Improving the Analysis and Reporting of Studies of Nature-Based Adventure Interventions: A Review of Studies Published in JAEOL

Claudio D. Rosa[1], Talisson Santos Chaves[2], Silvia Collado[3], and Nevin J. Harper[4]

[1]Department of Development and Environment, State University of Santa Cruz

[2]Department of Physiological Sciences, Federal University of São Carlos

[3]Department of Psychology and Sociology, University of Saragossa

[4]School of Exercise Science, Physical & Health Education, University of Victoria

Author note

Claudio D. Rosa http://orcid.org/0000-0002-1939-2716

Talisson Santos Chaves https://orcid.org/0000-0002-4795-6697

Silvia Collado https://orcid.org/0000-0002-3905-0617

Nevin J. Harper https://orcid.org/0000-0002-7828-2349

Correspondence concerning this article should be addressed to Claudio D. Rosa, claudio2008ilheus@hotmail.com, Phone: 55 73 991123063

Word count: 10938

Abstract

Nature-based adventure (NBA) interventions can be broadly understood as any intervention of adventure occurring in a natural setting not necessarily involving a licensed mental health professional. A growing body of evidence suggests that some NBA interventions improve people's health. Although encouraging, the strength of this evidence is often impaired by the non-random allocation of participants to groups and the lack of comparison groups. These design limitations need additional care during data analysis. Moreover, some studies of NBA interventions omit necessary information for the interpretation of their results, which may lead to inaccurate conclusions and hinder replicability. This information includes (a) a detailed description of the interventions and settings where they took place, (b) dropout numbers and reasons, (c) adverse events of NBA interventions (e.g., leg fracture), and (d) declaration of conflict of interest and research funding. To exemplify these analytical and reporting issues, we evaluated all quantitative studies assessing the effect of an NBA intervention published in the *Journal of Adventure Education and Outdoor Learning* from 2019 to 2021. Based on the present review and findings, we provide suggestions to improve the analysis of studies of NBA interventions and discuss the need to report these necessary elements in future studies.

Keywords: adventure, methodology, nature-based intervention, wilderness

Improving the Analysis and Reporting of Studies of Nature-Based Adventure Interventions: A Review of Studies Published in JAEOL

There is growing evidence indicating that spending time outdoors in contact with nature can enhance physical and mental health (Gladwell et al., 2013; Moreton et al., 2021; Roberts et al., 2020; Rosa et al., 2021). For example, time spent in nature has been related to lower psychological distress (Astell-Burt et al., 2014), higher positive emotional states (Corazon et al., 2019), and improved attention (Ohly et al., 2016) in the adult population. This has led researchers and practitioners to encourage people to spend time in natural environments as a way to enhance their wellbeing (White et al., 2019). Among the different activities that can be conducted in nature to improve people's health, nature-based adventure (NBA) interventions have received growing attention from the scientific community (Bettmann et al., 2016; Bowen & Neill, 2013; Feletti & Bonato, 2020; Fernee et al., 2017, 2021; Houge Mackenzie & Goodnow, 2021; Moreton et al., 2021; Trundle & Hutchinson, 2021).

NBA interventions refer to the facilitation of one or more adventure activities in a natural setting; an environment that has a significant presence of soil (e.g., mountains), vegetation (e.g., trees), and/or water, and that is not highly altered by humans (Wohlwill, 1983). Adventure activities share several characteristics such as uncertain outcomes, skill development and novelty (Bowen & Neill, 2013; Houge Mackenzie & Goodnow, 2021). What differentiates adventure activities in nature from other activities in natural environments is the risk they involve (Bowen & Neill, 2013; Houge Mackenzie & Goodnow, 2021; Javorski & Gass, 2013). For example, in terms of getting lost, walking in urban green spaces entitles lower risk than walking in a wilderness environment, though, urban

environments may present other risks such as getting hit by a car. Activities such as backpacking, camping, surfing, rock climbing and skiing are normally seen as adventurous (Bettmann et al., 2016; Bowen & Neill, 2013). Conceptually, NBA interventions can encompass wilderness therapy (therapy conducted in wilderness environments) and nature-based adventure therapy, which is the combination of adventure activities and therapy taking place in natural settings (Bowen & Neill, 2013; Fernee et al., 2017; Houge Mackenzie & Goodnow, 2021). Following the approach of previous researchers (Rosa et al., 2021), for this study, we refer to therapy in a broad sense, including any activity provided to improve an individual's health. Therapeutic activities may include, for example, group work, psychotherapy, and physical exercises – all of which may be integrated with NBA (Bowen & Neill, 2013; Fernee et al., 2017; Houge Mackenzie & Goodnow, 2021; Rosa et al., 2021). Hence, not all NBA interventions are necessarily led by licensed mental health professionals.

Two main theories may help to explain the effect of NBA interventions on people's health: Attention Restoration Theory (ART, Kaplan, 1995) and Stress Reduction Theory (SRT, Ulrich et al., 1991). The former theory articulates that 'any prolonged mental effort leads to directed attention fatigue and that a compatible natural environment may help people to recover their attentional capacity (Kaplan, 1995, p. 170). This recovery may promote reductions in stress and anxiety, and improvements in mood (Kaplan, 1995; Ulrich et al., 1991). This is in line with SRT which postulates that 'modern humans might have a biologically prepared readiness to quickly and readily acquire restorative responses with respect to many unthreatening natural settings' (Ulrich et al., 1991, p. 208). More recently, the relational restoration theory (RRT, Hartig, 2021) proposes that social resources (e.g., social capital) are depleted daily and can also be restored through people-environment transactions.

In addition to these theories that cover any form of contact with nature, there are more specific theories or frameworks explaining the health benefits of NBA activities (Fernee et al., 2017; Houge Mackenzie et al., 2021; Russell & Farnum, 2004). For example, Russell and Farnum (2004) developed the wilderness therapy treatment milieu model that was further elaborated by Fernee et al. (2017). Fernee et al.'s (2017) wilderness therapy clinical model argues that in wilderness therapy there are three therapeutic components: the wilderness environment, physical self (i.e., the challenges inherent to wilderness therapy that make individuals physically and psychologically stronger), and a psychosocial self that is related to psychological (e.g., counseling) and social (e.g., group activities) stimuli provided in wilderness therapy programs.

In support of these theories and frameworks, some randomized studies have found evidence that NBA interventions can improve people's health. For instance, Sturm et al. (2012) conducted a randomized crossover trial in Germany with adults at risk of committing suicide. In this study, the first group of participants took part in a 9-weeks of mountain hiking program (i.e., NBA intervention) and then moved on to a 9-weeks non-intervention phase. A second group followed the inverse procedure (i.e., non-intervention phase first). During the intervention, participants in both groups continued with their usual treatments, such as psychotherapy and pharmacotherapy. Considering the combined mean depression scores of both groups, there was a greater reduction in depressive symptoms during the mountain hiking phase than during the non-intervention phase. According to the authors (Sturm et al., 2012), each hike was led by two people: the director of the study and an additional person (e.g., a psychotherapist, nurse, or physician). Similarly, Shin et al. (2012) conducted a parallel-group randomized controlled trial (RCT) with alcoholics in South Korea. In this RCT, participants in the NBA group engaged in a 9-day forest program including mountain

climbing, tracking, and orienteering. The comparison group received no intervention. After the intervention period, the NBA group reported, on average, less difficulty in falling asleep and in staying asleep than the comparison group. The intervention providers were not described in this study (Shin et al., 2012). In a later study, Gelkopf et al. (2013) compared 40 sessions of sailing instruction and practice followed by two 3-day outdoor activities including outdoor sleeping and camping (i.e., the NBA group) with a waiting list control group. The authors found that participants in the NBA group held a higher mean emotional quality of life score after the intervention than those in the control group. The intervention was provided by four volunteers under the leadership of the project leader (Gelkopft et al., 2013). None of these volunteers had health or social related academic degrees and monthly supervision was received from a rehabilitation psychologist (Gelkopft et al., 2013).

Methodological Limitations of Studies Assessing NBA Interventions

Despite the existence of theories and some convincing evidence supporting the health benefits of NBA interventions, most of the conclusions regarding the positive health consequences of NBA interventions are impaired by methodological limitations, including a lack of comparison/control groups and, in the presence of a comparison group, the non-randomized allocation of participants (Bowen & Neill, 2013; Harper et al., 2021). For instance, of the 206 unique samples included in Bowen and Neil's (2013) systematic review only 16 (7.8%) were randomly assigned to groups. To exemplify, Kelley et al. (1997) compared people receiving treatment for mental illness plus participating in several NBA activities with people just receiving their treatment as usual. Since participants were not randomly allocated to groups, the authors' finding of a larger improvement in self-efficacy in the NBA group than in the usual care group might be partially explained by confounding variables such as participants' motivation (Kelley et al., 1997). As another example, Walter et

al. (2019) observed a reduction in the mean score of the Generalized Anxiety Disorder 7-Item Scale of participants who engaged in a 6-week program involving surfing. However, due to the lack of a control group, we cannot infer what should have happened to the participants had they not received this intervention (Hernán & Robins, 2020; Rosa & Delabrida, 2021).

