environment								
Farias 2018 BW	1.4	0.2	15	1.4	0.2	15	7.3%	0.00 [-0.72, 0.72]
Teas 2007	5.1	0.83	19	4.3	0.6	19	8.0%	1.08 [0.40, 1.77]
Subtotal (95% CI)			**34**			**34**	**15.3%**	**0.55 [-0.51, 1.61]**
Heterogeneity: Tau² = 0.46; Chi² = 4.57, df = 1 (P = 0.03); I² = 78%								
Test for overall effect: Z = 1.01 (P = 0.31)								
3.7.2 Mixed Outdoor Environment								
Calogiuri 2018	0	0	0	0	0	0		Not estimable
Carvalho 2010 (A) NM, BH	0.6	0.25	10	0.6	0.23	10	5.0%	0.00 [-0.88, 0.88]
Carvalho 2010 (A) NM, CL	0.6	0.25	10	0.61	0.24	10	5.0%	-0.04 [-0.92, 0.84]
Carvalho 2010 (A) S, BH	0.51	0.2	10	0.52	0.19	10	5.0%	-0.05 [-0.93, 0.83]
Carvalho 2010 (A) S, CL	0.51	0.2	10	0.51	0.18	10	5.0%	0.00 [-0.88, 0.88]
Carvalho 2010 (B) NM, BH	1.69	0.33	26	1.68	0.29	26	12.3%	0.03 [-0.51, 0.58]
Carvalho 2010 (B) NM, CL	1.69	0.33	26	1.66	0.34	26	12.3%	0.09 [-0.46, 0.63]
Carvalho 2010 (B) S, BH	1.31	0.22	26	1.3	0.23	26	12.3%	0.04 [-0.50, 0.59]
Carvalho 2010 (B) S,CL	1.31	0.22	26	1.26	0.21	26	12.2%	0.23 [-0.32, 0.77]
DaSilva 2011	0	0	0	0	0	0		Not estimable
Farias 2018 M	1.4	0.2	15	1.4	0.2	15	7.3%	0.00 [-0.72, 0.72]
Segabinazi 2014	11.8	0.8	18	12.2	0.8	18	8.5%	-0.49 [-1.15, 0.18]
Subtotal (95% CI)			**177**			**177**	**84.7%**	**0.00 [-0.20, 0.21]**
Heterogeneity: Tau² = 0.00; Chi² = 2.92, df = 9 (P = 0.97); I² = 0%								
Test for overall effect: Z = 0.04 (P = 0.97)								
Total (95% CI)			**211**			**211**	**100.0%**	**0.09 [-0.11, 0.29]**
Heterogeneity: Tau² = 0.01; Chi² = 11.65, df = 11 (P = 0.39); I² = 6%								
Test for overall effect: Z = 0.87 (P = 0.38)								
Test for subgroup differences: Chi² = 0.97, df = 1 (P = 0.33), I² = 0%								

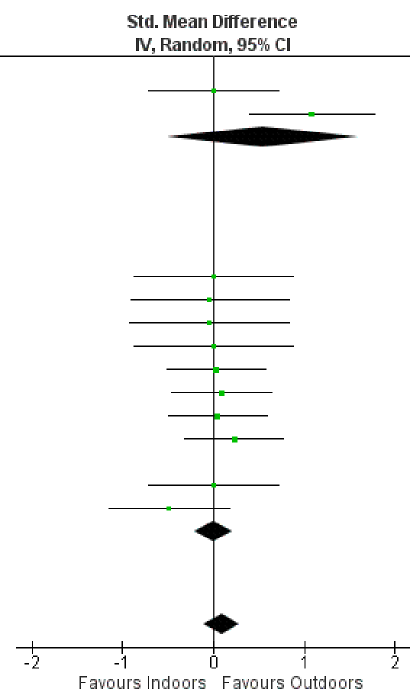


Figure 3.7 Forest plot of standardized mean difference in average speed between outdoor and indoor environments. Subgroups are based on quality of outdoor environment.

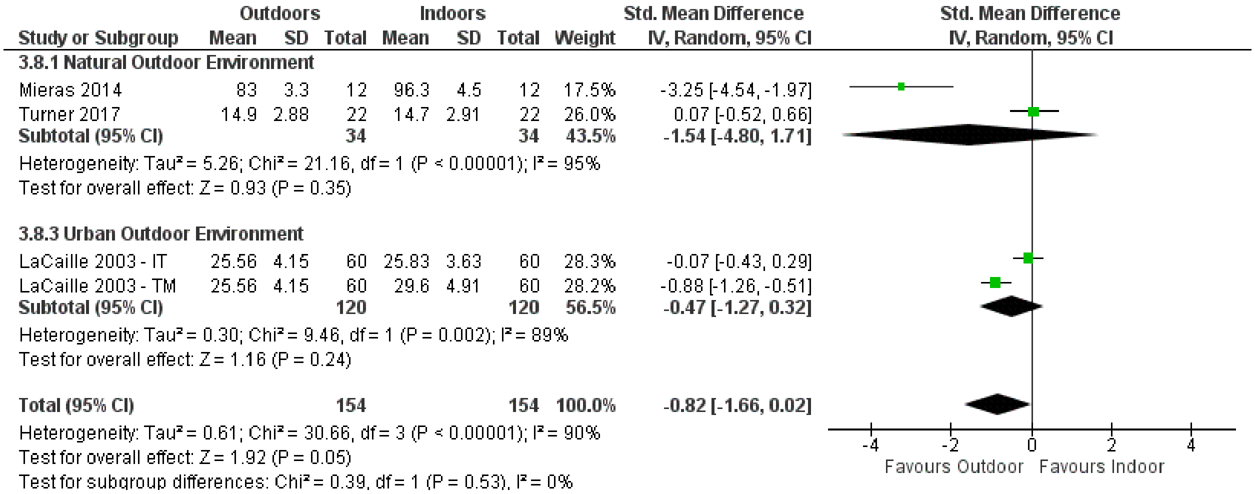


Figure 3.8 Forest plot of standardized mean difference in time to completion between outdoor and indoor environments. Subgroups are based on quality of outdoor environment.

Significant heterogeneity was detected in this analysis ($I^2 = 90\%$, $\chi^2 = 30.66$, df = 3, $p < 0.00001$). The only subgroup analysis that could be completed was by exercise environment type. In the pooled analysis of studies using natural outdoor environments (see Figure 3.8.1), the exercise environment did not significantly affect time to completion (Hedges *g* = -1.54, 95% CI = [-4.80, 1.71], $p = 0.35$) and significant heterogeneity was detected ($I^2 = 95\%$, $\chi^2 = 21.16$, df = 1, $p < 0.00001$). In the pooled analysis of studies using urban outdoor environments (see Figure 3.8.2), the exercise environment did not significantly affect time to completion (Hedges *g* = -0.47, 95% CI = [-1.27, 0.32], $p = 0.24$) and significant heterogeneity was detected ($I^2 = 89\%$, $\chi^2 = 9.46$, df = 1, $p = 0.002$). The subgroup analysis detected no significant difference in effects between subgroups ($I^2 = 0\%$, $\chi^2 = 0.39$, df = 1, $p = 0.53$).

The remaining five comparisons for all other performance measurements are depicted in the harvest plot in Figure 3.5 above. Across measures of power output, time to exhaustion, and distance, the majority of comparisons (*n* = 3, 60%) showed that exercising in an outdoor environment significantly improved performance[49,52,76] (i.e., greater power output, greater time to exhaustion, greater distance walked). Two comparisons (40%) identified no significant difference between exercise environments, although the calculated effect sizes of both comparisons suggest that a positive effect may be present[49,80].

3.4.4 Neuroendocrine Response

Participants' neuroendocrine response was measured by three studies[62,67,73], for a total of nine comparisons. Three comparisons (33%) over two studies measured the cortisol response to exercise[62,67], four comparisons (44%) in a single study measured catecholamine response[67], and a single comparison was made for each plasma beta-endorphins (11%) and blood lactate (11%) within one study[73]. In total, 3 of 9 of comparisons (33%) resulted in findings that the exercise environment significantly altered neuroendocrine output, and the remainder (n = 6, 67%) resulted in no significant effect (see Figure 3.9).

In only one of the three comparisons in which cortisol was assessed, reduced levels of cortisol were detected in the outdoor condition relative to an indoor condition with internal auditory stimuli. Lower noradrenaline was detected following outdoor exercise compared to indoor exercise with internal auditory stimuli, but the other three assessments of adrenaline and noradrenaline did not yield significant effects. Environment appeared to have no effect on plasma beta-endorphins. In the one comparison of blood lactate, outdoor exercise elicited a greater blood lactate than indoor exercise.

3.4.5 Cardiovascular Response

The effects of exercise on the cardiovascular system were measured by four articles[61,62,66,67]. Twenty-five unique comparisons were made for changes in systolic blood pressure[62,66,67] (n = 7, 28%), diastolic blood pressure[62,66] (n = 5, 20%), heart rate recovery[61,66] (n = 5, 20%), and heart rate variability[62] (n = 8, 32%) These comparisons are represented in the harvest plot below (see Figure 3.9). One study which assessed systolic and diastolic blood pressure

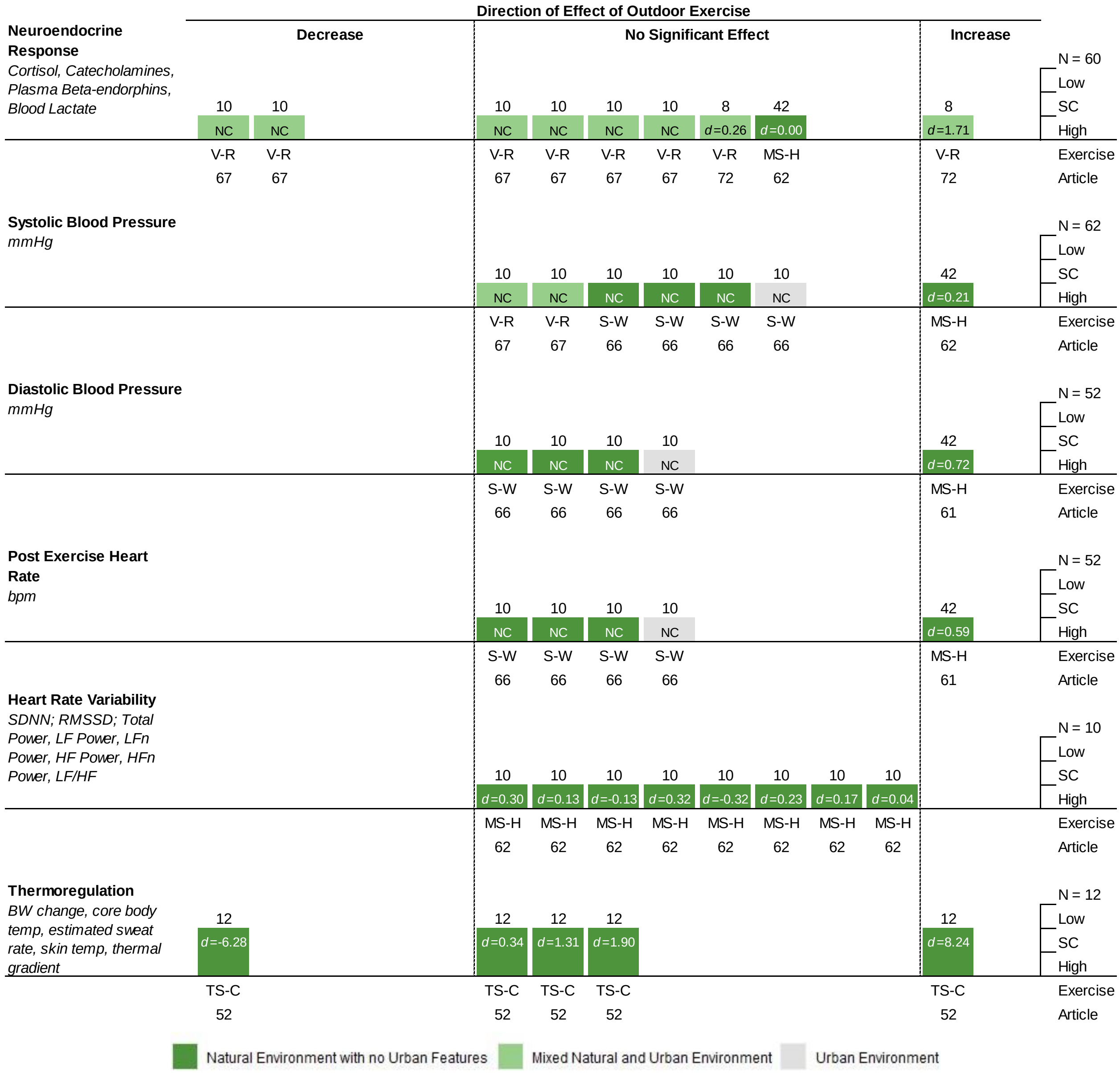


Figure 3.9 Harvest plot of direction of effect in neuroendocrine responses and cardiovascular responses (systolic blood pressure, diastolic blood pressure, heart rate recovery, heart rate variability) and thermoregulation between outdoor and indoor environments. Note: The number above each column indicates the total sample size for the study. The height of each column represents overall risk of bias of the study, with taller columns representing less risk of bias. Within the column, the Cohen's *d* represents effect size. Under each bar, exercise intensity and type is indicated below each column and separated by a hyphen with the article reference number below. Acronyms: NC: Not calculated due to insufficient reporting; V: Vigorous intensity; R: Running; MI: Moderate intensity; H: Hiking; S: Self-selected intensity; W: Walking; SDNN: Standard Deviation of normal-normal beat intervals of heart rate; RMSSD: Root Mean Square of Successive Differences between successive normal-to-normal beat intervals of heart rate; LF: Power in low frequency range; LFn: Power in low frequency range normalized; HF: Power in high frequency range; HFn: Power in high frequency range normalized; LF/HF: Ratio between power in low frequency and power in high frequency; BW: Body weight; TS: Self-selected training intensity; C: Cycling.

and heart rate recovery used four outdoor experimental conditions compared to one indoor control condition and thus contributes four comparisons to each outcome[66]. One study used two indoor control conditions that differed by auditory stimuli, which were both included as distinct comparisons in the assessment of systolic blood pressure[67]. Heart rate variability was measured using eight distinct operationalizations[62] (i.e., SDNN, RMSSD, total power, power in low frequency range, normalized low frequency power, power in high frequency range, normalized high frequency power, low frequency/high frequency ratio).

