Reviewing and Benchmarking Adventure Therapy Outcomes: Applications of Meta-Analysis

James T. Neill

This article summarizes evidence from meta-analyses of outdoor education, psychotherapy, and education which can be used to help determine the relative efficacy of adventure therapy programs. The major outdoor education meta-analysis by Hattie, Marsh, Neill, and Richards (1997) cites an effect size of .34 for immediate program effectiveness. Meta-analytic studies of innovative education programs in school settings report similar effects (Hattie et al., Hattie, 1992, 1993), while psychological training meta-analyses report slightly higher effects (Lipsey & Wilson, 1993), and meta-analyses of psychotherapy effectiveness find stronger effects (Casey & Berman, 1985; Smith, Glass, & Miller, 1980). This article derives outcome benchmarks for adventure therapy program outcomes taken from existing outdoor education, psychological, and educational meta-analyses, and recommends that the benchmarks be used in adventure therapy research and evaluation to allow for more accurate assessments of program's effectiveness. The suggested benchmarks should be used only as guides, and should be refined in future by meta-analyses of adventure therapy outcomes.

Keywords: Adventure Therapy, Research and Evaluation, Meta-analysis, Benchmarks

In modern Western psychology, therapeutic interventions for dealing with psychological problems have, approximately, a 100 year history. Today, mental health professionals find themselves dealing with an increasing range and depth of psychological problems. Fortunately, mental health professionals have at their disposal an increasingly diverse and sophisticated range of intervention methodologies. Increasingly, there is also demand from stakeholders, whether they are consumers, workers, insurance companies, government, or the general public, that mental health interventions be accountable. In other words, *bona fide* interventions need to offer and deliver measurable, positive impacts on clients and society.

This presents a significant challenge to the relatively new field of adventure therapy. In recent years, dialogue in professional development adventure therapy settings has largely centered on important issues which deal with defining adventure therapy, creating professional networks, qualifications, training, and ethics. These have been necessary emphases, indicative of a profession in the early stages of development. A key next step for advancing adventure therapy is the conduct of higher quality and more widespread research and evaluation. A critical task for this research is to establish a clear assessment of the relative efficacy of

James Neill is a faculty member in Outdoor Education at the University of New Hampshire. He also runs the "Outdoor Education Research and Evaluation Center" (http://www.wilderdom.com/research). *James' other publications can be downloaded from* http://www.wilderdom.com/JamesNeillpublications.htm. *James can be contacted via email:* james.neill@wilderdom.com.

adventure therapy programs. Meaningful benchmarks need to be established with which to compare the effects of various types of adventure therapy.

The current article aims to summarize existing meta-analytic evidence in outdoor education, psychology, and education in order to suggest some initial benchmarks which can be used to guide assessments of the relative impact of adventure therapy programs. In the future, these benchmarks should be updated by adventure therapy meta-analyses, and possibly, by combining primary outcome data from adventure therapy research and evaluation studies into a centralized database. In addition, this line of research should incorporate, where possible, qualitative insight from clinical studies.

How Can We Summarize Adventure Therapy Research Findings?

Getting a clear picture of the adventure therapy outcomes can be confusing because of the limited amount, variable quality, and difficult availability of adventure therapy research literature. A few sources, however, are particularly helpful. Gillis (1992, 2000; Gillis & Thomsen, 1996) has provided three well-considered, up-to-date critical reviews and the books by Gass (1993), and Davis-Berman and Berman (1994) provide probably the most significant collations and commentary on adventure therapy research and the field of adventure therapy, at least for North America. From an international perspective, the published proceedings of the first two International Adventure Therapy Conferences (Itin, 1998; Richards, in press) can be recommended, as can the growing range of Internet resources at sites such as the Wilderness Research Center (http://www.its.uidaho.edu/wrc/), Adventure Therapy Web (http://fdsa.gcsu.edu:6060/lgillis/AT/front.htm), Institute for Outdoor Leadership and Education (http://www.indiana.edu/~outdoor/IOLE.htm), and the Outdoor Education Research and Evaluation Center (http://www.wilderdom.com/research). Despite these sources, the adventure therapy field is notably undermined by a lack of well-organized, definitive, and widespread knowledge about the effectiveness of different types of adventure therapy programs.