The lack of randomization or even a comparison group is not exclusive to studies of NBA interventions. For example, of 13 controlled studies included in a recent meta-analysis of the effect of forest therapy interventions on depressive symptoms, only four of them involved random allocation to groups (Rosa et al., 2021). In fact, due to ethical considerations (e.g., to align the intervention to participants' preference) and high financial resources often needed to conduct randomized studies, most studies of health interventions are not randomized (Gabrielsen et al., 2016; von Elm et al., 2007). Even if a lack of randomization is a common practice in the evaluation of health interventions, one should keep in mind that randomization ensures that differences between groups at the baseline are due to chance. The larger the randomized groups, the more likely it is that they will be similar at baseline (Hernán & Robins, 2020). When groups are similar at baseline, group differences at the end of the intervention cannot be explained by prognostic variables (i.e., factors that predict the outcome, such as disease severity (Hernán & Robins, 2020; Sterne et al., 2016)). When participants are not randomly assigned to the different groups, systematic differences in prognostic variables at the baseline may occur so that a difference between groups at the end of the intervention may be explained by factors other than the intervention (Sterne et al., 2016, 2019).

In addition to the lack of randomization, many studies evaluating the effects of NBA interventions on people's health lack a comparison group (e.g., Bird, 2015; Bowen et al., 2016; Gabrielsen et al., 2019; Gillis & Simpson, 1991; Jakubec et al., 2016; Rogers et al.,

2014; Townsend et al., 2018; Walker et al., 2005; Walter et al., 2019). This methodological shortcoming implies that we cannot infer what would have occurred had the group not received the intervention, therefore causal inference is hard to make (Hernán & Robins, 2020). Because of this, causal claims reported in non-randomized studies should be taken with caution and additional care is needed during the analysis of these studies (Antonakis et al., 2010).

Besides these methodological limitations, some studies examining the effect of NBA interventions lack information needed to improve the interpretation of the intervention's effect, to conduct a risk-benefit assessment, and to replicate studies. This necessary information includes (a) a detailed description of the (co) interventions and settings where they took place, (b) the number and reasons for dropout, (c) adverse events or effects of NBA interventions, and (d) declaration of conflict of interest and research funding. Let's consider, for instance, that the main objective of an intervention study usually is to understand the effect of an intervention on a given outcome (Sterne et al., 2019). Such comprehension is important because practitioners (e.g., recreational therapists) would base their practice on rigorous intervention studies showing important health benefits (Higgins et al., 2019; Sterne et al., 2016, 2019). Nonetheless, without sufficient information about the participants, intervention, comparison group, and outcome, practitioners may not be able to evaluate what was the effect of the intervention and whether this intervention is appropriate for their clients (Moher et al., 2010). A detailed description of the intervention and the setting where it took place is also necessary because without it practitioners and researchers cannot replicate an intervention (Moher et al., 2010). Though other forms of replication exist, direct replication occurs when one repeats a study collecting data from a new sample in a way that duplicates as far as possible the conditions of the former study (American Psychological Association,

2020). Replication may increase certainty when findings are reproduced and promote innovation when they are not (Open Science Collaboration, 2015). Practitioners should also be aware of the risks of an intervention because the decision to deliver (or not) a specific intervention is taken by weighing potential benefits against these risks (Hernán & Robins, 2020; Higgins et al., 2019; Soga & Gaston, 2022).

The Present Study

Given the growing interest in the examination of the effect of NBA interventions in people's health (Benninger et al., 2020; Bettmann et al., 2016; Bowen & Neill, 2013; Greer & Vin-Raviv, 2019; Lackey et al., 2019), and the shortcomings mentioned above, we deem it important to highlight ways to methodically improve studies of NBA interventions and the reporting of findings. This can help future researchers and practitioners interested in the use of adventure activities in nature as a health resource. In this paper, we first provide suggestions to improve the analysis of studies of NBA interventions and then move on to reflecting on why reporting the above-mentioned elements (a) to (d) is relevant. Then, to better illustrate our suggestions, we assessed the analytical approach and reporting of all quantitative studies of NBA interventions published in the *Journal of Adventure Education and Outdoor Learning* (JAEOL, 2019-2021).

Method

This paper is an effort to describe some limitations of studies of NBA interventions and provide suggestions for future research. Our examples of limitations and suggestions to improve studies' data analysis and reporting are based on a selected review of previous studies of NBA interventions and methodological papers, so our study may be described as a narrative review (Greenhalgh et al., 2018). A systematic evaluation of all studies of NBA interventions should improve the understanding of the prevalence of the limitations we

discuss in this manuscript. Our main goal, however, is not to show the prevalence of these limitations but to provide suggestions to avoid them. Whereas our selective review of the literature does not allow us to fully comprehend the frequency of these limitations, we provide multiple examples to illustrate our arguments. We also took advantage of some systematic review methods, such as developing eligibility criteria (Shamseer et al., 2015), to evaluate research published in JAEOL during the 2019-2021 period. The methods used in this evaluation are detailed in a specific section below.

Suggestions to Improve Studies of NBA Interventions

As indicated above, in most studies of NBA interventions there is no comparison group and, in the presence of a comparison group, participants are often not randomly allocated to groups (Bowen & Neill, 2013). If these limitations in the studies' design are not taken into account during data analysis, studies may find inaccurate estimates of effect of the intervention on people's health (Antonakis et al., 2010; Sterne et al., 2016). When participants are not randomly assigned to groups, it is critical to consider the influence of prognostic variables in the results (Sterne et al., 2016), so we discuss herein the relevance and ways to consider such variables in the analyses. Also, some studies lacking randomization to groups or a comparison group run a within-group test of significance (e.g., Cave, 1979; Ilagan et al., 2020). Because the use of this test may lead to equivocal conclusions, we also discuss important elements to be considered when conducting a within-group test of significance (Bland & Altman, 2011). In the next sections, we explain these issues in greater detail. We acknowledge that considering prognostic variables and within-group tests of significance does not completely overcome the drawbacks of non-randomized studies (for further guidance see Antonakis et al., 2010; Marinescu et al., 2018; von Elm et al., 2007). Nonetheless, these are limitations of some studies assessing NBA interventions and they can

be minimized and/or considered when interpreting studies' results. As an example to illustrate the influence of prognostic variables, Vissel (2004) examined the effects of three NBA interventions on participants' mental health and found differences between groups in post-intervention scores in variables like self-esteem and psychopathy. Because participants were not randomized to the interventions, there were large differences in mean scores at the baseline. Thus, differences in mean scores post-intervention might be explained by differences in the baseline instead of implying differences in the effect of the interventions provided. As an example of a within-group test of significance, Cave (1979) conducted a study to compare the effects of a high-stress condition, a low-stress condition, and a control group on the Minnesota Multiphasic Personality Inventory (MMPI) scores. Instead of conducting comparisons of the post-test scores while statistically controlling for pre-test scores, the author conducted several within-group tests of significance. Unfortunately, these tests are not the most informative regarding differences in group effects because they assess the change in score within a group (Bland & Altman, 2011; Vetter & Mascha, 2018).

To deal with methodological issues such as the ones just described, we provide some suggestions related to data analyses. These suggestions are simple and can be followed in any study. There are more complex designs that can be used to achieve relatively accurate estimates of effect without randomizing participants into groups (Antonakis et al., 2010; Marinescu et al., 2018). These designs include the regression discontinuity design, differences-in-differences, and instrumental variables, which are covered elsewhere (Antonakis et al., 2010; Marinescu et al., 2018).

Prognostic Variables

When randomization is not possible, researchers can endeavor to ensure that groups are similar in prognostic variables at baseline (see DeMille et al., 2018). For instance, if

researchers are concerned that women and men will benefit differently from the NBA intervention, the groups can be balanced by gender. Nonetheless, in non-randomized studies, the groups will probably differ in some prognostic variables (Sterne et al., 2016; von Elm et al., 2007). Consider, for example, that the outcome of interest is the reduction of depressive symptomatology and, at baseline, participants in the experimental group suffer from more severe depression than those in the comparison group. In this case, a greater reduction in depressive symptomatology in the experimental group may occur because more depressed participants tend to naturally improve more than less depressed participants (American Psychiatric Association, 2014). This problem may also be described as regression to the mean, which is 'a statistical phenomenon where initial measurements of a variable in a nonrandom sample at the extreme ends of a distribution tend to be closer to the mean upon a second measurement' (D. M. Thomas et al., 2020, p. 256). As an example, Wall (1993) found a reduction of 59% in the depressive symptoms of students engaging in an NBA intervention from baseline to post-intervention. Because these students started the intervention with a high mean depression score, this large improvement may be partially explained by the phenomenon of regression to the mean, though it was not evaluated by Wall (1993). Regression to the mean should be considered in analyses of samples or subsamples (e.g., subgroup analysis) of participants with extreme scores. Ensuring groups are similar at the baseline is a remedy for the impact of regression to the mean in the study's findings. If this design control is not possible and there is more than one group in the study design, researchers can use the baseline score as a statistical control when comparing the groups. This can be done by using an analysis of covariance or multiple regression (Bland & Altman, 2011; DeMille et al., 2018), which are 'the preferred statistical approach to accounting for baseline measurements of the outcome' (Higgins et al., 2019, p. 252). For studies without a

comparison group, researchers can, for example, assess whether the changes in the outcome variable from pre-test to post-test correlate with the baseline values (D. M. Thomas et al., 2020).