The effects of the exercise environment on systolic and diastolic blood pressure were significant in one of seven comparisons and one of five comparisons, respectively. These findings demonstrated greater blood pressure following outdoor exercise compared to indoor exercise. In one of five assessments of heart rate recovery, significantly greater post-exercise resting heart rate was detected. No other effects were significant across cardiovascular responses, but potentially small effects exist for heart rate variability.

3.4.6 Thermoregulation

A single study measured the effect of the environment on five measures associated with thermoregulation, including body weight change, estimated sweat rate, core body temperature, skin temperature, and thermal gradient[52]. Significant differences were only detected in two of the five outcomes (40%), skin temperature and thermal gradient (see Figure 3.9 above). Skin temperature was lower in the outdoor environment and thermal gradient – the difference in temperature between the core body and the skin – was greater in the outdoor condition. However, the significantly greater wind speed detected in the outdoor condition may explain the differences in thermal gradient and power output between the outdoor and indoor conditons[52].

3.4.7 Enjoyment

Enjoyment of the exercise session was measured in seven studies[64,67,69,70,72,75,82], in which nine distinct pairwise comparisons could be completed. Instruments used for comparison include the PACES Activity Enjoyment Scale[64,82] (n = 3, 33%) as well as single-item ratings of enjoyment that varied slightly by scale type and magnitude (e.g., 1-10 Likert-type scale[67,69,70,72], 1-100% analog scale[75]) across all comparisons (n = 6, 67%). There was sufficient homogeneity to conduct a quantitative analysis of enjoyment (see Figure 3.10). One of the studies did not report grouped summary data, only including findings by gender and were considered as distinct pairwise comparisons for the analysis[64]. One study used two indoor controls that differed by the type of auditory stimuli and these were similarly categorized as distinct comparisons[67]. In the pooled analysis, an outdoor exercise environment was found to significantly increase enjoyment (Hedges g = 1.24, 95% CI = [0.59, 1.89], p = 0.0002). Significant heterogeneity was detected in this analysis (I^2 = 85%, χ^2 = 53.86, df = 8, p < 0.00001). The only subgroup analysis that could be completed was by environment type. In the pooled analysis of studies using natural outdoor environments

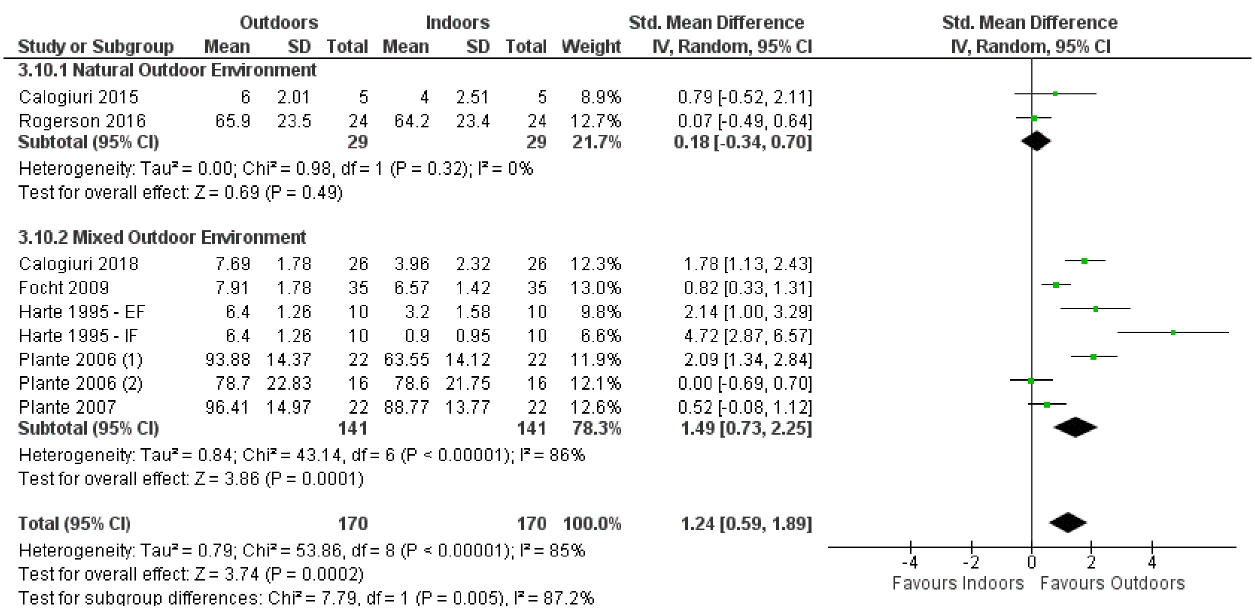


Figure 3.10 Forest plot of standardized mean difference in enjoyment between outdoor and indoor environments. Subgroups are based on quality of outdoor environment.

(see Figure 3.10.1), the exercise environment did not significantly affect enjoyment of exercise (Hedges $g = 0.18$, 95% CI = [-0.34, 0.70], $p = 0.49$) and no heterogeneity was detected ($I^2 = 0\%$). In the pooled analysis of studies using mixed outdoor environments (see Figure 3.10.2), the exercise environment did significantly improve enjoyment (Hedges $g = 1.49$, 95% CI = [0.73, 2.25], $p = 0.0001$) and significant heterogeneity was detected ($I^2 = 86\%$, $\chi^2 = 43.14$, df = 6, $p < 0.00001$). The subgroup analysis detected a significant difference in effects between subgroups ($I^2 = 87.2\%$, $\chi^2 = 7.79$, df = 1, $p = 0.005$).

3.4.8 Future Exercise Intentions

Two studies reported intention for future exercise using a visual analog scale ranging from 0-100% whereby participants were asked if they would be likely to "exercise regularly in a similar setting in the future" [72] or to "attend a free exercise session in the same place that [they] did [their] exercise" [75]. There was sufficient homogeneity to conduct a quantitative analysis of enjoyment (see Figure 3.11). In the pooled analysis, the exercise environment did not significantly affect future exercise intentions (Hedges $g = 0.40$, 95% CI = [-0.53, 1.33], $p = 0.40$). Significant heterogeneity was detected in this analysis ($I^2 = 83\%$, $\chi^2 = 6.03$, df = 1, $p = 0.01$).

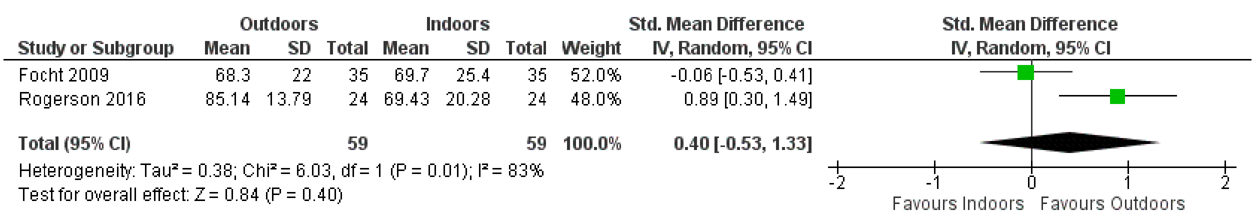


Figure 3.11 Forest plot of standardized mean difference in future intentions for exercise between outdoor and indoor environments.

3.4.9 Perception of the Environment

Two studies provided information on participants' perceptions of their environment[80,81] (see Figure 3.12). One study measured participants' perceived quality of the environment on a single item scale[81] (0-10). The study found that participants perceived the outdoor environment to be of significantly greater quality than the indoor environment. The second study used a 5-item subscale of the Belief about Green Exercise (BAGE) questionnaire to assess participants' attitudes

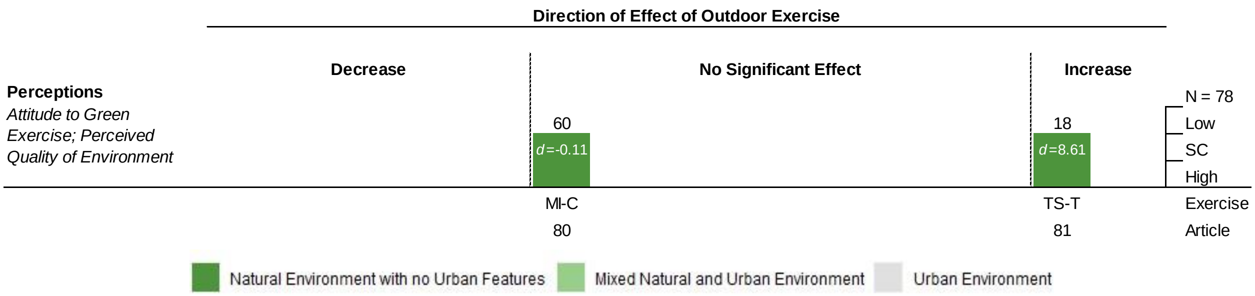


Figure 3.12 Harvest plot of direction of effect in perceptions of exercise between outdoor and indoor environments. The height of each column represents overall risk of bias of the study, with taller columns representing less risk of bias. Within the column, the Cohen's *d* represents effect size. Under each bar, exercise intensity and type is indicated below each column and separated by a hyphen with the article reference number below. Acronyms: MI: Moderate intensity; C: Cycling; TS: Self-selected training intensity; T: Taekwondo.

toward exercising in natural environments[80] (i.e., perceived expectations about behavioural outcomes of green exercise). Participants' BAGE scores were consistent across exercise environments as no significant differences were detected.

3.5 Overall Directionality of Findings

In total, 95 total comparisons were made; an overview of the overall direction of effects of all comparisons is depicted in Figure 3.13A. Comparisons were also categorized based on features of the exercise protocol including the outdoor exercise environment and indoor exercise environment (see Figure 3.13B), exercise intensity (see Figure 3.13C), and exercise type (see Figure 3.13D). The majority of comparisons (n = 60, 63%) revealed no significant effects, but 35 significant comparisons (37%) were identified that either represented increases (n = 28, 29%) or decreases (n = 7, 7%) in outcomes when assessed outdoors compared to indoors.

The quality of outdoor environments used for pairwise comparisons were classified as either mostly natural with very few urban features (i.e., green, blue, brown, mountainous; n = 43, 45%), a mix of natural and urban features (i.e., green/urban, blue/urban, university campus, outdoor track; n = 47, 50%), or mostly urban (n = 5, 5%). The majority of natural outdoor environment comparisons (*n* = 31, 72%) resulted in no effect between conditions, 10 of 43 comparisons (23%) resulted in increased values in the outdoor condition, and only two comparisons (5%) were identified where values decreased outdoors. In trials where the outdoor environment quality was mixed, about half of the 46 comparisons (n = 25; 53%) showed no effect between environments, 18 of 47 (38%) saw effects of positive direction when assessed outdoors, and four comparisons (9%) resulted in reduced values in the outdoor environment. One of the five comparisons (20%) made in urban-categorized outdoor environments resulted in a negative direction of effect, with the remaining four detecting no effect (80%).