Similar problems have been encountered in other related fields. For example, the value of psychotherapy has always been hotly debated and research has struggled to provide clarity. Problems associated with interpreting and summarizing many different research studies about psychotherapy led to a famous controversy. In 1961, Hans Eysenck published a seminal article in which he concluded that psychotherapy was not effective in changing clinical outcomes in clients, a claim which he continued to reiterate (1985). Eysenck's 1961 article sparked intense debate because there were claims that he had been biased in the articles he chose to review. In an attempt to resolve the issue, Gene Glass (1976, 1977) developed a technique in which he collated summary statistics (e.g., means, standard deviations, N, correlations) from different psychotherapy studies into a single database, coded each study's characteristics, and then statistically analyzed the overall outcomes and the relative influence of various independent variables (such as gender of client, length of treatment, etc.) on outcomes. Glass termed this procedure "meta-analysis" and claimed that it allowed for a more objective overview of quantitative data from many studies. Contrary to Eysenck's review, the Smith, Glass, and Miller (1980) meta-analysis found that psychotherapy, on average, had a moderately strong effect on clients' therapeutic outcomes (.68 standard deviations), akin to a 30% improvement in the measured outcomes.

Since 1980, meta-analysis has been used widely in medicine, psychology, and education, proving to be a powerful research summary tool, particularly in relation to understanding the effectiveness of treatments and interventions. While there has yet to be a major meta-analysis of adventure therapy outcomes, there have been four key meta-analyses of outdoor education programs (Cason & Gillis, 1994; Hans, 1997; 2000; Hattie, Marsh, Neill, & Richards, 1997; Marsh, 1999); innovative education in classrooms for affective outcomes (Hattie et al., 1997); and psychotherapeutic and psychological intervention research (Casey & Berman, 1985; Lipsey & Wilson, 1993; Smith, et al., 1980). The current article summarizes the results of these meta-analyses and thus proposes some initial benchmarks that can be used in future adventure therapy research and program evaluation studies.

Effect Sizes and Benchmarks

Meta-analyses report results in terms of "effect sizes" (ESs). In outcome studies, an ES is a measure of how much difference there is between measures of clients' well-being at two different points in time. Typically, the key ES of interest indicates the difference between measures before and after an intervention. It is also of interest to see whether there is maintenance, ongoing growth, or loss of any immediate or short-term change.

In statistical terms, an ES is the mean difference expressed in standard deviation units. An ES of zero indicates no change, a negative ES means a reduction in an outcome, and a positive ES generally refers to an improvement in an outcome measure. Various other interpretations of ESs have been suggested. For example, according to Wolf (1986) an ES of .50 represents a practical or clinically significant change while .25 is an educationally significant change. Cohen (1977) provided the most widely accepted guideline, suggesting that

.2 is small, .5 is moderate and .8 is large. Others warn that global guides should be avoided, and that in order to interpret a particular ES it is necessary to compare the result with closely related effect sizes. It should also be noted that a small ES can be very impressive if, for example, the outcome is difficult to change (e.g., a dysfunctional behavior) or if the outcome is highly valued (e.g., recidivism). On the other hand, a large ES does not necessarily mean that there is therapeutic value in the change, particularly if it is not related to the aims of the treatment (e.g., religious orientation or physical fitness).

It is also important to emphasize that program outcomes should not be compared against zero, no change. It makes more sense to compare program outcomes with minimum acceptable outcomes or other internally or externally established benchmarks. Benchmark comparisons can provide valuable diagnostic information about intervention strengths and weaknesses and help stakeholders to make informed decisions.

Contextualizing Adventure Therapy Outcomes

Currently, the most relevant benchmarks for evaluating adventure therapy programs can be derived from meta-analyses (Bunting & Donley, 2001; Cason & Gillis, 1994; Hans, 1997, 2000; Hattie et al., 1997; Marsh, 1999) and summaries of meta-analyses in outdoor education (Neill, 2002; Neill & Richards, 1998), innovative classroom-based education (Hattie et al., 1997), psychological training programs (Lipsey & Wilson, 1993), and psychotherapy (Casey & Berman, 1985; Smith, Glass, & Miller, 1980) (see Table 1).

The outdoor education meta-analytic studies represent results from approximately 12,000 participants and indicate small to medium overall outcomes (ESs between approximately .3 to .4) on typically measured outcomes such as changes in self-concept, self-confidence and locus of control. According to Hattie and al. (1997), these effects seem to increase even further over followup periods of up to 18 months, but no other meta-analyses provide comparable followup results. Despite these promising overall outcomes, there is wide variability, with ESs varying substantially between different types of outdoor education programs (Hattie, et al.). For example, only a small immediate effect of .17 (for 416 effects) was found for non-Outward Bound Australia research studies. The most effective outdoor education programs were those which where longer, involved adult-age participants, and were conducted by particular organizations, such as Outward Bound. Thus, the meta-analytic outdoor education results caution against blanket claims about the effectiveness of outdoor education and adventure therapy programs. Instead, they suggest promising potential and encourage closer analysis of particular types of adventure-based programming.