Equivocal use of Within-Group Significance Tests

Some non-randomized studies of NBA interventions base their conclusions on results from within-group (pre-post) significance tests (e.g., Cave, 1979; Ilagan et al., 2020). According to Bland and Altman (2011), this practice might lead to misinterpretation of the findings. For instance, when comparing two groups, researchers might conclude that the NBA intervention was superior to another type of intervention in terms of the benefits offered because the symptomatology of one group improved and that of the other group did not. This is checked by considering within-group significance tests. In this case, a researcher could have found a statistically significant improvement (commonly, $p < 0.05$) in the NBA group and a statistically non-significant result in the comparison group. This finding does not ensure that the two groups are statistically different at the end of the intervention (Bland & Altman, 2011). Neither it means that there was no change in the comparison group because a lack of statistical significance does not necessarily mean a lack of change in the sample (Bland & Altman, 2011; Cumming, 2014). In this case, a lack of statistical significance indicates that, with the confidence level adopted, the estimate found on the sample does not rule out a change of zero in the population (Cumming, 2014). Another issue to keep in mind when performing within-group tests of significance is that one group can start with a baseline score much lower or higher than the other group and, therefore, have much more room for improvement (Vickers, 2001). Thus, a larger improvement in a group might be due to participants' lower or higher baseline scores instead of due to the intervention (e.g., Ilagan et al., 2020; Magle-Haberek et al., 2012). To avoid inaccurate conclusions, it might be better to

analyze between-group differences while statistically controlling for the baseline score (Bland & Altman, 2011; Higgins et al., 2019; Vickers, 2001) as done by DeMille et al. (2018). In addition, we encourage researchers to interpret estimates of effect (i.e., standardized or non-standardized effect sizes) and their confidence intervals instead of focusing on the statistical significance of the results (see Amrhein et al., 2019). Estimates of effect help readers to judge the importance of the effect, and confidence intervals are informative regarding the precision of this effect in terms of inference to larger populations (Cumming, 2014).

Reporting of NBA Interventions: Some Necessary Elements

When assessing the effect of NBA interventions on people's health, there is essential information that needs to be reported to allow the comprehension and accurate evaluation of the findings (Montgomery et al., 2018; Seaman et al., 2020; Sterne et al., 2016, 2019). The Consolidated Standards of Reporting Trials (CONSORT) and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) provide guidelines about this (Moher et al., 2010; von Elm et al., 2007). Readers may also consult the American Psychological Association guidelines for quantitative (Appelbaum et al., 2018) and qualitative studies (Levitt et al., 2018) as well as the EQUATOR network (https://www.equator-network.org/), which compiles reporting guidelines according to the study's design.

Considering these guides, in this paper, we focus on four elements we believe are especially relevant for the interpretation and replicability of studies of NBA interventions. We describe these elements in the order they normally appear in an intervention study (Moher et al., 2010). The (a) first element is a detailed description of the interventions, (if any) co-interventions, and the setting where they take place. Co-interventions are other

treatments or preventive interventions received by study participants that are not part of the study (Berwanger et al., 2006). For example, a person engaged in an NBA intervention may also be receiving private counseling. If the private counseling is not reported, the readers cannot link change in health outcomes to this co-intervention and may believe that any change in outcome was due to the NBA intervention. Describing the elements involved in NBA interventions is especially important because these interventions often involve several nature-based activities (e.g., hiking, climbing, canoeing) taking place in different settings such as forests (Shin et al., 2012), mountains (Sturm et al., 2012), and sea (Gelkopf et al., 2013). Moreover, it is expected that different natural environments will have different effects on people's health, so knowing in detail the setting where the NBA intervention took place helps researchers to understand the possible impact of specific environments on people's health (Kaplan, 1995; Ulrich et al., 1991). Finally, detailed information of the (co) interventions and setting are necessary for possible replication efforts. The (b) second element that should be reported is the number of dropouts and reasons for dropping out. This information is especially relevant in studies of NBA interventions because the number and reasons for dropping out can vary widely across studies. For instance, whereas no participant dropped out from an NBA intervention involving a low ropes course and canoeing (Li et al., 2013), 12 of 45 (27%) participants dropped out from an intervention involving surfing (Pereira et al., 2020). Moreover, under certain conditions dropouts can bias effect estimates (Sterne et al., 2016, 2019). Adverse events are the (c) third element we believe is especially important to be reported in studies of NBA interventions. It is important to report these because serious adverse events may not be as rare in NBA interventions as in less risky interventions such as psychotherapy, walking, and slight changes in diet (Cave, 1979; Houge Mackenzie & Brymer, 2020; Moreton et al., 2021; Müller et al., 2011; Norton, 2008; Vissell,

2004; Wall, 1993). The (d) fourth element is a declaration of conflict of interest and research funding. There are a diversity of reasons for having a conflict of interest in research such as financial interest, professional promotion or recognition, beliefs, and speculation (Ioannidis, 2010). As an example of a financial conflict of interest in NBA research, running an NBA program may be a profitable business and some program owners may fund or work together with researchers to report the findings of their program (O'Mara-eves et al., 2016). Because the reported results may increase or decrease people's interest in the intervention, there is a financial conflict of interest in this kind of study (Higgins et al., 2019; Ioannidis, 2010). These four elements and their importance in studies of NBA interventions are explained in greater detail below.

Description of the (co) Interventions and the Settings where NBA Interventions Took Place

Many studies of NBA interventions report limited descriptive information about the interventions assessed (e.g., Meerts-Brandsma et al., 2020; Nurenberg, 1985; Rosenberg et al., 2014; Smith, 2010). There is a lack of detail about the place where the intervention took place and the activities provided during the intervention. Knowing the intervention in detail is essential to figure out which elements might have affected the outcomes of interest and to favor replicability (Moher et al., 2010; Montgomery et al., 2018; Tucker & Rheingold, 2010). For example, if a study of an NBA intervention does not report that participants also engaged in psychotherapy, we may erroneously think that the NBA intervention is the only reason for change in the outcome of interest. To avoid inaccurate interpretations of the findings, it is relevant to report whether co-interventions occurred and their characteristics. Similarly, when the setting where the NBA intervention took place is not described in detail and the activities conducted are vaguely explained, replicability is hindered (Moher et al., 2010; Montgomery

et al., 2018; Tucker & Rheingold, 2010). A detailed description of the intervention and setting may also be important for intervention or program fidelity/integrity, which refers to whether or not, and how well, the intervention or program is implemented as planned (Tucker & Rheingold, 2010). As Tucker and Rheingold explain (2010, p. 262) 'the more clearly defined a program or intervention is and the better the descriptions of specific aspects of it are, the greater the chance is that it will be implemented with fidelity.'. To favor the comprehension of the study's findings as well as its replication, researchers should provide a detailed description of the NBA intervention, the co-interventions (if any), including the setting where the intervention took place, and the activities provided to both, the experimental and the comparison group(s), if any.

The Number of Dropouts and Reasons

People's adherence to the intervention is relevant because those who drop out from the intervention may not get any benefit from it or might even have had negative experiences during the intervention (Sterne et al., 2016, 2019). For instance, in a study assessing the effect of a sailing intervention on quality of life, some participants dropped out because they did not like sailing (Gelkopf et al., 2013). Dropouts are common in interventions aimed at improving mental health because it generally takes some tries to find the most effective treatment(s) for each individual (Bernaras et al., 2019; Kappelmann et al., 2020; Lopresti, 2019). Despite this, many studies of NBA interventions do not report the number of participants who dropped out from the intervention (e.g., Bird, 2015; Eikenæs et al., 2006; Hyer et al., 1996; Jakubec et al., 2016; Shin et al., 2012). This practice might lead to a bias in the interpretation of the intervention's effect (see Sterne et al., 2016, 2019). Given the relevance of reporting the number of dropouts and the reasons for them, we encourage researchers to include this information in their studies, possibly through a flow diagram (see Moher et al., 2010). It is

also a good practice to consider the potential impact of the missing data due to dropouts on the studies' results or, in other words, to conduct attrition analyses (see Bell et al., 2013; Seaman et al., 2020 for further guidance).