Similarly, comparisons were completed for studies that conducted exercise in a laboratory-based indoor environment (*n* = 58, 61%), a traditional indoor exercise environment (*n* = 28, 30%), or other indoor environments (e.g., clinical setting; *n* = 9, 9%). Of those comparisons in which

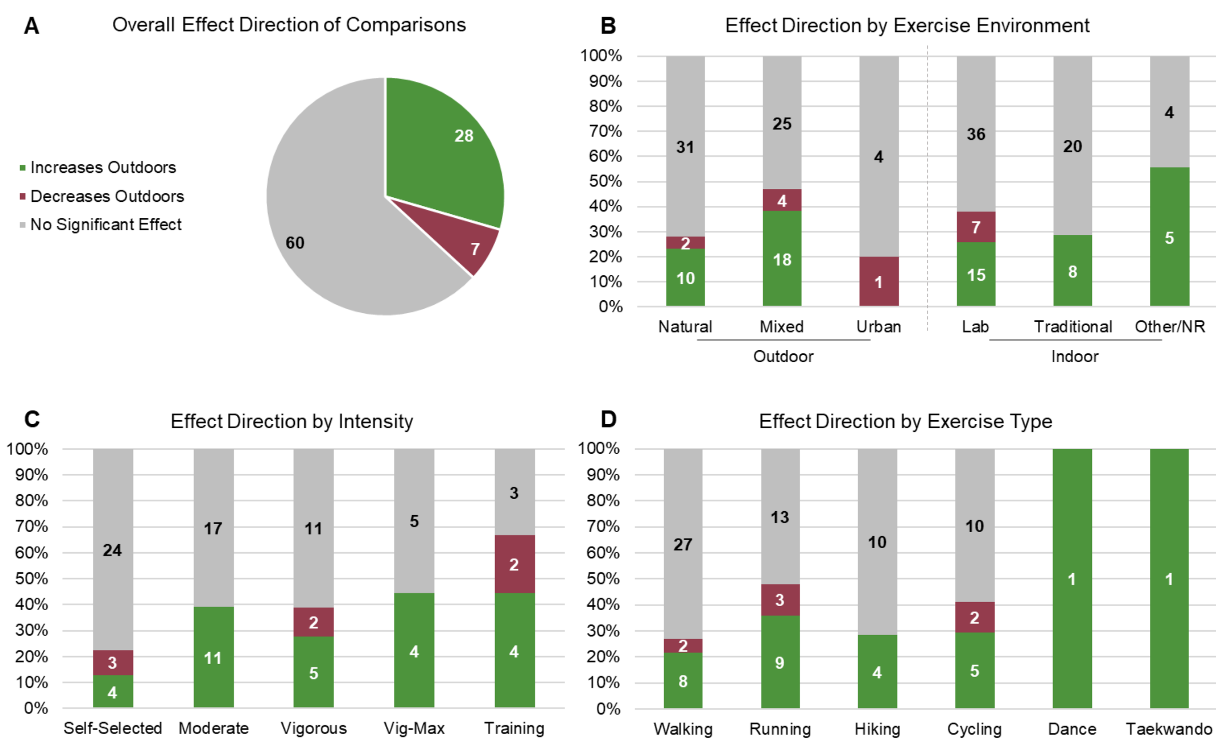


Figure 3.13 Summary of direction of effects detected between indoor and outdoor conditions for all comparisons (A) with sub-group categorizations based on the type of outdoor and indoor environments (B), exercise intensity (C), or exercise type (D).

exercise was conducted in a laboratory environment, the majority (*n* = 36, 62%) showed no effect, 15 of 58 comparisons (34%) showed significant increases in the outdoor condition and seven comparisons (4%) showed significant increases in the indoor condition. The majority of comparisons (*n* = 20, 71%) where participants exercised in traditional indoor exercise spaces demonstrated no significant difference between conditions and the remaining eight (29%) found significant differences that the outdoor environment elicited amplified measurements of outcomes.

When directionality of comparison effects was categorized based on exercise intensity, a possible trend was identified that is somewhat aligned with the dose-response of exercise intensity on benefits of exercise. Very few comparisons of self-selected intensity showed significant difference between conditions, where nearly 40-70% of comparisons at a moderate, vigorous, or training intensity demonstrated significant effects between the environments. When classified by exercise type, there did not appear to be any identifiable trends. It is possible that the increase in significant findings between running and walking were associated with the implicit difference in intensity between the two types of exercise. The only types of exercise in which the majority of comparisons saw significant changes were dancing and taekwondo, but each type of exercise was associated with just a single comparison.

Chapter 4: Discussion

4.1 Summary of Findings

The studies included in the present review examined whether the exercise environment effects differential physiological changes or subjective appraisals of the bout of exercise. Twenty-three distinct outcomes categorized into 9 groupings were assessed, including: (1) objective exercise intensity, (2) perceived exertion, (3) performance, (4) neuroendocrine response, (5) cardiovascular response, (6) thermoregulation, (7) enjoyment, (8) future exercise intentions, and (9) perception about the environment. Apparent from this systematic review of 24 studies, the authors used a wide variety of approaches to design studies. The result was a discordant assembly of methods, outcomes, and reporting, which made conclusions difficult to ascertain for the majority of the outcomes. In fact, only six of 23 outcomes were reported with the sufficient uniformity to conduct meta-analyses. These outcomes included a measure of objective exercise intensity (i.e., average heart rate), ratings of perceived exertion, two measures of performance (i.e., average speed and time to completion), enjoyment, and future intention for exercise. The findings for the remaining 17 outcomes were required to be synthesized using harvest plots and vote counting based on direction of effects that were statistically significant. The overall methodological diversity made it difficult to ascertain the existence of true effects.

4.1.1 Psychological effects of outdoor and indoor exercise

The present review included measures of perception of a single bout of exercise, including perceived exertion, enjoyment, future exercise intentions, and perceptions of the environment. From the meta-analyses and syntheses of results performed on the psychological outcomes, it can be concluded that exercise appears to be significantly more enjoyable and perceptually easier than indoor exercise but results for future exercise intentions and perceptions of the environment were inconclusive.

In the analysis of the five studies that examined perceived exertion under controlled physiological challenge, outdoor exercise resulted in lower perceived exertion, suggesting that the bout of exercise was psychologically easier to complete. Additionally, under similar physiological challenge (e.g., equivalent heart rate or oxygen consumption), participants across the seven studies found outdoor exercise to be more enjoyable than indoor exercise. The relationship between perceived exertion and enjoyment could not be analyzed in the present review, but previous work by Ekkekakis has identified that a relationship exists, such that lower perceived exertion is associated with greater enjoyment[36]. However, recent evidence suggests this relationship can be disrupted through the use of dissociative strategies[35], or by engaging in interval exercise, as opposed to continuous training[85,86]. Taken together with such evidence, it is implied that relationships exist between the exercise environment, perceived exertion, and enjoyment, when objective exercise intensity is controlled. Therefore, perceived exertion may mediate the effect of the environment on perceptions of enjoyment, but more research is needed to fully elucidate the nature of these relationships.

4.1.2 Physiological effects of outdoor and indoor exercise

The present review included physiological effects of exercise, including objective exercise intensity, neuroendocrine responses, cardiovascular responses, and thermoregulation. Studies examining effects on these outcomes during or after of a single bout of outdoor versus indoor exercise were impossible to summarize through a meta-analytical lens due to a high degree of methodological diversity between them. For the majority of these studies, results split between either no differences between the two environments or increased in the outdoor exercise environment. Three outcomes were able to be quantitatively analyzed: average heart rate during exercise, average speed during exercise, and time to completion. All three meta-analyses resulted in non-significant findings.

Outcomes that had to be synthesized using vote-counting included other measures of objective intensity (i.e., relative intensities, maximal heart rate, time spent in moderate to vigorous physical activity), performance (i.e., power output, time to exhaustion, distance), neuroendocrine responses (i.e., cortisol, catecholamines, beta-endorphins, blood lactate), cardiovascular responses (i.e., systolic and diastolic blood pressure, heart rate recovery, heart rate variability) and thermoregulation (i.e., body weight change, core body temperature, estimated sweat rate, skin temperature, thermal gradient). The findings for objective intensity were inconclusive, as the direction of identified effects were exhibited in both directions. Synthesis for performance were similarly inconsistent, but calculated effects suggest that there may be an effect toward improved performance in outdoor environments.

Thermoregulatory responses (e.g., greater thermal gradient) were identified in the outdoor environment. Higher thermal gradient is an indication that the outdoor environment may reduce the thermal strain on the body compared to indoor exercise where lesser convective cooling is present, allowing individuals to exercise more intensely and achieve improved performance outdoors[52]. This is aligned with findings from another study that included a comparison between environments for blood lactate concentration while exercise intensity was objectively controlled using percentage of maximal oxygen consumption[73]. In this study, there were no significant differences detected for either intensity or perceived exertion. Yet, elevated blood lactate values in the outdoor condition were detected, indicating that the environment may have resulted in greater physiological load caused by exercise. Finally, neuroendocrine and cardiovascular responses to exercise were primarily significantly null.

4.1.3 Implication of Findings

In most circumstances, exercise still presents the potential for improved physical and mental health and wellbeing on par with pharmacological treatment strategies[5,20], despite the limited additive physiological benefits identified by the present review as result of exercising outdoors. Yet, if acute exercise feels easier and is more enjoyable when performed outdoors, it may be more likely to be repeated. This effect was not reflected in the included analysis of future intentions for exercise. Moreover, future intentions have been linked to actual engagement, but to a relatively minor degree[87,88]. Therefore, we defer to other findings in the literature, such as research by Ekkekakis and others that has empirically shown that more enjoyable exercise predicts increased likelihood of actual future engagement in exercise[36,89]. So, while results were inconclusive for self-reported intentions for future exercise in the current review, promoting outdoor exercise might still lead to more sustained engagement in the public. Even though a single bout of outdoor exercise appears to be at the very least equally physiologically beneficial as indoor exercise, outdoor exercise may be more sustainable and lead to physiological benefits that accrue over time following repeated bouts.

4.2 Overall Quality of Evidence

4.2.1 Limitations of the included studies

Despite the paucity of evidence supporting potential synergistic physiological benefits of outdoor exercise, there may still be true effects that were undetectable due to the presence of confounding variabes and methodological concerns. A particular challenge about study designs that conduct exercise outdoors compared to indoors is the inability to control for features of the environment, including wind, noise, and other people. Such confounding variables make it difficult to investigate the cause of differences between exercise environments. For example, the presence of wind may result in a greater thermal gradient, reduced thermal strain, and greater power output at a similar exertion, as seen in the study conducted by Mieras and colleagues[52]. Future studies examining differences in exercise environment should identify, measure, and account for such extraneous variables in analysis.

Included studies may have been limited by their statistical power, which is influenced by factors such as study design and sample size. Studies were required to assess outcomes during or after a single bout of exercise performed in each an outdoor environment and an indoor environment, but there was little agreement in terms of study design. A repeated-measures, crossover design, like that used in most studies, provides greater statistical power to a small sample size as participants are used as their own control. However, studies which did not use counterbalancing, randomization, or both, are of lower quality, which limits the internal validity of this research and reduces the weight and interpretability of the evidence.

Overall, the participant samples that were utilized tended to be small, but according to power analyses conducted in a little more than half the studies, the sample sizes were appropriate for a statistical power of at least 80%. The majority of these sample sizes were completed retrospectively, which may introduce a degree of confirmation bias, but is preferred over the omission of a calculation altogether. Yet, two of these studies did not analyze samples large enough to satisfy the sample size calculation they provided[62,80]. Insufficient statistical power is one possible explanation for limited statistically significant differences. Methodological diversity elsewhere may have also contributed to the restricted interpretations including characteristics of the sample, like the gender, age, and health of participants, features of the environment, and features of the exercise protocol.

The sex or gender of analyzed participants may have influenced findings. For example, sex differences have been identified in outcomes such as blood pressure[90] and heart rate variability[91]. Only one study utilized a participant sample that was half male and half female[71], yet in total the aggregate sample was approximately evenly distributed between females (57%) and males (43%). Authors may have chosen to utilize a single sex to control for variation in some outcomes where sex differences have been previously identified, but information detailing these issues was seldom reported by the included articles. Additionally, across studies, there was poor differentiation between sex and gender. While the present review included percentages of males and females that were reported by authors, in most cases these authors conflated sex and gender and little information was provided on how this information was collected. Two studies reported environment X gender interaction effects for walking speed[71] and enjoyment[64], such that the difference between walking speed across environments was greater in men than in women and that females' enjoyment increased outdoors more than males', compared to indoors. With such little information about sex/gender differences, conclusions cannot be drawn about the effect of sex and gender on summarized outcomes, but the external validity of the findings should extend to both males and females.

With a weighted mean ± SD age of 29.36 ± 1.65 years and more than half the studies having samples younger than the weighted mean, a question arises about the generalizability of the presented findings. Similarly, the health status of the samples included a disproportionate number of participants who were identified as being healthy, active, or athletes. Previous research has suggested greater benefits of exercise in older populations and those with poorer health and health behaviours, but this effect was unable to be detected by the present review. Subgroup analysis nor effective synthesis of findings were conducted on outcomes based on age or baseline-characteristics of participants as a result of the limited number of studies which assessed each outcome and the few consistent characteristics that might have been examined. Therefore, it is impossible to conclude that age or health status had a moderating effect on the findings summarized above. However, due to the reported age and average health status of analyzed participants, these findings can likely only be generalized to younger populations who are mostly healthy and already active.

Finally, most of the articles were identified as having an overall high risk of bias, with the remainder identified as having some concerns overall. Primarily, these risk of bias rulings were a result of articles' insufficient reporting, poor or non-existent randomization procedures, and inadequate accounting for order and carryover effects. Beyond issues that were captured by the employed risk of bias tool, there are some other quality issues with the included studies. For example, while no studies were assessed as having evidence of missing outcome data, effect sizes and information to calculate effect sizes (i.e., mean ± SD and allocation ratios of participants) were often not reported. The quality of the reviewed evidence was deemed to be low, which made synthesis difficult, and the findings presented should be cautiously interpreted.

4.2.2 Limitations of the present review

The primary limitation of the present review is related to the scope; conclusions about the long-term effects of outcomes are not able to be made. Only acute studies were included to understand the acute mechanisms by which outdoor exercise may provide beneficial health effects compared to indoor exercise. Further, few psychological outcomes were included that may act in concert or parallel with the physiological effects of outdoor exercise. Theories like the Attention

Restoration Theory demonstrate the possibility of psychological changes as a result of exercising outdoors[15]. Outcomes like general affective valence or activation, feelings of energy, revitalization, and fatigue, and other moods and emotions could shed light on other benefits of outdoor exercise and why it might be perceived to be easier and more enjoyable. Therefore, it is likely that limiting the scope in the ways done here – to only physiological or psychological outcomes directly linked to the reported bout of exercise – limits the conclusions that can be garnered about how specifically the exercise environment may impact health over the short and long term.

In contrast, the present review used eligibility criteria which included controlled studies, regardless of randomization. The review of these non-randomized controlled trials permitted the synthesis of a greater amount of data with a greater number of outcomes. However, while this criterion was aligned with the two previous reviews[9,10], inclusion of these additional studies was at the cost of the overall quality of evidence. A considerable number of studies that were assessed to be high risk of bias was a result of the lack of randomization that was employed. As the expansion of the field continues to increase the overall number of studies examining outdoor exercise, future reviews should constrain the scope to randomized studies.

The method of vote counting using effect directions has some considerable limitations associated with it. As discussed in the Cochrane Handbook, vote counting can be an appropriate method to synthesize data in order to answer the question “Is there any evidence of an effect?”[54]. The determination to use vote counting was a result of minimal consistent reporting, the small number of studies, and the high number of outcomes assessed. The Cochrane Handbook warns against vote counting based on statistical significance as it may provide misleading conclusions[54]. Conclusions based solely on significance omit the effects of studies that are simply underpowered.

To attenuate this limitation, the present review employed a hybrid version of vote counting which summarized both significant findings as well as directions of effect. Further, the accompanying harvest plots included calculations of estimated effects and integrated defining features of studies to assist with interpretation.

4.2.3 Degree of agreement with previous systematic reviews

4.2.3.1 Included Studies

The present review set out to update the reviews conducted by Thompson Coon and colleagues (2011) and Lahart, Darcy, Gidlow, & Calogiuri, (2019). A decade since the first review on this topic has allowed for research in the area to proliferate, allowing for a much greater study pool from which it is possible to synthesize and interpret findings. Unlike both previous reviews, the scope of the present review is much narrower due to the following changes: (1) exclusion of psychological and affective measures not directly related to a bout of exercise (e.g., positive and negative affect, attention); (2) exclusion of studies which measure outcomes over a longitudinal study design; and (3) a required operational definition of the intervention condition for outdoor exercise to be undertaken outdoors, not including studies using virtual reality or sensory occlusions.

The search strategy and screening process used by Thompson Coon and colleagues (2011) collected 11 articles, while Lahart and colleagues (2019) included 31 articles that assessed the findings of 28 trials. In comparison to the latter, only 16 articles included in their review met the inclusion criteria of the present study[49,52,73,75,77–79,82,61–64,67,69,70,72], resulting in an additional eight studies which contained novel findings not summarized to date in either the 2011 or 2019 reviews[65,66,68,71,74,76,80,81]. Of the 15 studies which were included by Lahart and colleagues (2019) but not in the present review, five studies utilized an inappropriate study design (e.g., longitudinal intervention), five studies only measured outcomes not included in the present review (e.g., positive and negative affect, concentration), four studies only included a 'green' condition that used virtual reality in a laboratory setting, and one study was not peer reviewed. The primary reason for the discrepancy in search results was from a difference in how the intervention and control conditions were operationally defined.

The 8 novel studies not yet included in a review until the current one include three studies that were not published at the time of Lahart and colleagues' search conducted on June 28, 2018[65,66,80]. As their search was intended as an update to the review conducted by Thompson Coon and colleagues in 2011[9], the search by Lahart and colleagues did not extend before January 1, 2010, which resulted in the omission of one study published in 2004[68]. In addition, four studies appeared to meet their search criteria but were omitted without any stated reason[71,74,76,81]. It is possible that Lahart and colleagues' (2019) erroneously excluded these studies because they appeared to utilize more athletic or active populations, with greater emphasis on performance-type outcomes. While this is simply speculation, these studies met all the inclusion criteria of the present review. Further, the inclusion of these studies is warranted due to the implications that performance-type measures of exercise may have on the subsequent health effect of performed exercise. For example, power output during cycling, average speed, and time to completion of time-trial type exercise protocols can all be used to infer individuals' cardiorespiratory fitness[83].

4.2.3.2 Discrepancy of findings

The present review is unable to comment on some of the outcomes of the previous reviews, as Lahart and colleagues included psychological outcomes like affect, moods, emotions, attention, and memory in addition to those summarized here[10]. Our review did not include more diffuse psychological outcomes that were not directly associated with engagement in or perceptions of the bout of exercise (e.g., enjoyment of the bout). The summarized findings of the present review largely agree with those previously by Lahart and colleagues[10]. The authors reported that green exercise was associated with heightened enjoyment and lower perceived exertion of physical activity. However, all other findings were inconclusive.

The most significant discrepancy in the summarized findings between the present review and the review conducted by Lahart and colleagues was related to perceptions of exertion of outdoor exercise compared to indoor exercise[10]. The previous review synthesized seven studies that measured perceived exertion[52,61,67,72,73,75,78], of which six reported null findings[52,61,72,73,75,78]. Yet, the authors conflated subjective intensity as a control parameter of the exercise protocol and perceived exertion as an outcome. In other words, when a prescribed level of subjective intensity was provided (i.e., the effort participants should be using during their bout of exercise), there should be limited expected differences in perceived intensity between environmental conditions. In four of the seven studies examined by Lahart and colleagues, subjective intensity was used as a control parameter[52,61,72,78]. In each of these four studies, null differences in exertion between outdoor and indoor exercise were reported, as opposed to being interpreted as properly controlled subjective intensity across experimental conditions. In the present review, the syntheses of subjective intensity and perceived exertion are separated to prevent this conflation. The seven studies that were included in Lahart and colleagues review[10] were all included in the present review, but separated based on whether perceived exertion was included as a parameter used to control for exercise intensity or as an outcome. The analysis of perceived exertion contained herein as an outcome included the three studies from the 2019 review that used objective intensity control parameters to set their protocols[67,73,75], in addition to two novel studies not yet synthesized[69,80]. To further strengthen the reliability of this finding, the present review was able to quantitatively analyze the outcome via meta-analysis, as opposed to relying on vote-counting based on effect direction. The effect estimate may not be accurate, as there was substantial statistical heterogeneity in the reported findings, but this synthesis does support the conclusion that the exercise environment has a significant effect on perceived exertion. Future research should continue this investigation into how perceived exertion can be affected by extrinsic factors, but authors should contextualize findings around the parameters used to control exercise intensity.

Similarly, Lahart and colleagues (2019) also used vote-counting of effect direction to support their claim that enjoyment of outdoor exercise was significantly greater than both indoor exercise or indoor exercise with simulated nature exposure. Five studies examined by Lahart and colleagues supported their finding[64,67,70,72,82] with one study showing no effect[75]. Each of the studies previously examined were included in the present synthesis, with the addition of one study not previous identified[69]. The outcome was then meta-analyzed in the current review to be able to quantify the previously identified effect of the environment on enjoyment of exercise. Though while it was determined that the methodology and reporting of studies were sufficiently uniform to conduct a quantitative analysis, the meta-analysis detected significant, considerable statistical heterogeneity. The identified heterogeneity suggests that the pooled effect size may not be reliable due to inconsistent effect sizes of individual studies and a small number of comparisons, but an effect which favours greater enjoyment in outdoor exercise is still likely.

4.3 Future Directions

The term “green exercise” was first used by Pretty and colleagues in 2003 to describe exercise performed outdoors while exposed to nature[18]. Even before that, authors hypothesized models that explained the beneficial effects of nature exposure[15,17]. Previous authors and research groups have used this definition to guide their research questions and the scope of studies and reviews[8–11,92]. It is possible that this definition of green exercise as exercise in exposure to nature is too narrow as the present study shows that there may not be a significant difference in the health effects obtained when exercising in natural, mixed, and urban environments and exercise in general. Future studies and reviews should address this potential limitation when examining 'green exercise'. To do this, one approach would be to quantify defining features of outdoor spaces (e.g., green-ness, biodiversity, environmental quality) that are required to effect modulation in perceived exertion, enjoyment, and foster the potential for health improvement compared to indoor exercise and begin a conversation as to the mechanisms involved.

Within the scope of assessing physiological outcomes, there were many potential health-related outcomes and biomarkers either not examined or examined with minimal consistency. For example, it is well understood that exercise has a hypotensive effect on blood pressure, but blood pressure was only measured in two studies. Other potential effects that have been identified in studies of shinrin-yoku (i.e., forest bathing), but not measured by studies included in this review, include changes to immune function [44,50], metabolic biomarkers [48], and other hormones including catecholamines, dopamine, and DHEA [48]. However, this is not an exhaustive list as there are many outcomes that have never been examined in relation to outdoor exercise. For example, Brain Derived Neurotrophic Factor (BDNF) has been shown to increase with acute bouts of exercise and that these increases become more sensitive as individuals become more physically active [93]. BDNF plays a role in brain health and promotes memory and cognitive function in the general population and individuals with pathological conditions[94]. Yet, BDNF has not been investigated in the context of outdoor exercise or environmental quality. The omission of such outcomes and the significant methodological diversity when measuring outcomes should be addressed in future studies to allow for more conclusive results using meta-analytic strategies.

The evidence that does exist on outdoor exercise is of questionable quality. Future studies should take an open science approach, including the pre-registration of study methods and statistical analysis plan, and publication of summary data and de-identified participant data. As inter-individual differences (e.g., sex and gender, age, cardiorespiratory fitness, affinity for outdoor environments) may exist within a sample, studies should ensure proper randomization is conducted regardless of design, and counterbalancing and adequate washout in crossover designs should also be included whenever possible. The physiological outcomes that should be measured might require larger sample sizes to reduce variance and detect smaller effect sizes.

This research is potentially relevant to public health and public policy. If exercising in outdoor spaces may make exercise feel easier and more enjoyable and subsequently result in health improvements due to greater physical activity engagement, then the accessibility of those spaces should be prioritized to support exercise, active travel, and other forms of physical activity. As discussed above, larger and more rigorous studies are required to understand the true effect of outdoor exercise and a coordinated effort between university-based researchers and public health agencies may support greater quality evidence to effect change on public policy.

4.4 Conclusion

The present systematic review and meta-analysis provided a mixture of narrative and quantitative syntheses of the evidence which exists, to date, comparing the differential effects of a single bout of outdoor exercise compared to an indoor exercise bout on physiological measures, perceived exertion, and enjoyment. The current literature that exists in this field is of primarily low quality and rigour, and future studies should address these methodological deficiencies. Such variability in study design made it difficult to assess whether factors, such as quality of the environment or specific details of the exercise protocol (i.e., intensity, time, or type of activity), moderate the effect of outdoor exercise. It was found that when compared to indoor exercise, outdoor exercise provided greater enjoyment and lower perceptions of exertion during exercise of a controlled intensity, but that inconclusive evidence exists supporting additive benefits to physiological outcomes of outdoor exercise. In sum, exercise has been previously shown to provide substantial benefits to short and long-term health, and exercising outdoors is one strategy to make exercise feel easier and more enjoyable, thus potentially fostering sustained physical activity engagement in the long-term.

References

1. *Health, Place and Nature*. *UK Sustainable Development Commission* http://www.sd-commission.org.uk/data/files/publications/Outdoor_environments_and_health.pdf (2008).
2. Center for Disease control and Prevention. *Strategies to Prevent Obesity and Other Chronic Diseases: The CDC Guide to Strategies to Increase Physical Activity in the Community.* http://www.cdc.gov/obesity%0Ahttp://www.cdc.gov/obesity/downloads/PA_2011_WEB.pdf (2011).
3. De Vries, S., Verheij, R. A., Groenewegen, P. P. & Spreeuwenberg, P. Natural environments - Healthy environments? An exploratory analysis of the relationship between greenspace and health. *Environ. Plan. A* **35**, 1717–1731 (2003).
4. Pretty, J. *et al*. Green exercise in the UK countryside: Effects on health and psychological well-being, and implications for policy and planning. *J. Environ. Plan. Manag.* **50**, 211–231 (2007).
5. Pedersen, B. K. & Saltin, B. Exercise as medicine - Evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scand. J. Med. Sci. Sport.* **25**, 1–72 (2015).
6. Twohig-Bennett, C. & Jones, A. The health benefits of the great outdoors: A systematic review and meta-analysis of greenspace exposure and health outcomes. *Environ. Res.* **166**, 628–637 (2018).
7. Pretty, J., Peacock, J., Sellens, M. & Griffin, M. The mental and physical health outcomes of green exercise. *Int. J. Environ. Health Res.* **15**, 319–337 (2005).
8. Barton, J. & Pretty, J. What is the best dose of nature and green exercise for improving mental health- A multi-study analysis. *Environ. Sci. Technol.* **44**, 3947–3955 (2010).

9. Thompson Coon, J. *et al.* Does participating in physical activity in outdoor natural environments have a greater effect on physical and mental wellbeing than physical activity indoors? A systematic review. *Environ. Sci. Technol.* **45**, 1761–1772 (2011).

10. Lahart, I., Darcy, P., Gidlow, C. & Calogiuri, G. The effects of green exercise on physical and mental wellbeing: A systematic review. *Int. J. Environ. Res. Public Health* **16**, (2019).

11. Rogerson, M. *et al.* Regular doses of nature: The efficacy of green exercise interventions for mental wellbeing. *Int. J. Environ. Res. Public Health* **17**, 1–15 (2020).

12. Akers, A. *et al.* Visual color perception in green exercise: Positive effects on mood and perceived exertion. *Environ. Sci. Technol.* **46**, 8661–8666 (2012).

13. Marselle, M. R., Irvine, K. N., Lorenzo-Arribas, A. & Warber, S. L. Moving beyond green: Exploring the relationship of environment type and indicators of perceived environmental quality on emotional well-being following group walks. *Int. J. Environ. Res. Public Health* **12**, 106–130 (2015).