Adverse Events or Effects of NBA Interventions

An adverse event is 'an unfavourable or harmful outcome that occurs during, or after, the use of a drug or other intervention, but is not necessarily caused by it' and an adverse effect is 'an adverse event for which the causal relation between the intervention and the event is at least a reasonable possibility' (Higgins et al., 2019, p. 494). NBA activities entail a series of risks that are rarer or even not present in non-adventurous activities. These risks may vary according to the activities (e.g., drowning while surfing or falling while skiing) and the risk management practices adopted by the intervention providers (Javorski & Gass, 2013). For example, the risk of having a concussion may be higher when practicing motocross than when sailing (Feletti & Bonato, 2020). Whereas many activities outdoors can be relatively safe (Javorski & Gass, 2013; Rosa et al., 2021), serious adverse events have been related to some NBA interventions (Cave, 1979; Dobud & Harper, 2018; Müller et al., 2011; Vissell, 2004). For instance, an intervention led by ski instructors to assess the effects of guided skiing on adults aged 60 to 76 in Austria reported that two participants had a leg fracture and three felt knee pain during the period of the intervention (Müller et al., 2011). Moreover, of the 21 participants who filled a questionnaire about falls, 17 reported at least one fall. Despite this, NBA intervention studies often do not report whether adverse events occurred or did not occur (e.g., Gelkopf et al., 2013; Li et al., 2013; Pereira et al., 2020; Shin et al., 2012; Sturm et al., 2012; Townsend et al., 2018; Walter et al., 2019). This hinders the assessment of the risk-benefit of the intervention and, hence, impairs the decision of promoting or disregarding such intervention. For example, knowing that serious adverse events can occur during a

guided skiing intervention may prevent a practitioner to use this kind of NBA activity to improve their clients' physical fitness. Alternatively, this practitioner may use an activity that is thought to be safer like backpacking. Thus, we encourage researchers to report the number and characteristics of adverse events occurred during their NBA interventions.

Conflict of Interest and Research Funding in NBA Interventions

Conflicts of interest can arise when individuals involved in research on the impact of a NBA intervention desire to obtain specific results (Higgins et al., 2019; Ioannidis, 2010). Such desire may be motivated by financial interest, job promotion or recognition, or even out of personal beliefs and speculation (Higgins et al., 2019; Ioannidis, 2010). The conflict of interest may be stronger in studies of NBA interventions when researchers have a financial interest in the intervention, such as when testing private-pay adventure programs they are affiliated with, than when working with non-profit organizations (DeMille et al., 2018; Higgins et al., 2019; Ioannidis, 2010). O'Mara-eves et al. (2016) found in their review of outdoor adventure programs that 13 of the 16 included studies were funded by private non-governmental organizations. Reporting conflict of interest and funding is important because these might influence the design, procedure, data analysis, and results reporting of studies assessing a health intervention (Higgins et al., 2019). There is evidence that studies tend to present more favorable results or interpretations of the interventions' effect when they have a financial interest in it (Higgins et al., 2019). The fact that a conflict of interest exists in a study does not make it unworthy. It just reflects the need to be aware of it when considering the study's results (Higgins et al., 2019; Ioannidis, 2010). Pre-registration may improve the trustworthiness of studies in which conflict of interest is unavoidable because it implies the registration of the study's participants, intervention, measures, and outcomes of interest before beginning the study (Moher et al., 2010; Page et al., 2021; Sterne et al., 2016, 2019). It

is also important that researchers report the study's results independent of whether the findings align or not with their interests. Though it is required by many journals that authors report their conflict of interest and funding, it is common to find both published and unpublished studies of NBA interventions that did not report whether or not the authors have a conflict of interest or received funding for their study (e.g., Gelkopf et al., 2013; Shin et al., 2012).Following existing reporting guidelines, we recommend reporting information about the study's conflict of interest and funding, even when there isn't any (e.g., 'The authors declare they have no conflict of interest').

Analytical Approach and Reporting of Studies of NBA Interventions Published in the Journal of Adventure Education and Outdoor Learning (2019-2021)

To better illustrate the above-mentioned suggestions to improve the analysis and reporting of studies of NBA interventions, we assessed all quantitative studies of NBA interventions published in the *Journal of Adventure Education and Outdoor Learning* (JAEOL) in the last three years (2019-2021). We chose JAEOL because it is one of the main outlets publishing research about NBA, so it would be more likely to find eligible studies in this journal as compared to others. The period was chosen to favor feasibility but still provide a screenshot of the most recent practice. The first author alone performed the selection process and extracted the data from included studies. The involvement of another researcher in these phases is recommended to reduce the probability of errors, nonetheless, it was not feasible for this study (Higgins et al., 2019). Fortunately, due to the small scale of this review and the simplicity of the data extraction, we do not believe any important error has occurred (Robson et al., 2019).

We were unable to expand our evaluation to other journals or to a longer period due to resource constraints (e.g., time and funding) but this does not affect our objective (i.e., to

illustrate the limitations and suggestions described in previous sections by using as examples studies published in a leading journal in the field). Nevertheless, this review should not be seen as representative of all published NBA studies or even of the studies published in JAEOL. Because we did not have an a priori hypothesis, we did not register a protocol for this review (Schultz et al., 2018).

The first step to conduct this review was to develop criteria to deem which studies are eligible for this evaluation and which are not. These criteria are detailed in Table 1. In the second step, the first author of this paper screened all studies for eligibility (see Figure 1). Of the 92 studies published in the period of interest (2019-2021), six were deemed eligible and evaluated (Chang et al., 2019; Hackett et al., 2021; Ilagan et al., 2020; Meerts-Brandsma et al., 2020; Mutz et al., 2019; G. J. Thomas, 2019). As an example of the application of the eligibility criteria, we excluded Kourtesopoulou and Kriemadis (2021) because the setting where the intervention took place was not described in enough detail to judge whether it was a natural setting.

<Please, insert Table 1 and Figure 1 about here>

In the third step, the first author extracted information regarding the main characteristics of the eligible studies (Table 2) and information relevant for the evaluation of the analytical approach and reporting of these studies (Table 3). This information was limited to the topics discussed in this paper.

<Please, insert Tables 2 and 3 about here>

The main characteristics reported in Table 2 highlight the diversity of participants and outcomes that are investigated in studies of NBA interventions. For example, participants groups included adult mentors, cadet women, year nine students, and German-speaking adolescents. Outcomes included confidence in skills, happiness, and perceived stress. The

sample size of included studies ranged from 12 to 261 participants. In line with the broader literature on wilderness therapy (Bowen & Neill, 2013), five studies adopted a one-group, pre-post design where there is no comparison group and participants are assessed before and after the intervention and one study used a non-randomized controlled design where participants are not randomly assigned to groups. No study included a follow-up assessment. Four studies occurred in the United States of America (USA), one in Australia, and one in France.

Regarding information relevant for the evaluation of the analytical approach (Table 3), two of the six studies conducted some statistical control for prognostic variables and all studies performed a within-group test of significance. For instance, assuming that the change in their outcome of interest may vary according to sex, Chang et al. (2019) analyzed whether their findings differed between males and females. Similarly, Mutz et al. (2019) assessed whether their findings differed between high-media consumers and low-to-moderate media consumers. However, this type of statistical control does not eliminate the possibility of confounding because the groups analyzed (e.g., male and female) may differ in prognostic variables. Ilagan et al.'s study (2020) is an example of how the statistical control for prognostic variables and the critical interpretation of the within-group test of significance could have improved the trustworthiness of the conclusions. In this study, the authors assessed the change in the happiness of cadet women who participated in a 3-day backpacking intervention (experimental group) and the change in the happiness of cadet women who did not participate in this intervention (comparison group). The authors found a moderate improvement in the experimental group after the intervention (d = -0.55) and virtually no change in the comparison group (d = -0.04). Based on these results, the researchers concluded that 'cadets who participate in the backpacking activity will have

higher OHQ [Oxford Happiness Questionnaire] scores than cadets who do not participate' (Ilagan et al., 2020, p. 7). Nevertheless, the scores of cadets in both groups were similar at the end of the intervention (experimental = 4.43; comparison = 4.35). This similarity in scores at the end of the intervention occurred because the experimental group started the intervention with baseline scores considerably lower than the comparison group (experimental = 3.99; comparison = 4.32). Thus, the experimental group had more room for improvement than the comparison group, being the baseline score a possible prognostic variable in this study. In this case, the authors could have assessed between-group differences in happiness after the intervention while statistically controlling for the baseline happiness of the participants (Vickers, 2001). Given the authors' design, this approach is the most appropriate to support their conclusion of between-group differences (Bland & Altman, 2011; Vickers, 2001).

Concerning the reporting of the six included studies (Table 3), no study described the (co) interventions and the setting they took place in enough detail to allow direct replication, half of these studies reported the number of participants who dropped out, no study reported information about adverse events, all studies presented a declaration of conflict statement, and two of the six studies did not report any information related to funding.

Whereas the information describing the (co) interventions and the settings where they took place was not enough for direct replication in any included study, the amount of information varied substantially across studies. For example, Hackett et al. (2021) reported the kind of activity performed in each location in their study (e.g., rock climbing and tree identification occurred in Riverside Park, Milwaukee). Despite providing considerable information regarding their intervention and setting, these authors did not describe the activities in enough detail to allow direct replication (e.g., the duration of each activity, the number of people in each session, and the responsible staff). On the other side, Meerts-

Brandsma et al. (2020) provided little information about the setting where their intervention took place.

Regarding dropout, the findings (Table 3) support our argument that dropout varies widely from study to study of NBA interventions and that the number of participants who dropped out is sometimes not reported. Moreover, the reasons why participants dropped out were not reported in any of the included studies. Even more concerning is the fact that no study reported information about adverse events. This lack of information may be because no adverse event occurred, but it may also be that the authors did not get information about adverse events or decided not to include this information (e.g., believed it was not relevant to report).

A positive finding was that all reviewed studies provided a conflict of interest statement, probably because it is a norm of the JAEOL. We also believe that reporting funding information is a norm of this journal but two studies did not report information about funding. It may have occurred because no statement appears in the study when the authors did not report funding on the journal's submission page. However, it is impossible to discern, when there is no information in the study about funding, if there was no funding for the study or if the authors did not report their funding. Thus, it may be a good practice to report information about funding even when there was no funding, for example, by writing that the study received no funding.

Conclusion

Many studies of NBA interventions argue that these interventions provide health benefits to participants (e.g., Gelkopf et al., 2013; Ilagan et al., 2020). These conclusions are supported by theories explaining the health benefits of contact with nature more broadly (Kaplan, 1995; Ulrich et al., 1991), and specific frameworks describing the health benefits of

NBA activities (Fernee et al., 2017; Houge Mackenzie et al., 2021; Russell & Farnum, 2004). Nonetheless, methodological limitations in the study design and data analysis and a lack of information when reporting the results sometimes make the positive effects found in favor of NBA interventions questionable (Sterne et al., 2016, 2019). In this paper, we discussed limitations that may occur in the design and data analysis of studies of NBA interventions, argue the need to report some information, and illustrated our suggestions by using as examples studies published in JAEOL from 2019 to 2021. This analysis of JAEOL allowed us to exemplify our suggestions with recent studies published in a leading journal dedicated to adventure research. We focused on quantitative studies of NBA interventions but many of our suggestions may also be useful for quantitative researchers from other fields. For example, any non-randomized controlled study should consider the impact of prognostic variables and base arguments regarding group differences in between-group analyses rather than within-group significance tests (Bland & Altman, 2011; Sterne et al., 2016).

Our study is not without limitations, the main ones being the use of a broad definition of NBA, and the fact that we did not discuss how analysis and reporting vary across studies depending on the kind of adventure intervention provided to participants (e.g., involving mental health professionals versus not involving these professionals). Our decision to use a broad definition of NBA was based on our aim to highlight methodological challenges appearing across studies in the field and to provide suggestions that might be useful for a wide range of studies. Whereas it was not the focus of our study to describe how dropout differs according to interventions' characteristics, nor to discuss in detail the different reasons behind the conflict of interest, we encourage future research on these topics. Our aim in focusing on the challenges presented in the paper is not to criticize previous work in the field that has achieved a growing recognition of the role of NBA interventions in the promotion of

people's health. Rather, we hope that outlining these challenges and research guidelines will encourage future studies to control for prognostic variables, avoid within-study tests of significance when having a comparison group, and report more detailed information about the intervention, setting, dropouts, adverse events, conflict of interest and research funding. In the end, improving the analysis and reporting of studies of NBA interventions will make these studies more trustworthy, favor replicability, and better inform the decision of using or not NBA interventions to improve people's health.

Declaration Statement

The authors declare they have no conflict of interest.

References

American Psychiatric Association (Ed.). (2014). *Manual diagnóstico e estatístico de transtornos mentais: DSM-5 [Diagnostic and Statistical Manual of Mental Disorders: DSM-5]* (5th ed.). Artmed.

American Psychological Association. (2020). *Publication Manual of the American Psychological Association* (7th ed.).

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086–1120. https://doi.org/10.1016/j.leaqua.2010.10.010

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A. M., & Rao, S. M. (2018). Reporting Standards for Quantitative Research in Psychology: The APA Publications and Communications Board Task Force Report. *American Psychologist*, *1*(2), 26–46. https://doi.org/10.1037/amp0000151

Astell-Burt, T., Mitchell, R., & Hartig, T. (2014). The association between green space and mental health varies across the lifecourse. A longitudinal study. *Journal of Epidemiology and Community Health*, *68*(6), 578–583. https://doi.org/10.1136/jech-2013-203767

Bell, M. L., Kenward, M. G., Fairclough, D. L., & Horton, N. J. (2013). Differential dropout and bias in randomised controlled trials: When it matters and when it may not. *BMJ*, *346*, e8668–e8668. https://doi.org/10.1136/bmj.e8668

Benninger, E., Curtis, C., Sarkisian, G. V., Rogers, C. M., Bender, K., & Comer, M. (2020). Surf therapy: A scoping review of the qualitative and quantitative research evidence.

Global Journal of Community Psychology Practice, *11*(2), 1–26.

Bernaras, E., Jaureguizar, J., & Garaigordobil, M. (2019). Child and adolescent depression: A review of theories, evaluation instruments, prevention programs, and treatments. *Frontiers in Psychology*, *10*(543). https://doi.org/10.3389/fpsyg.2019.00543

Berwanger, O., Guimarães, H. P., Avezum, Á., & Piegas, L. S. (2006). Os dez mandamentos do ensaio clínico randomizado – princípios para avaliação crítica da literatura médica [The ten commandments of the randomized clinical trial – principles for critical appraisal of the medical literature]. *Revista Brasileira de Hipertensão*, *13*(1), 65–70.

Bettmann, J. E., Gillis, H. L., Speelman, E. A., Parry, K. J., & Case, J. M. (2016). A meta-analysis of wilderness therapy outcomes for private pay clients. *Journal of Child and Family Studies*, *25*(9), 2659–2673. https://doi.org/10.1007/s10826-016-0439-0

Bird, K. (2015). Research Evaluation of an Australian Peer Outdoor Support Therapy Program for Contemporary Veterans' Wellbeing. *International Journal of Mental Health*, *44*(1–2), 46–79. https://doi.org/10.1080/00207411.2015.1009752

Bland, J. M., & Altman, D. G. (2011). Comparisons against baseline within randomised groups are often used and can be highly misleading. *Trials*, *12*(1), 264. https://doi.org/10.1186/1745-6215-12-264

Bowen, D. J., & Neill, J. T. (2013). A meta-analysis of adventure therapy outcomes and moderators. *The Open Psychology Journal*, *6*(1), 28–53. https://doi.org/10.2174/1874350120130802001

Bowen, D. J., Neill, J. T., & Crisp, S. J. R. (2016). Wilderness adventure therapy effects on the mental health of youth participants. *Evaluation and Program Planning*, *48*, 49–59.

Cave, S. E. R. (1979). *Evaluation of level of stress and group cohesiveness in the wilderness experience using the mmpi and sociograms* [Doctoral dissertaion, The University of

New Mexico]. https://www.proquest.com/openview/4726634b304f1429a4597dcba8a16eb3/1?pq-origsite=gscholar&cbl=18750&diss=y

Chang, Y., Davidson, C., Conklin, S., & Ewert, A. (2019). The impact of short-term adventure-based outdoor programs on college students' stress reduction. *Journal of Adventure Education and Outdoor Learning*, *19*(1), 67–83. https://doi.org/10.1080/14729679.2018.1507831

Corazon, S. S., Sidenius, U., Poulsen, D. V., Gramkow, M. C., & Stigsdotter, U. K. (2019). Psycho-physiological stress recovery in outdoor nature-based interventions: A systematic review of the past eight years of research. *International Journal of Environmental Research and Public Health*, *16*(10), 1711. https://doi.org/10.3390/ijerph16101711

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

DeMille, S., Tucker, A. R., Gass, M. A., Javorski, S., VanKanegan, C., Talbot, B., & Karoff, M. (2018). The effectiveness of outdoor behavioral healthcare with struggling adolescents: A comparison group study a contribution for the special issue: Social innovation in child and youth services. *Children and Youth Services Review*, *88*, 241–248. https://doi.org/10.1016/j.childyouth.2018.03.015

Dobud, W. W., & Harper, N. J. (2018). Of Dodo birds and common factors: A scoping review of direct comparison trials in adventure therapy. *Complementary Therapies in Clinical Practice*, *31*, 16–24. https://doi.org/10.1016/j.ctcp.2018.01.005