14. Marselle, M. R., Irvine, K. N., Lorenzo-Arribas, A. & Warber, S. L. Does perceived restorativeness mediate the effects of perceived biodiversity and perceived naturalness on emotional well-being following group walks in nature? *J. Environ. Psychol.* **46**, 217–232 (2016).

15. Kaplan, R. & Kaplan, S. *The experience of nature. A psychological perspective*. (1989). doi:10.4324/9781351054386-1.

16. Herzog, T. R., Maguire, C. P. & Nebel, M. B. Assessing the restorative components of environments. *J. Environ. Psychol.* **23**, 159–170 (2003).

17. Ulrich, R. S. *et al.* Stress recovery during exposure to natural and urban environments. *J. Environ. Psychol.* **11**, 201–230 (1991).

18. Pretty, J., Griffin, M., Sellens, M. & Pretty, C. *Green Exercise : Complementary Roles of Nature , Exercise and Diet in Physical and Emotional Well-Being and Implications for Public Health Policy*. (2003).

19. Jahn, N. europepmc: R Interface to the Europe PubMed Central RESTful Web Service. (2020).

20. Rhodes, R. E., Janssen, I., Bredin, S. S. D., Warburton, D. E. R. & Bauman, A. Physical activity: Health impact, prevalence, correlates and interventions. *Psychol. Heal.* **32**, 942–975 (2017).

21. Ibrahim, E. M. & Al-Homaidh, A. Physical activity and survival after breast cancer diagnosis: meta-analysis of published studies. *Med. Oncol.* **28**, 753–765 (2011).

22. Meyerhardt, J. A. *et al.* Physical activity and male colorectal cancer survival. *Arch. Intern. Med.* **169**, 2102–2108 (2009).

23. Fransen, M. *et al.* Exercise for osteoarthritis of the knee. *Cochrane Database of Systematic Reviews* vol. 2015 (2015).

24. Thompson, P. D. *et al.* The acute versus the chronic response to exercise. *Med. Sci. Sports Exerc.* **33**, S438–S445 (2001).

25. Dimkpa, U. Post-exercise heart rate recovery: an index of cardiovascular fitness. *J. Exerc. Physiol.* **12**, 10–22 (2009).

26. Hötting, K., Schickert, N., Kaiser, J., Röder, B. & Schmidt-Kassow, M. The Effects of Acute Physical Exercise on Memory, Peripheral BDNF, and Cortisol in Young Adults. **2016**, (2016).

27. Tsai, C.-L. *et al.* Executive function and endocrinological responses to acute resistance exercise. *Front. Behav. Neurosci.* **8**, 262 (2014).

28. Anderson, T. & Wideman, L. Exercise and the cortisol awakening response: A systematic review. *Sport. Med. - Open* **3**, (2017).

29. Statistics Canada. *Canadian health measures survey: Directly measured physical activity of Canadians, 2012 to 2013. Health Fact Sheet* (2015).

30. Annesi, J. J. Effects of a cognitive behavioral treatment package on exercise attendance and drop out in fitness centers. *Eur. J. Sport Sci.* **3**, 1–16 (2003).

31. Ajzen, I. The theory of planned behavior. *Organ. Behav. Hum. Decis. Process.* **50**, 179–211 (1991).

32. Downs, D. S. & Hausenblas, H. A. The theories of reasoned action and planned behavior applied to exercise: A Meta-analytic Update. *J. Phys. Act. Heal.* **2**, 76–97 (2005).

33. Fen, Y. S. & Sabaruddin, N. A. An extended model of theory of planned behaviour in predicting exercise intention. *Int. Bus. Res.* **1**, 108–122 (2008).

34. Jekauc, D. Enjoyment during exercise mediates the effects of an intervention on exercise adherence. *Psychology* **06**, 48–54 (2015).

35. Jones, L., Karageorghis, C. I. & Ekkekakis, P. Can high-intensity exercise be more pleasant? Attentional dissociation using music and video. *J. Sport Exerc. Psychol.* **36**, 528–541 (2014).

36. Ekkekakis, P., Parfitt, G. & Petruzzello, S. J. The pleasure and displeasure people feel when they exercise at different intensities: Decennial update and progress towards a tripartite rationale for exercise intensity prescription. *Sports Medicine* vol. 41 641–671 (2011).

37. Barreto-Silva, V., Bigliassi, M., Chierotti, P. & Altimari, L. R. Psychophysiological effects of audiovisual stimuli during cycle exercise. *Eur. J. Sport Sci.* **18**, 560–568 (2018).

38. Ulrich, R. S. *Biophilia, biophobia, and natural landscapes*. (1993).

39. Ulrich, R. S. View through a window may influence recovery from surgery. *Science (80-.).* **224**, 420–421 (1984).

40. Gladwell, V. F. *et al.* The effects of views of nature on autonomic control. *Eur. J. Appl. Physiol.* **112**, 3379–3386 (2012).

41. Antonelli, M., Barbieri, G. & Donelli, D. Effects of forest bathing (shinrin-yoku) on levels of cortisol as a stress biomarker: a systematic review and meta-analysis. *Int. J. Biometeorol.* **63**, 1117–1134 (2019).

42. Hansmann, R., Hug, S. M. & Seeland, K. Restoration and stress relief through physical activities in forests and parks. *Urban For. Urban Green.* **6**, 213–225 (2007).

43. Joseph, R. P. & Maddock, J. E. Observational Park-based physical activity studies: A systematic review of the literature. *Prev. Med. (Baltim).* **89**, 257–277 (2016).

44. Park, B. J., Tsunetsugu, Y., Kasetani, T., Kagawa, T. & Miyazaki, Y. The physiological effects of Shinrin-yoku (taking in the forest atmosphere or forest bathing): Evidence from field experiments in 24 forests across Japan. *Environ. Health Prev. Med.* **15**, 18–26 (2010).

45. Park, B. J. *et al.* Physiological effects of Shinrin-yoku (taking in the atmosphere of the forest) - Using salivary cortisol and cerebral activity as indicators-. *J. Physiol. Anthropol.* **26**, 123–128 (2007).

46. Tsunetsugu, Y. *et al.* Physiological effects of Shinrin-yoku (taking in the atmosphere of the forest) in an old-growth broadleaf forest in Yamagata Prefecture, Japan. *J. Physiol. Anthropol.* **26**, 135–142 (2007).

47. Halliwill, J. R. Mechanisms and Clinical Implications of Post-exercise Hypotension in

Humans. *Exerc. Sport Sci. Rev.* **29**, 65–70 (2001).

48. Li, Q. *et al.* Acute effects of walking in forest environments on cardiovascular and metabolic parameters. *Eur. J. Appl. Physiol.* **111**, 2845–2853 (2011).

49. Carvalho, C., Sunnerhagen, K. S. & Willén, C. Walking speed and distance in different environments of subjects in the later stage post-stroke. *Physiother. Theory Pract.* **26**, 519–527 (2010).

50. Li, Q. Effect of forest bathing trips on human immune function. *Environ. Health Prev. Med.* **15**, 9–17 (2010).

51. Wooller, J. J., Barton, J., Gladwell, V. F. & Micklewright, D. Occlusion of sight, sound and smell during Green Exercise influences mood, perceived exertion and heart rate. *Int. J. Environ. Health Res.* **26**, 267–280 (2016).

52. Mieras, Heesch & Slivka. Physiological and psychological responses to outdoor vs. laboratory cycling. *J. strength Cond. Res.* **28**, 2324–2329 (2014).

53. Liberati, A. *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med.* **6**, e1000100 (2009).

54. *Cochrane Handbook for Systematic Reviews of Interventions*. (Cochrane, 2021).

55. Sterne, J. A. C. *et al.* RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ* **366**, 1–8 (2019).

56. McGuinness, L. A. robvis: An R package and web application for visualising risk-of-bias assessments. (2019).

57. Shi, J. *et al.* Optimally estimating the sample standard deviation from the five-number summary. *Res. Synth. Methods* **11**, 641–654 (2020).

58. Luo, D., Wan, X., Liu, J. & Tong, T. Optimally estimating the sample mean from the sample size, median, mid-range, and/or mid-quartile range. *Stat. Methods Med. Res.* **27**, 1785–1805 (2018).

59. Shi, J. *et al.* Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. http://www.math.hkbu.edu.hk/~tongt/papers/median2mean.html (2020).

60. Review Manager (RevMan). (2014).

61. Niedermeier, M., Einwanger, J., Hartl, A. & Kopp, M. Affective responses in mountain hiking - A randomized crossover trial focusing on differences between indoor and outdoor activity. *PLoS One* **12**, 1–18 (2017).

62. Niedermeier, M., Grafetstätter, C., Hartl, A. & Kopp, M. A randomized crossover trial on acute stress-related physiological responses to mountain hiking. *Int. J. Environ. Res. Public Health* **14**, (2017).

63. Chu, C. Y., Lu, S. Y. & Lin, K. F. Influences of exercise experience and exercise settings on heart rate responses during self-selected intensity exercises. *J. Exerc. Sci. Fit.* **8**, 73–77 (2010).

64. Plante, T. G., Cage, C., Clements, S. & Stover, A. Psychological benefits of exercise paired with virtual reality: Outdoor exercise energizes whereas indoor virtual exercise relaxes. *Int. J. Stress Manag.* **13**, 108–117 (2006).

65. Farias, T. D. B. *et al.* Psycho-affective responses of physically active older adults during exercise in different environments. *J. Phys. Educ. Sport* **18**, 1640–1645 (2018).

66. Navalta, J. W., Bodell, N. G., Tanner, E. A., Aguilar, C. D. & Radzak, K. N. Effect of exercise in a desert environment on physiological and subjective measures. *Int. J. Environ.*

Health Res. **00**, 1–11 (2019).

67. Harte, J. L. & Eifert, G. H. The effects of running, environment, and attentional focus on athletes' catecholamine and cortisol levels and mood. *Psychophysiology* **32**, 49–54 (1995).

68. LaCaille, R. A., Masters, K. S. & Heath, E. M. Effects of cognitive strategy and exercise setting on running performance, perceived exertion, affect, and satisfaction. *Psychol. Sport Exerc.* **5**, 461–476 (2004).

69. Calogiuri, G., Nordtug, H. & Weydahl, A. The potential of using exercise in nature as an intervention to enhance exercise behavior: Results from a pilot study. *Percept. Mot. Skills* **121**, 350–370 (2015).

70. Calogiuri, G. *et al.* Experiencing nature through immersive virtual environments: Environmental perceptions, physical engagement, and affective responses during a simulated nature walk. *Front. Psychol.* **8**, 1–14 (2018).

71. Dasilva, S. G. *et al.* Psychophysiological responses to self-paced treadmill and overground exercise. *Med. Sci. Sports Exerc.* **43**, 1114–1124 (2011).

72. Focht, B. C. Brief walks in outdoor and laboratory environments: Effects on affective responses, enjoyment, and intentions to walk for exercise. *Res. Q. Exerc. Sport* **80**, 611–620 (2009).

73. McMurray, R. G., Berry, M. J., Vann, R. T., Hardy, C. J. & Sheps, D. S. The effect of running in an outdoor environment on plasma beta endorphins. *Ann. Sport. Med.* **3**, 230–233 (1988).

74. Peserico, C. S. & Machado, F. A. Comparação entre desempenhos de corrida time trial realizados em pista e esteira. *Rev. Bras. Cineantropometria e Desempenho Hum.* **16**, 456–464 (2014).

75. Rogerson, M., Gladwell, V. F., Gallagher, D. J. & Barton, J. L. Influences of green outdoors versus indoors environmental settings on psychological and social outcomes of controlled exercise. *Int. J. Environ. Res. Public Health* **13**, (2016).

76. Slapsinskaite, A. *et al.* Cycling outdoors facilitates external thoughts and endurance. *Psychol. Sport Exerc.* **27**, 78–84 (2016).

77. Teas, J., Hurley, T., Msph, S. G. & Mph, K. O. Walking outside Improves Mood for Healthy Postmenopausal Women. *Clin. Med. Oncol.* **1**, CMO.S343 (2007).

78. Turner, T. L. & Stevinson, C. Affective outcomes during and after high-intensity exercise in outdoor green and indoor gym settings. *Int. J. Environ. Health Res.* **27**, 106–116 (2017).

79. Byrka, K. & Ryczko, N. Positive effects of dancing in natural versus indoor settings: The mediating role of engagement in physical activity. *J. Environ. Psychol.* **57**, 25–33 (2018).

80. Flowers, E. P., Freeman, P. & Gladwell, V. F. Enhancing the acute psychological benefits of green exercise: An investigation of expectancy effects. *Psychol. Sport Exerc.* **39**, 213–221 (2018).

81. Jang, S. & So, W. Y. The effect of short-term outdoor taekwondo training on the concentration and mood of taekwondo players. *J. Mens. health* **13**, e68–e75 (2017).

82. Plante, T. G. *et al.* Does exercise environment enhance the psychological benefits of exercise for women? *Int. J. Stress Manag.* **14**, 88–98 (2007).

83. American College of Sports Medicine. *ACSM's guidelines for exercise testing and prescription*. (Wolters Kluwer Health, 2018).

84. Garber, C. E. *et al.* Quantity and quality of exercise for developing and maintaining cardiorespiratory, musculoskeletal, and neuromotor fitness in apparently healthy adults.

Med. Sci. Sport. Exerc. **43**, 1334–1359 (2011).

85. Oliveira, B. R. R., Santos, T. M., Kilpatrick, M., Pires, F. O. & Deslandes, A. C. Affective and enjoyment responses in high intensity interval training and continuous training: A systematic review and meta-analysis. *PLoS One* **13**, e0197124 (2018).