Eikenæs, I., Gude, T., & Hoffart, A. (2006). Integrated wilderness therapy for avoidant personality disorder. *Nordic Journal of Psychiatry*, *60*(4), 275–281.

https://doi.org/10.1080/08039480600790093

Feletti, F., & Bonato, M. (2020). The incidence of pediatric and adolescent concussion in action sports: A systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, *17*(23), 8728. https://doi.org/10.3390/ijerph17238728

Fernee, C. R., Gabrielsen, L. E., Andersen, A. J. W., & Mesel, T. (2017). Unpacking the black box of wilderness therapy. *Qualitative Health Research*, *27*(1), 114–129. https://doi.org/10.1177/1049732316655776

Fernee, C. R., Gabrielsen, L. E., Andersen, A. J. W., & Mesel, T. (2021). Emerging stories of self: Long-term outcomes of wilderness therapy in Norway. *Journal of Adventure Education and Outdoor Learning*, *21*(1), 67–81. https://doi.org/10.1080/14729679.2020.1730205

Gabrielsen, L. E., Eskedal, L. T., Mesel, T., Aasen, G. O., Hirte, M., Kerlefsen, R. E., Palucha, V., & Fernee, C. R. (2019). The effectiveness of wilderness therapy as mental health treatment for adolescents in Norway: A mixed methods evaluation. *International Journal of Adolescence and Youth*, *24*(3), 282–296. https://doi.org/10.1080/02673843.2018.1528166

Gabrielsen, L. E., Fernee, C. R., Aasen, G. O., & Eskedal, L. T. (2016). Why randomized trials are challenging within adventure therapy research: Lessons learned in Norway. *Journal of Experiential Education*, *39*(1), 5–14. https://doi.org/10.1177/1053825915607535

Gelkopf, M., Hasson-Ohayon, I., Bikman, M., & Kravetz, S. (2013). Nature adventure rehabilitation for combat-related posttraumatic chronic stress disorder: A randomized control trial. *Psychiatry Research*, *209*(3), 485–493.

https://doi.org/10.1016/j.psychres.2013.01.026

Gillis, H. L., & Simpson, C. (1991). Project Choices: Adventure-Based Residential Drug Treatment for Court-Referred Youth. *Journal of Addictions & Offender Counseling*, *12*(1), 12–27. https://doi.org/10.1002/j.2161-1874.1991.tb00077.x

Gladwell, V. F., Brown, D. K., Wood, C., Sandercock, G. R., & Barton, J. L. (2013). The great outdoors: How a green exercise environment can benefit all. *Extreme Physiology and Medicine*, *2*, 1–7. https://doi.org/10.1186/2046-7648-2-3

Greenhalgh, T., Thorne, S., & Malterud, K. (2018). Time to challenge the spurious hierarchy of systematic over narrative reviews? *European Journal of Clinical Investigation*, *48*(6), e12931. https://doi.org/10.1111/eci.12931

Greer, M., & Vin-Raviv, N. (2019). Outdoor-based therapeutic recreation programs among military veterans with posttraumatic stress disorder: Assessing the evidence. *Military Behavioral Health*, *7*(3), 286–303. https://doi.org/10.1080/21635781.2018.1543063

Hackett, K. A., Ziegler, M. C., Olson, J. A., Bizub, J., Stolley, M., Szabo, A., Heller, E., & Beyer, K. M. M. (2021). Nature Mentors: A program to encourage outdoor activity and nature engagement among urban youth and families. *Journal of Adventure Education and Outdoor Learning*, *21*(1), 35–52. https://doi.org/10.1080/14729679.2020.1730203

Harper, N. J., Fernee, C. R., & Gabrielsen, L. E. (2021). Nature's Role in Outdoor Therapies: An Umbrella Review. *International Journal of Environmental Research and Public Health*, *18*(10), 5117. https://doi.org/10.3390/ijerph18105117

Hartig, T. (2021). Restoration in nature: Beyond the conventional narrative. In A. R. Schutte, J. Torquati, & J. R. Stevens (Eds.), *Nature and psychology: Biological, cognitive, developmental, and social pathways to well-being*. Springer.

Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. CRC Press.

Higgins, J. P. T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2019). *Cochrane Handbook for Systematic Reviews of Interventions* (2nd ed.). John Wiley & Sons.

Houge Mackenzie, S., & Brymer, E. (2020). Conceptualizing adventurous nature sport: A positive psychology perspective. *Annals of Leisure Research*, *23*(1), 79–91. https://doi.org/10.1080/11745398.2018.1483733

Houge Mackenzie, S., & Goodnow, J. (2021). Adventure in the age of COVID-19: Embracing microadventures and locavism in a post-pandemic world. *Leisure Sciences*, *43*(1–2), 62–69. https://doi.org/10.1080/01490400.2020.1773984

Houge Mackenzie, S., Hodge, K., & Filep, S. (2021). How does adventure sport tourism enhance well-being? A conceptual model. *Tourism Recreation Research*, 1–14. https://doi.org/10.1080/02508281.2021.1894043

Hyer, L., Boyd, S., Scurfield, R., Smith, D., & Burke, J. (1996). Effects of Outward Bound Experience as an adjunct to inpatient PTSD treatment of war veterans. *Journal of Clinical Psychology*, *52*(3), 263–278. https://doi.org/10.1002/(SICI)1097-4679(199605)52:3<263::AID-JCLP3>3.0.CO;2-T

Ilagan, G., Ilagan, J., Jocius, R., Jefferson, R., Bennett-Mintz, J., McCormick, K., & Farrell, M. (2020). Happiness outcomes among cadet women backpackers. *Journal of Adventure Education and Outdoor Learning*, *20*(4), 285–297. https://doi.org/10.1080/14729679.2019.1660194

Ioannidis, J. P. A. (2010). Meta-research: The art of getting it wrong. *Research Synthesis Methods*, *1*(3–4), 169–184. https://doi.org/10.1002/jrsm.19

Jakubec, S. L., Carruthers Den Hoed, D., Ray, H., & Krishnamurthy, A. (2016). Mental well-being and quality-of-life benefits of inclusion in nature for adults with disabilities and

751 their caregivers. *Landscape Research*, *41*(6), 616–627.
752 https://doi.org/10.1080/01426397.2016.1197190
753 Javorski, S. E., & Gass, M. A. (2013). 10-Year Incident Monitoring Trends in Outdoor
754 Behavioral Healthcare: Lessons learned and future directions. *Journal of Therapeutic*
755 *Schools & Programs*, *VI*(I), 112–128.
756 Kaplan, S. (1995). The restorative benefits of nature: Toward an integrative framework.
757 *Journal of Environmental Psychology*, *15*(3), 169–182. https://doi.org/10.1016/0272-
758 4944(95)90001-2
759 Kappelmann, N., Rein, M., Fietz, J., Mayberg, H. S., Craighead, W. E., Dunlop, B. W.,
760 Nemeroff, C. B., Keller, M., Klein, D. N., Arnow, B. A., Husain, N., Jarrett, R. B.,
761 Vittengl, J. R., Menchetti, M., Parker, G., Barber, J. P., Bastos, A. G., Dekker, J., Peen,
762 J., … Kopf-Beck, J. (2020). Psychotherapy or medication for depression? Using
763 individual symptom meta-analyses to derive a Symptom-Oriented Therapy (SOrT)
764 metric for a personalised psychiatry. *BMC Medicine*, *18*(170).
765 https://doi.org/10.1186/s12916-020-01623-9
766 Kelley, M. P., Coursey, R. D., & Selby, P. M. (1997). Therapeutic adventures outdoors: a
767 demonstration of benefits for people with mental illness. *Psychiatric Rehabilitation*
768 *Journal*, *20*(4), 61–73.
769 Kourtesopoulou, A., & Kriemadis, A. (2021). Exploring the influence of Outdoor
770 Management Development (OMD) program on leadership and teamwork competencies.
771 *Journal of Adventure Education and Outdoor Learning*, *21*(3), 247–260.
772 https://doi.org/10.1080/14729679.2020.1784763
773 Lackey, N. Q., Tysor, D. A., McNay, G. D., Joyner, L., Baker, K. H., & Hodge, C. (2019).
774 Mental health benefits of nature-based recreation: A systematic review. *Annals of*

Leisure Research, 1–15. https://doi.org/10.1080/11745398.2019.1655459

Levitt, H. M., Bamberg, M., Creswell, J. W., Frost, D. M., & Suárez-orozco, C. (2018). Reporting standards for qualitative primary, qualitative meta-analytic, and mixed methods research in Psychology : The APA Publications and Communications Board Task Force Report. *American Psychologist*, *73*(1), 26–46. https://doi.org/http://dx.doi.org/10.1037/amp0000151

Li, W. H. C., Chung, J. O. K., & Ho, E. K. Y. (2013). Effectiveness of an adventure-based training programme in promoting the psychological well-being of primary schoolchildren. *Journal of Health Psychology*, *18*(11), 1478–1492. https://doi.org/10.1177/1359105312465102

Lopresti, A. L. (2019). It is time to investigate integrative approaches to enhance treatment outcomes for depression? *Medical Hypotheses*, *126*, 82–94. https://doi.org/10.1016/j.mehy.2019.03.008

Magle-Haberek, N. A., Tucker, A. R., & Gass, M. A. (2012). Effects of program differences with wilderness therapy and residential treatment center (RTC) programs. *Residential Treatment For Children & Youth*, *29*(3), 202–218. https://doi.org/10.1080/0886571X.2012.697433

Marinescu, I. E., Lawlor, P. N., & Kording, K. P. (2018). Quasi-experimental causality in neuroscience and behavioural research. *Nature Human Behaviour*, *2*(12), 891–898. https://doi.org/10.1038/s41562-018-0466-5

Meerts-Brandsma, L., Sibthorp, J., & Rochelle, S. (2020). Using transformative learning theory to understand outdoor adventure education. *Journal of Adventure Education and Outdoor Learning*, *20*(4), 381–394. https://doi.org/10.1080/14729679.2019.1686040

Moher, D., Hopewell, S., Schulz, K. F., Montori, V., Gøtzsche, P. C., Devereaux, P. J.,

Elbourne, D., Egger, M., & Altman, D. G. (2010). CONSORT 2010 explanation and elaboration: Updated guidelines for reporting parallel group randomised trials. *International Journal of Surgery*, *10*(1), 28–55. https://doi.org/https://doi.org/10.1136/bmj.c869

Montgomery, P., Grant, S., Mayo-Wilson, E., Macdonald, G., Michie, S., Hopewell, S., & Moher, D. (2018). Reporting randomised trials of social and psychological interventions: the CONSORT-SPI 2018 Extension. *Trials*, *19*(407). https://doi.org/10.1186/s13063-018-2733-1

Moreton, S. G., Brennan, M. K., Nicholls, V. I., Wolf, I. D., & Muir, D. L. (2021). Exploring potential mechanisms underpinning the therapeutic effects of surfing. *Journal of Adventure Education and Outdoor Learning*, 1–18. https://doi.org/10.1080/14729679.2021.1884104

Müller, E., Gimpl, M., Poetzelsberger, B., Finkenzeller, T., & Scheiber, P. (2011). Salzburg Skiing for the Elderly Study: Study design and intervention - health benefit of alpine skiing for elderly. *Scandinavian Journal of Medicine & Science in Sports*, *21*, 1–8. https://doi.org/10.1111/j.1600-0838.2011.01336.x

Mutz, M., Müller, J., & Göring, A. (2019). Outdoor adventures and adolescents' mental health: daily screen time as a moderator of changes. *Journal of Adventure Education and Outdoor Learning*, *19*(1), 56–66. https://doi.org/10.1080/14729679.2018.1507830

Norton, C. L. (2008). Understanding the Impact of Wilderness Therapy on Adolescent Depression and Psychosocial Development. *Illinois Child Welfare*, *4*(1), 166–178.

Nurenberg, S. J. G. (1985). *Psychological development of borderline adolescents in wilderness therapy*. Smith College School for Social Work.

O'Mara-eves, A., Fiennes, C., & Oliver, S. (2016). *The relationship between short-term*

outcomes and long-term outcomes produced by outdoor adventure programmes: A rapid review. UCL Intstitute of Education and Giving Evidence.

Ohly, H., White, M. P., Wheeler, B. W., Bethel, A., Ukoumunne, O. C., Nikolaou, V., & Garside, R. (2016). Attention Restoration Theory: A systematic review of the attention restoration potential of exposure to natural environments. *Journal of Toxicology and Environmental Health, Part B*, *19*(7), 305–343. https://doi.org/10.1080/10937404.2016.1196155

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716

Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *BMJ*, *160*. https://doi.org/10.1136/bmj.n160

Pereira, A. I., Ferreira, C., Oliveira, M., Evangelista, E. S., Ferreira, J., Roberto, M. S., Tereso, S., Pereira, A. M., Neves, S., & Crespo, C. (2020). Effectiveness of a combined surf and psychological preventive intervention with children and adolescents in residential childcare: A randomized controlled trial. *Revista de Psicología Clínica Con Niños y Adolescentes*, *7*(2), 22–31. https://doi.org/10.21134/rpcna.2020.07.2.3

Roberts, A., Hinds, J., & Camic, P. M. (2020). Nature activities and wellbeing in children and young people: A systematic literature review. *Journal of Adventure Education and Outdoor Learning*, *20*(4), 298–318. https://doi.org/10.1080/14729679.2019.1660195

Robson, R. C., Pham, B., Hwee, J., Thomas, S. M., Rios, P., Page, M. J., & Tricco, A. C.

(2019). Few studies exist examining methods for selecting studies, abstracting data, and appraising quality in a systematic review. *Journal of Clinical Epidemiology*, *106*, 121–135. https://doi.org/10.1016/j.jclinepi.2018.10.003

Rogers, C. M., Mallinson, T., & Peppers, D. (2014). High-Intensity Sports for Posttraumatic Stress Disorder and Depression: Feasibility Study of Ocean Therapy With Veterans of Operation Enduring Freedom and Operation Iraqi Freedom. *American Journal of Occupational Therapy*, *68*(4), 395–404. https://doi.org/10.5014/ajot.2014.011221

Rosa, C. D., & Delabrida, Z. (2021). Método experimental e ensaios clínicos: Metassíntese de artigos de revisão publicados em português [Experimental method and clinical trials: Metassynthesis of review articles published in Portuguese]. *Psico*, *52*(4), e36259. https://doi.org/10.15448/1980-8623.2021.4.36259

Rosa, C. D., Larson, L. R., Collado, S., & Profice, C. C. (2021). Forest therapy can prevent and treat depression: Evidence from meta-analyses. *Urban Forestry & Urban Greening*, *57*, 126943. https://doi.org/10.1016/j.ufug.2020.126943

Rosenberg, R. S., Lange, W., Zebrack, B., Moulton, S., & Kosslyn, S. M. (2014). An Outdoor Adventure Program for Young Adults with Cancer: Positive Effects on Body Image and Psychosocial Functioning. *Journal of Psychosocial Oncology*, *32*(5), 622–636. https://doi.org/10.1080/07347332.2014.936652

Russell, K. C., & Farnum, J. (2004). A concurrent model of the wilderness therapy process. *Journal of Adventure Education & Outdoor Learning*, *4*(1), 39–55. https://doi.org/10.1080/14729670485200411

Schultz, A., Goertzen, L., Rothney, J., Wener, P., Enns, J., Halas, G., & Katz, A. (2018). A scoping approach to systematically review published reviews: Adaptations and recommendations. *Research Synthesis Methods*, *9*(1), 116–123.

https://doi.org/10.1002/jrsm.1272

Seaman, J., Dettweiler, U., Humberstone, B., Martin, B., Prince, H., & Quay, J. (2020). Joint recommendations on reporting empirical research in outdoor, experiential, environmental, and adventure education journals. *Journal of Experiential Education*, *43*(4), 348–364. https://doi.org/10.1177/1053825920969443

Shamseer, L., Moher, D., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L. A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. *BMJ*, *349*, g7647–g7647. https://doi.org/10.1136/bmj.g7647

Shin, W. S., Shin, C. S., & Yeoun, P. S. (2012). The influence of forest therapy camp on depression in alcoholics. *Environmental Health and Preventive Medicine*, *17*(1), 73–76. https://doi.org/10.1007/s12199-011-0215-0

Smith, N. T. (2010). *Effectiveness of a guided wilderness experience on improved client self-esteem, self-efficacy, anxiety, and depression* [Master's thesis, University of Utah]. https://www.proquest.com/openview/142424a1ea5ceb27d24d9ba3d37c3af5/1?pq-origsite=gscholar&cbl=18750

Soga, M., & Gaston, K. J. (2022). The dark side of nature experience: Typology, dynamics and implications of negative sensory interactions with nature. *People and Nature*, *4*(5), 1126–1140. https://doi.org/10.1002/pan3.10383

Sterne, J. A., Hernán, M. A., Reeves, B. C., Savović, J., Berkman, N. D., Viswanathan, M., Henry, D., Altman, D. G., Ansari, M. T., Boutron, I., Carpenter, J. R., Chan, A.-W., Churchill, R., Deeks, J. J., Hróbjartsson, A., Kirkham, J., Jüni, P., Loke, Y. K., Pigott, T. D., … Higgins, J. P. (2016). ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*, *i4919*. https://doi.org/10.1136/bmj.i4919

Sterne, J. A., Savović, J., Page, M. J., Elbers, R. G., Blencowe, N. S., Boutron, I., Cates, C. J., Cheng, H.-Y., Corbett, M. S., Eldridge, S. M., Emberson, J. R., Hernán, M. A., Hopewell, S., Hróbjartsson, A., Junqueira, D. R., Jüni, P., Kirkham, J. J., Lasserson, T., Li, T., … Higgins, J. P. T. (2019). RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*, l4898. https://doi.org/10.1136/bmj.l4898