86. Greene, D. R., Greenlee, T. A. & Petruzzello, S. J. That feeling I get: Examination of the exercise intensity-affect-enjoyment relationship. *Psychol. Sport Exerc.* **35**, 39–46 (2018).

87. Gomes, A. R., Gonçalves, A. M., Maddux, J. E. & Carneiro, L. The intention-behaviour gap: An empirical examination of an integrative perspective to explain exercise behaviour. *Int. J. Sport Exerc. Psychol.* **16**, 607–621 (2018).

88. Rhodes, R. E. & Dickau, L. Experimental evidence for the intention-behavior relationship in the physical activity domain: A meta-analysis. *Heal. Psychol.* **31**, 724–727 (2012).

89. Dishman, R. K., Sallis, J. F. & Orenstein, D. R. The determinants of physical activity and exercise. *Public Health Reports* vol. 100 158–171 (1985).

90. Joyner, M. J., Wallin, B. G. & Charkoudian, N. Sex differences and blood pressure regulation in humans. *Exp. Physiol.* **101**, 349–355 (2016).

91. Koenig, J. & Thayer, J. F. Sex differences in healthy human heart rate variability: A meta-analysis. *Neuroscience and Biobehavioral Reviews* vol. 64 288–310 (2016).

92. Bowler, D. E., Buyung-Ali, L. M., Knight, T. M. & Pullin, A. S. A systematic review of evidence for the added benefits to health of exposure to natural environments. *BMC Public Health* **10**, 456 (2010).

93. Szuhany, K. L., Bugatti, M. & Otto, M. W. A meta-analytic review of the effects of exercise on brain-derived neurotrophic factor. *J. Psychiatr. Res.* **60**, 56–64 (2015).

94. Miranda, M., Morici, J. F., Zanoni, M. B. & Bekinschtein, P. Brain-derived neurotrophic

factor: A key molecule for memory in the healthy and the pathological brain. *Front. Cell. Neurosci.* **13**, 1–25 (2019).

Appendices

Appendix A Summary of Methods of Included Articles

Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment: Outdoors	Exercise Environment: Indoors	Exercise Protocol: Outdoors	Exercise Protocol: Indoors	Outcomes
Byrka & Ryczko, 2018[79] **Poland, University**	Dancers with >3mo of salsa dance experience 64 (29.42 ± 8.94) 79% female 21% male	RCT	Gravel area in public park surrounded by old trees	Dance studio with mirrors and natural light	40min dance session led by instructor blinded to hypothesis 10min warm-up 5min simple dance movements 20min main choreography 5min stretching Self-selected intensity		Objective Exercise Intensity
Calogiuri et al., 2015[69] **Norway, University**	Healthy employees from 2 workplaces (Assessment group only) 5 (33.4 ± 4.0) 80 % female 20 % male	N-RXT	Forest area near workplace with trees and grass, few built elements	Gym hall, curtain-covered windows and artificial light	25min of biking around 6082m track including steep uphill section that had to be walked Vigorous intensity (70% HRR)	25min biking on cycle ergometer – bench step ups provided to reproduce uphill walking section Vigorous intensity (70% HRR)	Perceived Exertion, Enjoyment; Objective Exercise Intensity Confirmation
Calogiuri et al., 2018[70] **Norway, University**	Healthy university students and employees 26 (26 ± 8) 54% female 46% male	N-RXT	Paved trail along a large river near the university, with built elements like buildings and football field	Laboratory visual outdoor stimuli displayed using virtual reality headset	10min walk on out-and-back route Comfortable, self-selected intensity	10min walk on manually-driven treadmill Comfortable, self-selected intensity	Objective Exercise Intensity, Performance, Enjoyment; Perceived Exertion Confirmation
Carvalho et al., 2010[49] **Sweden, University-related Hospital**	Recently recovered stroke patients stratified by clinical walking speed: (A) speed <0.8m/s; (B) speed ≥0.8m/s. **A**:10 (60 ± 3) 40% female 60% male **B**: 26 (60 ± 4) 31% female 69% male	RXT	Walkway in calm garden and quiet neighborhood with concrete tile walking surface	**Basement:** Empty corridor absent of other patients or health professionals **Clinical:** Corridor in rehabilitation unit	**30m walk test** 1. Self-selected intensity 2. Maximal intensity **6min walk test** Maximal Intensity		Performance

Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment		Exercise Protocol		Outcomes
			Outdoors	Indoors	Outdoors	Indoors	
Chu et al., 2010[63] **Taiwan, University**	Healthy university students stratified into a high-exercise-experience group (HG) or low-exercise-experience group (LG) based on university major and experience exercising **HG**: 12 (22.33 ± 0.65) **LG**: 14 (21.64 ± 1.01) 100% male	N-RXT CB	Outdoor running track	NR	5min walk 20min run around track 5min walk Walk: light intensity Run: moderate, self-selected intensity	Treadmill (0% incline): 5min walk 20min run 5min walk Walk: light intensity (4.8km/h) Run: moderate, self-selected intensity	Objective Exercise Intensity
Dasilva et al., 2011[71] **Italy, University**	Healthy volunteers between 18-30 years old 34 (**F**: 22.5 ± 2.6) (**M**: 24 ± 3.3) 50% female 50% male	N-RXT CB	Standard outdoor 400m tartan track	Laboratory	20min walk Self-selected intensity	20min treadmill walk Self-selected intensity	Objective Exercise Intensity, Performance; Perceived Exertion Confirmation
Farias et al., 2018[65] **Brazil, University**	Older adults (>60) recruited from "My Best Age" study 15 (65.4 ± 5.1) 73% female 27% male	RXT	**Natural:** Beach **Mixed:** Olympic track-field	Gym	5min warm-up 20min walk 5min recovery Self-selected intensity		Performance; Perceived Exertion Confirmation
Flowers et al., 2018[80] **UK, University**	Undergraduate students 60 (19.9 ± 4.26) 32% female 68% male	RCT	Edge of a large natural sports field, consisting of mostly flat and grass covered expansive area with interspersed trees and hedge perimeter	Laboratory facing a blank screen and a light grey wall, with equipment and furniture removed from view	15min of cycling on ergometer Moderate intensity (50% HRR)		Perceived Exertion, Performance, Attitudes to Green Exercise

Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment		Exercise Protocol		Outcomes
			Outdoors	Indoors	Outdoors	Indoors	
Focht, 2009[72] **USA, University**	Physically-active university students 35 (22.43 ± 2.74) 100% female	RXT	University campus using paths & sidewalks immediately surrounding laboratory building	Laboratory	10min walk on marked paths Moderate, self-selected intensity	10min treadmill walk Moderate, self-selected intensity	Objective Exercise Intensity, Enjoyment, Intention for Future Exercise; Perceived Exertion Confirmation
Harte & Eifert, 1995[67] **Australia, University**	Amateur triathletes & marathon runners 10 (27.1 ± NR) 100% male	N-RXT CB	University campus	Laboratory with high-set windows	12km run around pre-determined route Vigorous intensity (16km/hr)	45min treadmill run with 2 conditions: **Ext Stim** = outdoor sounds; **Int Stim** = breathing sounds) Vigorous intensity; 15min at 10-12km/hr + 30min 15-18km/hr	Perceived Exertion, Adrenaline, Noradrenaline, Cortisol, Systolic Blood Pressure, Enjoyment
Jang & So, 2017[81] **South Korea, University**	Experienced taekwondo players 18 (22.55 ± 0.98) 22 % female 78 % male	Q-RCT	Mountain near the university (Subjective appraisal of environment = 9.22 ± 1.09)	Taekwondo training gymnasium (Subjective appraisal of environment = 1.33 ± 0.7)	40min taekwondo training; reverse turning kicks in response to a whistle: 3x10min continuous exercise with 5min breaks Self-selected training intensity		Quality of Environment; Perceived Exertion Confirmation
LaCaille et al., 2004[68] **USA, University**	Individuals who ran ~15miles/wk 60 (26.8 ± 8.93) 63% female 37% male	N-RXT CB	Outdoor flat road route	**TM:** Exercise laboratory **OG:** Indoor 200m track	5km run Self-selected intensity	5km run on treadmill (1% grade) or indoor track Self-selected intensity	Performance; Perceived Exertion Confirmation
McMurray et al., 1988[73] **USA, University**	Experienced Runners (VO_{2max}: 59 ± 3SEM mL/kg) 8 (NR) 100% male	N-RXT	Fairly level, predetermined course; first and last mile completed on 400m running track	Laboratory	10-mile run 70% VO_{2max}	10-mile treadmill run 70% VO_{2max}	Objective Exercise Intensity, Perceived Exertion, Plasma Beta Endorphins, Blood Lactate; Objective Exercise Intensity Confirmation

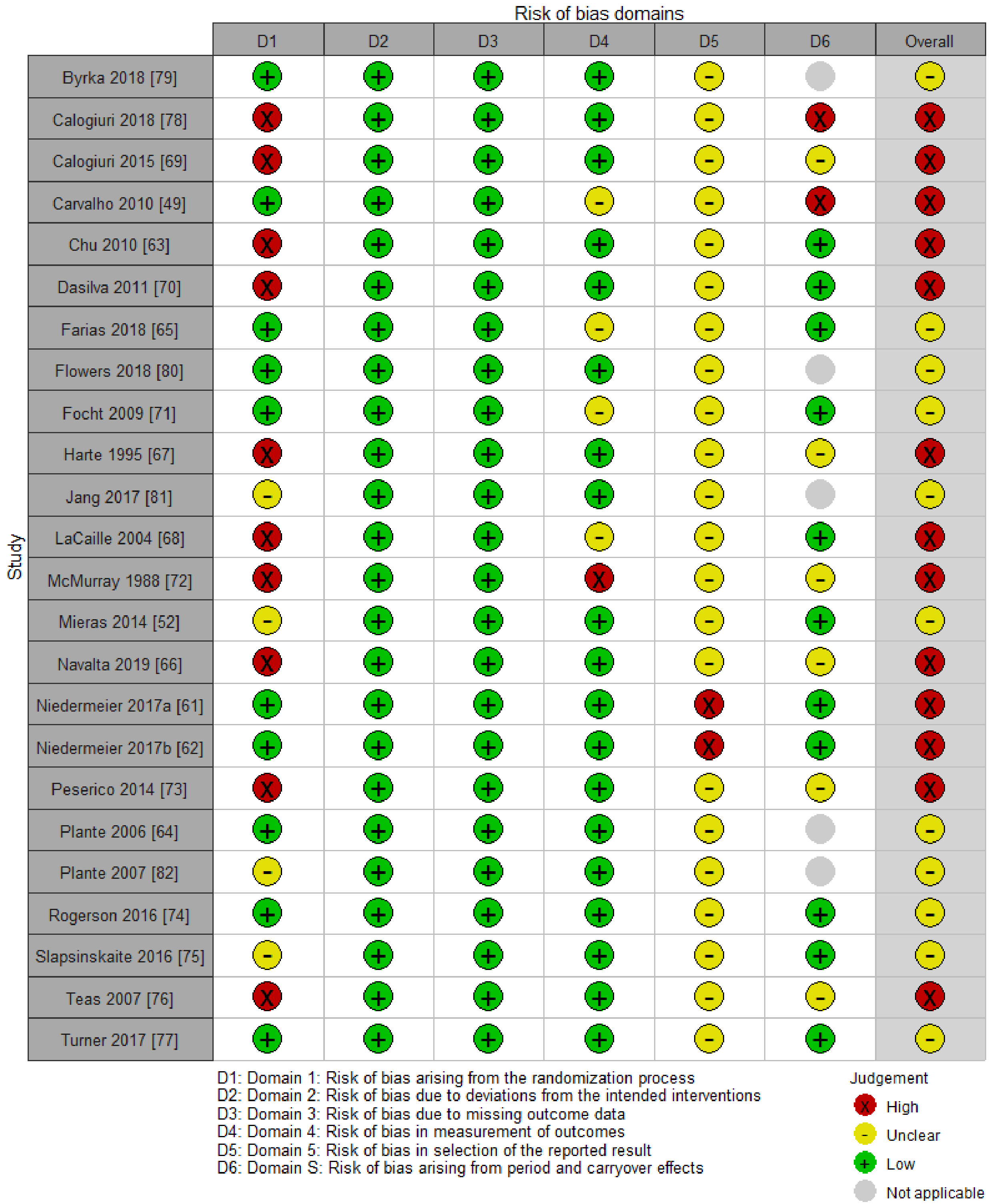
Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment		Exercise Protocol		Outcomes
			Outdoors	Indoors	Outdoors	Indoors	
Mieras et al., 2014[52] **USA, University**	Recreationally-trained cyclists 12 (37 ± 2) 100% male	RXT	Paved, recreational trail	Exercise physiology laboratory	40km training ride on relatively flat, paved trail Constant, self-selected training intensity	40km training ride on cycle ergometer matched to outdoor trail with software Constant, self-selected training intensity	Objective Exercise Intensity, Performance, Body weight change, Core Body Temperature, Estimated Sweat Rate, Skin Temperature, Thermal Gradient Perceived Exertion Confirmation
Navalta et al., 2019[66] **USA, University**	NR 10 (29.2 ± 7.3) 70% female 30% male	N-RXT	**Green:** Mountain forest hiking path in National Recreation Area **Brown:** Desert Trail in National Conservation Area **Below sea level:** Train near dried lakebed in National Park **Urban:** On campus, near busy intersection	Exercise physiology laboratory	30min walk Self-selected intensity		Heart Rate Recovery, Diastolic Blood Pressure, Systolic Blood Pressure
Niedermeier et al., 2017a[61] **Austria, University**	Healthy, active adults 42 (32 ± 12) 48% female 52% male	RXT	Mountain hiking area with single trails and forest roads to mountain hut with a view	Fitness centre	12km hiking route (600m elevation gain); approx. 90min uphill & 70min downhill; Moderate, self-selected intensity	90min treadmill walk (10% incline) + 70min treadmill walk (no incline); Moderate, self-selected intensity	Objective Exercise Intensity, Heart Rate Recovery; Perceived Exertion Confirmation

Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment		Exercise Protocol		Outcomes
			Outdoors	Indoors	Outdoors	Indoors	
Niedermeier et al., 2017b[62] **Austria, University**	Healthy, active adults 42 (32 ± 12) 48% female 52% male	RXT	Mountain hiking area with single trails and forest roads to mountain hut with a view	Fitness centre	12km hiking route (600m elevation gain); approx. 90min uphill & 70min downhill; Moderate, self-selected intensity	90min treadmill walk (10% incline) + 70min treadmill walk (no incline); Moderate, self-selected intensity	Objective Exercise Intensity, Cortisol, Diastolic Blood Pressure, Systolic Blood Pressure, HRV; Perceived Exertion confirmation
Peserico & Machado, 2014[74] **Brazil, University**	Experienced endurance runners 18 (25.4 ± 3.3) 100% male	N-RXT	Outdoor 400m athletics track	NR	1hr distance trial Vigorous-maximal self-selected intensity	1hr distance trial on treadmill (1% incline) Vigorous-maximal self-selected intensity	Objective Exercise Intensity, Performance Perceived Exertion Confirmation
Plante et al., 2006[64] **USA, University**	Undergraduate psychology students 112 (NR) 58 % female 42 % male	RCT	University campus	Laboratory; visual outdoor stimuli displayed by projector	20min walk on pre-determined route on campus, Moderate intensity (~4.8km/h)	20min treadmill walk with projection following same route as outdoor Moderate intensity (4.3-5.6km/h)	Enjoyment
Plante et al., 2007[82] **USA, University**	Undergraduate students 88 (19.31 ± 0.94) 100% female	RCT	University campus	University fitness facility	20min walk around predetermined route Moderate intensity; HR between 120-140 bpm (60-70% HR_{Max})	20min walk on treadmill Moderate intensity; HR between 120-140 bpm (60-70% HR_{Max})	Enjoyment
Rogerson et al., 2016[75] **UK, University**	Healthy adults, recruited in pairs 24 (35.1 ± 20.1) 79% female 21% male	RXT	University sports field, lined with trees	Laboratory	15min of cycling on ergometer Moderate intensity (50% HRR)		Perceived Exertion, Enjoyment, Intention for Future Exercise
Slapsinskaite et al., 2016[76] **Spain, University**	Physically-active university students 13 (21.69 ± 2.81#) 23% female 77% male	RXT	Ergometer near roadway and park with sparse vegetation and mountain view	Human performance laboratory with no view of outdoors	Incremental + constant power protocol to exhaustion on cycle ergometer Constant power phase at vigorous intensity (RPE=15)		Performance

Author, Date Location, Setting	Sample N (Age ± SD) % Sex/Gender*	Study Design	Exercise Environment		Exercise Protocol		Outcomes
			Outdoors	Indoors	Outdoors	Indoors	
Teas et al., 2007[77] **USA, Public Cancer Centre**	Healthy, non-smoking postmenopausal women 19 (58 ± 4) 100% female	N-RXT	Grassy area lined with brick paths, old trees, and flower beds	Exercise laboratory located in basement, with environment similar to commercial gym	1hr walk around pre-determined route Comfortable, self-selected intensity	1hr walk on treadmill Comfortable, self-selected intensity	Performance
Turner & Stevinson, 2017[78] **UK, University**	Healthy members of running clubs 22 (33 ± 8.3) 36% female 64% male	RXT	Large woodland area with waymarked trails; other users present	Large, modern, air-conditioned fitness suite; other users present	6km run with small climb and short descent; 3km at self-selected steady state + 3km as fast as possible self-selected intensity	6km treadmill run at 1% incline 3km at self-selected steady state + 3km as fast as possible self-selected intensity	Performance; Perceived Exertion Confirmation

Appendix A Summary of methods table. Note: *: Sex or gender not consistently distinguished or reported. Acronyms: RCT: Randomized comparative trial; N-RXT: Non-randomized crossover trial; HRR: Heart rate reserve; RXT: Randomized crossover trial; CB: Counterbalanced; NR: Not reported; F: Female, M: Male; Q-RCT: Quasi-randomized comparative trial; TM: Treadmill; OG: Over-ground (i.e., through space as on indoor track); #: denotes calculated value; RPE: Rating of Perceived Exertion (Borg Scale 6-20).

Appendix B Risk of Bias Traffic Light Figure



Appendix C Summary of Findings

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results			
							Before Mean (SD)	**During Mean (SD)**	**After Mean (SD)**	**Findings**
EIC	69	G	Gym	C	V	HR based on HRR (bpm)		O: 126c (14.58#) I: 120c (20.62#)		NS ($d_{\#}$ = 0.34)
	73	GU	Lab	R	V	Mean VO_2 (mL/kg/min)		O: 41.9 (7.07#) I: 40.9 (3.68#)		NS ($d_{\#}$ = 0.18)
Perceived Exertion Confirmation	**52**	G	Lab	C	TS	Borg Scale (6-20)		O: 13.7 (.4) I: 13.7 (.3)		NS ($d_{\#}$ = 0.00)
	78	G	FC	R	SVM	Borg Scale (6-20)	-	NR	-	NS
	81	Mt	Gym	T	TS	Borg Scale (6-20)		O: 9.22 (2.63) I: 15.11 (2.08)		↓Outdoors vs. Indoors (p<.001; $d_{\#}$ = -2.48)
	61	Mt	FC	O:H I:W	MS	Borg Scale (6-20)	O: 7.2 (2.31#) I: 7.3 (1.82#)	O: 11.0 (2.98#) I: 11.5 (2.48#)	O: 9.8 (1.82#) I: 10.3 (2.15#)	NS first half ($d_{\#}$ = -0.18) NS second half ($d_{\#}$ = -0.25)
	62	Mt	FC	O:H I:W	MS	Borg Scale (6-20)		O: 10.4 (1.6) I: 10.9 (2.1)		NS ($d_{\#}$ = -0.27)
	70	GU	Lab (VR)	W	S	Borg Scale (6-20)	-	NR	-	↑Indoor vs Outdoor (p<.001)
	72	GU	Lab	W	MS	Borg Scale (6-20)		O: 10 (1.28) I: 10.33 (1.6)		NS ($d_{\#}$ = -0.23)
	68	U	Lab IT	R	S	Borg Scale (6-20)			O: 13.28(1.85) I-TM: 14.75(2.01) I-IT: 13.93(1.84)	↓Outdoor vs TM (p<.01, d = -0.76) ↓Outdoor vs. IT (p<.05, d = -0.35)
	65	NW OT	Gym	W	S	OMNI Scale (0-10)		$Total_{5min}$: 3.4 (.2)	$Total_{20min}$: 4.2 (.2)	NS
	71	T	Lab	W	S	Borg Scale (6-20)	-	NR	-	*Environment X Time Interaction:* ↓last 10 min vs first 10 min, Outdoor vs Indoor (p<.01, η_p^2 = .262) *Main Effect of Environment:* ↓Outdoor vs Indoor (p<.05, η_p^2 = .166)
	74	T	NR	R	SVM	Borg Scale (6-20)			O: 19 (0.8) I: 19 (1.1)	NS ($d_{\#}$ = 0.00)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results			
							Before Mean (SD)	During Mean (SD)	After Mean (SD)	Findings
Exercise Intensity	***Relative Values***									
	72	GU	Lab	W	MS	% Max HR		O: 59.21 (8.1) I: 57.11 (8.3)		NS ($d_{\#}$ = 0.26)
	63 (HG)	T	NR	R	MS	%HRR		O: 80.55 (6.96) I: 65.24 (10.74)		↑ Outdoors (p<.001; $d_{\#}$ = 1.69)
	63 (LG)	T	NR	R	MS	%HRR		O: 88.49 (10.31) I: 82.91 (12.45)		↑ Outdoors (p=.02; $d_{\#}$ = 0.49)
	71	T	Lab	W	S	%HRR	-	NR	-	*Environment X Time Interaction:* ↓last15min vs first 5min, Outdoor vs Indoor (p<.01, η_p^2 = .142)
						$\%VO_2R$	-	NR	-	*Environment X Time Interaction:* ↓last15min vs first 5min, Outdoor vs Indoor (p<.01, η_p^2 = .408) *Main Effect of Environment:* ↓Outdoor vs Indoor (p<.01, η_p^2 = .231)
	Absolute Values									
	52	G	Lab	C	TS	Mean HR (bpm)		O: 152 (4) I: 143 (6)		↑Outdoors vs. Indoors (p<.05; $d_{\#}$ = 1.77)
	61	Mt	FC	O:H I:W	MS	Mean HR (bpm)		O: 110.9 (17.19#) I: 105.0 (14.38#)		↑Outdoors vs. Indoors during exercise (p<.006; $d_{\#}$ = 0.37)
	62	Mt	FC	O:H I:W	MS	Mean HR (bpm)		O: 111 (17) I: 105 (14)		NS ($d_{\#}$ = 0.39)
	70	GU	Lab (VR)	W	S	Mean HR (bpm)	-	NR	-	NS
						Max HR (bpm)	-	NR	-	NS
	73	GU	Lab	R	V	Mean HR (bpm)		O: 152 (14.14#) I: 160 (14.14#)		NS ($d_{\#}$ = -0.57)
	74	T	NR	R	SVM	Mean HR (bpm)		O: 175 (7.8) I: 178 (7.5)		↑Outdoors vs. Indoors (p=*.016*; $d_{\#}$ = -0.39)
						Max HR (bpm)		O: 188 (7.1) I: 184 (8.1)		↑Outdoors vs. Indoors (p=*.03*; $d_{\#}$ = 0.53)
	Time Spent in Moderate-Vigorous Physical Activity									
	79	G	DS	D	S	3D ActiGraph Accelerometer (minutes of intensity ≥MV)		O: 18.11 (8.69) I: 11.91 (7.00)		↑Outdoor vs. Indoors (p<.01; $d_{\#}$ = 0.78)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results: Before Mean (SD)	Results: During Mean (SD)	Results: After Mean (SD)	Results: Findings
Perceived Exertion	[69]	G	Gym	C	V	Borg Scale (6-20)		O: 8c (3.02#) I: 12# (3.02#)		↓Outdoors vs. Indoors (p=.04; $d_{\#}$ = -1.33)
	80	G	Lab	C	MI	Borg Scale (6-20)		O: 11.20 (1.57) I: 12.40 (2.32)		NS ($d_{\#}$ = -0.61)
	[75]	G	Lab	C	MI	Borg Scale (6-20)		Halfway: O: 11.9 (2.5) I: 12.2 (2.5)	Last Minute: O: 12.9 (2.2) I: 13.5 (2.2)	NS at halfway ($d_{\#}$ = -0.12) NS during last minute ($d_{\#}$ = -0.27)
	[73]	GU	Lab	R	V	Borg Scale (6-20)			O: 14.2 (5.09#) I: 12.8 (2.26#)	NS ($d_{\#}$ = 0.36)
	67	C	Lab	R	V	Borg Scale (0-8)	O: .8 (0.95#) I-IS: .9 (0.63#) I-ES: .6 (0.63#)		O: 4.1 (1.26) I-IS: 7.3 (0.95) I-ES: 5.7 (0.95)	*Main Effect of Condition, covaried by Pre:* ↓Outdoor vs. Indoor-IS (p<.01; $d_{\#}$ = -2.87) NS Outdoor vs Indoor-ES ($d_{\#}$ = -1.43)
Performance	***Average Speed***									
	[77]	G	FC	R	CS	Mean speed (km/h)		O: 5.1 (0.83#) I: 4.3 (0.60)		NS ($d_{\#}$ = 1.10)
	[70]	GU	Lab	W	S	Mean speed (min/km)	-	NR	-	NS
	49 (A)	GU	BH CL	W	S	***30-meter Walk Test*** Self-selected Speed (m/s)		O: 0.51 (0.2) I-BH: 0.52 (0.19) I-CL: 0.51 (0.18)		NS ($d_{\#BH}$ = -0.05; $d_{\#CL}$ = -0.05)
					NM	***30-meter Walk Test*** Max Speed (m/s)		O: 0.6 (0.25) I-BH: 0.6 (0.23) I-CL: 0.61 (0.24)		NS ($d_{\#BH}$ = 0.00; $d_{\#CL}$ = -0.04)
	49 (B)	GU	BH CL	W	S	***30-meter Walk Test*** Self-selected Speed (m/s)		O: 1.31 (0.22) I-BH: 1.30 (0.23) I-CL: 1.26 (0.21)		NS ($d_{\#BH}$ = 0.05; $d_{\#CL}$ = 0.23)
					NM	***30-meter Walk Test*** Max Speed (m/s)		O: 1.69 (0.33) I-BH: 1.68 (0.29) I-CL: 1.66 (0.34)		NS ($d_{\#BH}$ = 0.03; $d_{\#CL}$ = 0.09)
	65	NW OT	Gym	W	S	Mean speed (m/s)		O-NW: 1.4 (0.2) O-OT: 1.4 (0.2) I: 1.4 (0.2)		NS ($d_{\#}$ = 0.00)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results			
							Before Mean (SD)	During Mean (SD)	After Mean (SD)	Findings
Performance	[71]	T	Lab	W	S	Mean speed (m/s)	-	NR	-	*Environment X Time Interaction:* ↑first15min vs last 5min, Outdoor vs Indoor (p<.01, η_p^2 = .467) *Main Effect of Environment:* ↑Outdoor vs Indoor (p<.01, η_p^2 = .666) *Environment X Gender Interaction:* ↑difference between males vs females, indoors vs outdoor (p<.05, η_p^2 = .122)
	[74]	T	NR	R	SVM	Mean speed (m/s)		O: 11.8 (0.8) I: 12.2 (0.8)		↑Outdoors vs. Indoors (p=*.001;* $d_{\#}$ = -0.50)
	Power Output									
	80	G	Lab	C	MI	Power Output (Watts)		O: 121.60 (40.36) I: 106.20 (35.98)		NS ($d_{\#}$ = 0.40)
	52	G	Lab	C	TS	*Power Output* CycleOps Power Tap (Watts)	-	NR	-	↑Outdoors vs. Indoors (p<.001)
	Time to Completion									
	52	G	Lab	C	TS	Time to completion (min)			O: 83 (3.3) I: 96.3 (4.5)	↓Outdoors vs. Indoors (p<.001; $d_{\#}$ = -3.37)
	[78]	G	FC	R	SVM	Time to completion (min)		O: 14.9 (2.88) I: 14.7 (2.91)		NS ($d_{\#}$ = 0.07)
	68	U	Lab IT	R	S	Time to completion (min)		O: 25.56 (4.15) I-TM: 29.60 (4.91) I-IT: 25.83 (3.63)		↑TM vs Outdoor (p<.001, d=0.89) NS Outdoor vs IT ($d_{\#}$ = -0.07) ↑TM vs IT (p<.001, d=0.88)
	Time to Exhaustion									
	[76]	GU	Lab	C	SVM	Time to exhaustion (min)		O: 12.54 (0.14#) I: 11.35 (0.11#)		↑Outdoors vs. Indoors (p<.05, d=.56)
	Distance									
	49 (A)	GU	BH	W	S	***6-minute Walk Test*** Distance (meters)		O: 178 (64) I-BH: 175 (67)		NS ($d_{\#}$ = 0.05)
	49 (B)	GU	BH	W	S	***6-minute Walk Test*** Distance (meters)		O: 463 (84) I-BH: 452 (82)		↑Outdoor vs Indoor (p=.01; $d_{\#}$ = 0.13)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results: Before Mean (SD)	Results: During Mean (SD)	Results: After Mean (SD)	Results: Findings
Neuroendocrine Response	***Cortisol***									
	62	Mt	FC	O:H I:W	MS	Salivary Assay	O: 4.7 (3.7) I: 5.0 (3.5)		O: 1.8 (1.2) I: 1.8 (1.1)	*Environment x Time Interaction* (p=.032; $d_{\#}$ = 0.00)): NS comparisons for environment
	67	UC	Lab	R	V	Urinary Assay	-	NR	-	↓Outdoor vs. Indoor-IS (p<0.01)
	Catecholamines									
	67	UC	Lab	R	V	***Adrenaline*** Urinary Assay	-	NR	-	NS
						Noradrenaline Urinary Assay	-	NR	-	↓Outdoor vs. Indoor-IS (p<0.01)
	Other									
	73	GU	Lab	R	V	***Plasma beta-endorphin Concentration*** Radioimmunoassay (pm/L)	O: 6.5 (1.98#) I: 6.0 (1.98#)		O: 21.8 (10.75#) I: 19.0 (10.75#)	NS ($d_{\#}$ = 0.26)
						Blood Lactate (mmol/L)			O: 4.1 (1.7#) I: 1.8 (.85#)	↑Outdoors vs. Indoors (p<.05; $d_{\#}$ = 1.71)
Cardiovascular Response	***Systolic Blood Pressure***									
							Estimated Mean (SE)			
	66	G U B BSL	Lab	W	S	Automated Sphygmomanometer (mmHg)	*Covaried by baseline measure, distance walked, & environmental conditions		O-G: 116.3 (8.3) O-B: 118.8 (3.1) O-BSL: 124.4 (6.3) O-U: 118.7 (3.8) I: 118.9 (2.5)	NS
	62	M	FC	O:H I:W	MS	Oscillometric Device (mmHg)	O: 127.2 (11.6) I: 123.5 (13.2)		O: 121.3 (11.2) I: 119.0 (11.0)	*Main Effect of Environment:* ↑Outdoors vs. Indoors (p=.006; $d_{\#}$ = 0.21)
	67	C	Lab	R	V	Automated sphygmomanometer (mmHg)	-	NR	-	NS
	Diastolic Blood Pressure									
							Estimated Mean (SE)			
	66	G U B BSL	Lab	W	S	Automated Sphygmomanometer (mmHg)	*Covaried by baseline measure, distance walked, & environmental conditions		O-G: 75.8 (5.6) O-B: 73.4 (2.1) O-BSL: 73.1 (4.2) O-U: 70.2 (2.6) I: 69.7 (1.8)	NS
	62	Mt	FC	O:H I:W	MS	Oscillometric Device (mmHg)	O: 77.7 (7.6) I: 75.8 (7.4)		O: 78.3 (7.8) I: 72.6 (8.1)	*Main Effect of Environment:* ↑Outdoors vs. Indoors (p<.001; $d_{\#}$ = 0.72)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results			
							Before Mean (SD)	During Mean (SD)	After Mean (SD)	Findings
Heart Rate Recovery										
							Estimated Mean (SE)			
Cardiovascular Response	**66**	G U B BSL	Lab	W	S	Post-Exercise Resting Heart Rate (bpm)	*Covaried by baseline measure, distance walked, & environmental conditions		O-G: 62.6 (10.9) O-B: 65.4 (4.1) O-BSL: 79.1 (8.2) O-U: 64.9 (5.1) I: 67.4 (3.4)	NS
	61	Mt	FC	O:H I:W	MS	Post-Exercise Resting Heart Rate (bpm)	O: 71.5 (12.73#) I: 72.8 (11.57#)		O: 75.2 (19.18#) I: 75.3 (18.19#)	*Environment X Time Interaction:* ↑Outdoors vs. Indoors Pre→Post (p=.001, d=.59) *Main Effect of Environment:* NS ($d_{\#}$ = -0.01)
Heart Rate Variability										
	62	Mt	FC	O:H I:W	MS	SDNN (ms)	O: 76.3 (40.9) I: 70.7 (40.6)		O: 85.4 (47.5) I: 77.9 (41.6)	NS ($d_{\#}$ = 0.30)
						RMSSD (ms)	O: 58.4 (44.5) I: 51.5 (43.9)		O: 68.6 (56.6) I: 57.3 (41.0)	NS ($d_{\#}$ = 0.13)
						Total power (ms^2)	O: 9497 (10,715) I: 7906 (10,552)		O: 10,504 (10,254) I: 10,055 (10,574)	NS ($d_{\#}$ = -0.13)
						Power in Low Frequency Range (ms^2)	O: 2331 (2399) I: 1973 (2934)		O: 2967 (2729) I: 2614 (2505)	NS ($d_{\#}$ = 0.32)
						Normalized Low Frequency Power	O: 63.8 (17.7) I: 66.5 (14.7)		O: 66.4 (16.4) I: 71.1 (13.1)	NS ($d_{\#}$ = -0.32)
						Power in High Frequency Range (ms^2)	O: 1785 (2538) I: 1489 (2659)		O: 2409 (3568) I: 1548 (1996)	NS ($d_{\#}$ = 0.23)
						Normalized High Frequency Power	O: 36.2 (17.7) I: 33.5 (14.7)		O: 33.6 (16.4) I: 28.9 (13.1)	NS ($d_{\#}$ = 0.17)
						Low Frequency/High Frequency Power Ratio	O: 292.3 (314.9) I: 293.1 (275.0)		O: 313.7 (316.9) I: 356.8 (361.1)	NS ($d_{\#}$ = 0.04)
Thermoregulation	**52**	G	Lab	C	TS	***Body Weight*** Change over trial (Kg)	All: 82.1 (4.8)		O: 82.5 (4.8) I: 80.9 (4.7)	NS ($d_{\#}$ = 0.34)
						Estimated Sweat Rate (ml/h)		O: 1506.7 (152.3) I: 1332.5 (109.6)		NS ($d_{\#}$ = 1.31)
						Core Body Temperature Ingestible Capsule (^{O}C)		O: 37.7 (0.2) I: 37.4 (0.1)		NS ($d_{\#}$ = 1.90)
						Skin Temperature Thermistor Patch (^{O}C)		O: 31.4 (0.3) I: 33.0 (0.2)		↓Outdoors vs. Indoors (p<.001; $d_{\#}$ = -6.28)
						Thermal Gradient Core/Skin Diff (^{O}C)		O: 6.5 (0.3) I: 4.4 (0.2)		↑Outdoors vs. Indoors (p<.001; $d_{\#}$ = 8.24)

Grouping	Article	Outdoor Environment	Indoor Environment	Exercise Type	Exercise Intensity	Measure	Results			
							Before Mean (SD)	During Mean (SD)	After Mean (SD)	Findings
Enjoyment	[69]	G	Lab	C	V	Single Item (0-10)			O: 6c (2.01#) I: 4c (2.51#)	NS ($d_{\#}$ = 0.88)
	[75]	G	Lab	C	MI	Single Item (0-100%)			O: 65.9 (23.5) I: 64.2 (23.4)	NS ($d_{\#}$ = 0.07)
	[70]	GU	Lab (VR)	W	S	Single Item (0-10)			O: 7.69 (1.78) I: 3.96 (2.32)	↑Outdoor vs Indoor (*p*<.001; ($d_{\#}$ = 1.80)
	[72]	GU	Lab	W	MS	Single Item (1-10)			O: 7.91 (1.17) I: 6.57 (1.42)	↑Outdoor vs Indoor (*p*<.001, *d*=1.03)
	67	UC	Lab	R	V	Single Item (1-10)			O: 6.4 (1.26#) I-IS: .9 (.95#) I-ES: 3.2 (1.58#)	↑Outdoor vs. Indoor (*p*<.01; $d_{\#I\text{-}IS}$ = 4.93; $d_{\#I\text{-}ES}$ = 2.24)
	64 (F)	UC	Lab (VR)	W	MI	PACES Activity Enjoyment Scale (7-126)			O: 93.88 (14.37) I: 63.55 (14.12)	*Main Effect of Environment* : ↑Outdoors vs. Indoors (*p*<.05)
	64 (M)	UC	Lab (VR)	W	MI	PACES Activity Enjoyment Scale (7-126)			O: 78.70 (22.83) I: 78.60 (21.75)	
	82	UC	FC	W	MI	PACES Activity Enjoyment Scale (7-126)			I: 88.77 (13.77) O: 96.41 (14.97)	↑Outdoors vs. Indoors (*p*<.05)
FEI	[75]	G	Lab	C	MI	Single Item (0-100%)			O: 68.3 (22.0) I: 69.7 (25.4)	NS ($d_{\#}$ = -0.06)
	[72]	GU	Lab	W	MS	Single Item (0-100%)			O: 85.14 (13.79) I: 69.43 (20.28)	↑Outdoors (*p*<.001, *d*=.92)
Perception	**80**	G	Lab	C	MI	***Green Exercise Attitude*** BAGE questionnaire (5 item, 1-7)	O: 5.67 (0.70) I: 5.58 (0.68)		O: 6.22 (0.67) I: 6.28 (0.44)	NS ($d_{\#}$ = -0.11)
	81	Mt	Gym	T	TS	***Perceived Quality*** Single Item (1-10)			O: 9.22 (1.09) I: 1.33 (0.70)	↑Outdoors vs. Indoors (*p*<.001; $d_{\#}$ = 8.61)

Appendix C Summary of findings table. Note. Ordered by quality of environment within outcome grouping. Acronyms: SD: Standard deviation; EIC: Exercise Intensity Confirmation; G: Green natural environment; C: Cycling; V: Vigorous intensity; HR: Heart Rate; HRR: Heart rate reserve; bpm: beats per minute; O: Outdoor condition; I: Indoor condition; #: Denotes calculated value; NS: Not significant; $d_{\#}$: Cohen's *d* effect size calculated from provided summary data; GU: Green-urban mixed environment; R: Running; TS: Self-selected training intensity; FC: Fitness centre; SVM: Self-selected vigorous-maximal intensity; NR: Not reported; Mt: Mountainous natural environment; T: Taekwondo; H: Hiking; W: Walking; MS: Moderate, self-selected intensity; VR: Virtual reality or visual stimulus; S: Self-selected intensity; U: Urban environment; IT: Indoor track; TM: Treadmill; NW: Natural environment by water; OT: Outdoor Track; HG: High exercise experience group (stratified sample); LG: Low exercise experience group (stratified sample); VO_2R: Oxygen consumption reserve; DS: Dance studio; D: Dance; MI: Moderate intensity; I-IS: Indoor condition with internal auditory stimuli; I-ES: Indoor condition with external auditory stimuli; CS: Comfortable, self-selected intensity; BH: Basement hallway; CL: Clinical setting; NM: Near maximal to maximal exercise intensity; UC: University campus; F: Female (stratified sample); M: Male (stratified sample); B: Brown natural environment; BSL: Brown natural environment below sea level; SE: Standard error; FIE: Future Intention for Exercise; BAGE: Belief about Green Exercise.