Sturm, J., Plöderl, M., Fartacek, C., Kralovec, K., Neunhäuserer, D., Niederseer, D., Hitzl, W., Niebauer, J., Schiepek, G., & Fartacek, R. (2012). Physical exercise through mountain hiking in high-risk suicide patients. A randomized crossover trial. *Acta Psychiatrica Scandinavica*, *126*(6), 467–475. https://doi.org/10.1111/j.1600-0447.2012.01860.x

Thomas, D. M., Clark, N., Turner, D., Siu, C., Halliday, T. M., Hannon, B. A., Kahathuduwa, C. N., Kroeger, C. M., Zoh, R., & Allison, D. B. (2020). Best (but oft-forgotten) practices: Identifying and accounting for regression to the mean in nutrition and obesity research. *The American Journal of Clinical Nutrition*, *111*(2), 256–265. https://doi.org/10.1093/ajcn/nqz196

Thomas, G. J. (2019). Effective teaching and learning strategies in outdoor education: Findings from two residential programmes based in Australia. *Journal of Adventure Education and Outdoor Learning*, *19*(3), 242–255. https://doi.org/10.1080/14729679.2018.1519450

Townsend, J., Hawkins, B. L., Bennett, J. L., Hoffman, J., Martin, T., Sotherden, E., & Bridges, W. (2018). Preliminary long-term health outcomes associated with recreation-based health and wellness programs for injured service members. *Cogent Psychology*, *5*. https://doi.org/10.1080/23311908.2018.1444330

Trundle, G., & Hutchinson, R. (2021). The phased model of adventure therapy: Trauma-

focussed, low arousal, & positive behavioural support. *Journal of Adventure Education and Outdoor Learning*, *21*(1), 82–92. https://doi.org/10.1080/14729679.2020.1736109

Tucker, A. R., & Rheingold, A. (2010). Enhancing fidelity in adventure education and adventure therapy. *Journal of Experiential Education*, *33*(3), 258–273. https://doi.org/10.5193/JEE33.3.258

Ulrich, R. S., Simons, R. F., Losito, B. D., Fiorito, E., Miles, M. A., & Zelson, M. (1991). Stress recovery during exposure to natural and urban environments. *Journal of Environmental Psychology*, *11*(3), 201–230. https://doi.org/10.1016/S0272-4944(05)80184-7

Vetter, T. R., & Mascha, E. J. (2018). Unadjusted bivariate two-group comparisons. *Anesthesia & Analgesia*, *126*(1), 338–342. https://doi.org/10.1213/ANE.0000000000002636

Vickers, A. J. (2001). The use of percentage change from baseline as an outcome in a controlled trial is statistically inefficient: A simulation study. *BMC Medical Research Methodology*, *1*(1), 6. https://doi.org/10.1186/1471-2288-1-6

Vissell, R. (2004). *Effects of wilderness therapy on youth at risk's concept of self and other: A deeper understanding of the journey* [Doctoral dissertation, Institute of Transpersonal Psychology]. https://www.proquest.com/openview/4687d4a75bd44944ee37c08417885e4a/1?pq-origsite=gscholar&cbl=18750&diss=y

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2007). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: Guidelines for Reporting Observational Studies. *Annals of Internal Medicine*, *147*(8), 573. https://doi.org/10.7326/0003-4819-147-8-200710160-

00010

Walker, A. J., Onus, M., Doyle, M., Clare, J., & McCarthy, K. (2005). Cognitive rehabilitation after severe traumatic brain injury: A pilot programme of goal planning and outdoor adventure course participation. *Brain Injury*, *19*(14), 1237–1241. https://doi.org/10.1080/02699050500309411

Wall, S. M. (1993). *The effects of a wilderness survival school on adolescent depression scores* [Doctoral dissertation, Pacific University]. https://commons.pacificu.edu/work/sc/73aece02-3353-46ea-893b-b849410f9987

Walter, K. H., Otis, N. P., Ray, T. N., Glassman, L. H., Michalewicz-Kragh, B., Powell, A. L., & Thomsen, C. J. (2019). Breaking the surface: Psychological outcomes among U.S. active duty service members following a surf therapy program. *Psychology of Sport & Exercise*, *45*, N.PAG-N.PAG. https://doi.org/10.1016/j.psychsport.2019.101551

White, M. P., Alcock, I., Grellier, J., Wheeler, B. W., Hartig, T., Warber, S. L., Bone, A., Depledge, M. H., & Fleming, L. E. (2019). Spending at least 120 minutes a week in nature is associated with good health and wellbeing. *Scientific Reports*, *9*(7730). https://doi.org/10.1038/s41598-019-44097-3

Wohlwill, J. F. (1983). The concept of nature: A psychologist's view. In I. Altman & J. F. Wohlwill (Eds.), *Behavior and the natural environment* (pp. 5–37). Plenum Press. https://doi.org/10. I007/978-1-4613-3539-9

963 **Table 1**

964 *Eligibility criteria for inclusion in our analysis of studies of NBA interventions published in*

965 *JAEOL (2019-2021) based on population (P), intervention (I), comparison groups of interest*

966 *(C), outcomes (O), and study designs (S)*

PICOS	Description
Population	Studies with humans at any age, healthy or unhealthy
Intervention	Studies that provide or consider any form of NBA intervention. We define NBA intervention as the promotion of one or more activities of adventure in a natural setting. We considered a natural setting as an environment that has a significant presence of soil (e.g., mountains), vegetation (e.g., trees), and/or water, and that was not highly altered by humans
Comparison groups of interest	Studies with any comparison/control group and studies without a control group
Outcomes	Studies assessing any quantitative outcome
Study design	Quantitative randomized and non-randomized studies of interventions

967 *Note.* JAEOL = Journal of Adventure Education and Outdoor Learning; NBA = Nature-based

968 adventure

969

970 **Table 2**

971 *Main characteristics of the studies included in our analysis of studies of NBA interventions*

972 *published in the JAEOL (2019-2021)*

First author (year)	Participants	N	Mean age or age range	Women %	Time (T) in which data was collected	Outcome measures	Study design	Country
Hackett (2021)	Adult mentor and child mentee pairs and families	44	8 to 60	88.46	T1: Within two weeks before the program starting T2: One to four weeks after the program	Outdoor Skills Confidence Scale	One-group, pre-post	United States of America
Ilagan (2020)	Cadet women backpackers	12	18 to 21	100.00	T1: Two nights before the trip T2: Last night of the trip	Oxford Happiness Questionnaire	Non-randomized controlled trial	United States of America
Meerts-Brandsma (2020)	National Outdoor Leadership School (NOLS) students	139	23	57.78	T1: Before the arrival at NOLS T2: After the course	Learning Activities Survey and The University of Rhode Island Change Assessment scale (URICA)	One-group, pre-post	United States of America
Thomas (2019)	Year nine students	261	NI	NI	T1: At the school before the program T2: Near the end of the program	The Life Effectiveness Questionnaire	One-group, pre-post	Australia
Mutz (2019)	German-speaking adolescents	108	17.8	38.84	T1: Before the activities T2: In the evening of the last day	Perceived Stress Questionnaire; a life satisfaction question; and an adapted measure of hedonic balance	One-group, pre-post	France
Chang (2019)	College students	33	20.67	66.66	T1: Meeting in the classroom T2: Before the field trip T3: At the conclusion of the field trip	Salivary cortisol and the Perceived Stress Questionnaire	One-group, pre-post	United States of America

973 *Note.* JAEOL = Journal of Adventure Education and Outdoor Learning; NBA = Nature-based

974 adventure; NI = No information

975

Table 3

Description of the studies of NBA interventions published in JAEOL (2019-2021) regarding the analytical approach and the reporting elements discussed in this paper

First author (year)	Statistical control for any prognostic variable	Within-group test of significance	Description of the (co) interventions and setting where the (co) interventions took place	Number of dropouts	Adverse events	Declaration of conflict of interest	Funding source declaration
Hackett (2021)	No	Yes	Insufficient detail for replication	2/28	NI	Declared	Reported
Ilagan (2020)	No	Yes	Insufficient detail for replication	NI	NI	Declared	Reported
Meerts-Brandsma (2020)	No	Yes	Insufficient detail for replication	NI	NI	Declared	NI
Thomas (2019)	No	Yes	Insufficient detail for replication	NI	NI	Declared	Independentstudy
Mutz (2019)	Yes	Yes	Insufficient detail for replication	0/108	NI	Declared	NI
Chang (2019)	Yes	Yes	Insufficient detail for replication	10/46	NI	Declared	Reported

Note. JAEOL = Journal of Adventure Education and Outdoor Learning; NBA = Nature-based adventure; NI = No information.

Figure 1

Flow diagram illustrating the process of selection of studies eligible for the analysis