It is possible to derive from outdoor education meta-analyses categories of reporting (i.e., particular types of programs and particular types of outcomes) which more clearly pertain to adventure therapy programs. Such results are presented in Table 2. According to the adventure therapy results in Table 2, the overall adventure therapy effects appear to be positive and ongoing. While adventure therapy outcomes appear to be stronger than for outdoor education, they are not nearly as strong as for individual psychotherapy (compare Table 2 with Table 1). This suggests that there is considerable room for developing more effective adventure therapeutic practices and achieving strong outcomes more consistently.

An exceptional ES (1.05) was reported for "clinical scales" used with adolescents (12 effects; see Table 2; Cason & Gillis, 1994). The studies which contributed to

Table 1. Effects Sizes from Outdoor Education, Education, and Psychological Meta-analyses

Focus	Client group	No. of effects	ES Program	ES Follow-up
Overall adventure education programming[a]	Adolescents	147	.31	-
Overall outdoor education[b]	All	1728	.34	.17
Overall camping (self-concept & self-esteem)[c]	Children/ Adolescents	37	.21	-
Innovative classroom interventions (affective outcomes)[d]	Adolescents	24780	.28	-
Overall psychological training[e]	All	9400	.47	-
Overall psychotherapy[f]	Adults	1766	.68	-
Overall psychotherapy[g]	Children	75 Studies not reported	.71	-

Note. [a]Cason & Gillis, 1994; [b]Hattie, et al., 1997; [c]Marsh, 1999; [d]Hattie, J.A., 1992; [e]Lipsey & Wilson, 1993; [f]Smith, et al., 1980; [g]Casey & Berman, 1985.

Table 2. Effects Sizes for Adventure Therapeutic Outcomes, Derived from Meta-analyses and Program Evaluations of Outdoor and Adventure-based Programs

Outcome	Client Group	No. of Effects	ES Program	ES Follow-up
Behavior (rated by others)[a]	Adolescents	23	.40	n.a.
Clinical scales[a]	Adolescents	12	1.05	n.a.
Multiple outcomes[a]	Emotionally or physically challenged	21	.42	n.a.
Multiple outcomes[b]	Adjudicated	148	.30	n.a.
Multiple outcomes[b]	Delinquents	80	.33	.32
Aggression reduction[b]	Multiple groups	7	.33	.72
Neurosis reduction[b]	Multiple groups	33	.31	.24
Recidivism[b]	Multiple groups	3	.55	.10
Well-being[b]	Multiple groups	43	.24	-.09
Locus of Control (towards internal)[c]	Adults and adolescents	7	.38	n.a.

Note. [a]Cason & Gillis, 1994; [b]Hattie, et al., 1997; [c]Hans, 1997; 2000.

this overall high ES are worthy of closer investigation. Cason and Gillis do not provide identifiable referencing for these studies but they explain that the majority of the high ESs for clinical scales emerged from studies with residential populations in treatment centers or adjudicated youth in alternative correctional placements. In these residential settings, participants are more likely to experience adventure programming as part of a broader intervention spectrum, which is likely to contribute to more emotionally intense experience and stronger outcomes (Cason & Gillis). This claim would seem to be supported by the Hattie and al. (1997) meta-analysis results in which short-term outdoor education program effects were similar for delinquent (.33) and normal (.35) participants, but during the follow-up period were stronger for delinquents (.32) than for other outdoor education participants (.14). However, some caution should be applied to avoid over-interpreting these comparisons between "normal" and "delinquent" groups. Adventure therapy participants are likely to start lower on psychological measurement scales and thus have greater "regression to the mean" effects and more statistical "room for growth."

Other useful comparative ESs for adventure therapy outcomes can be drawn from the Smith and al. (1980) meta-analysis of adult psychotherapy outcomes (.68), and Casey and Berman's (1985) meta-analysis of child psychotherapy outcomes (.71) (see Table 1). It should be remembered that these impressive outcomes for psychotherapy are generally derived from one-on-one therapeutic work conducted by trained therapists in periodic session-based intervention regimens.

Lipsey and Wilson (1993) report a more moderate, but still impressive ES for psychological change and training programs (ES = .47), based on summarizing 156 meta-analyses, with over one million participants. Such a comprehensively derived ES could be a very useful general benchmark for evaluation purposes in adventure therapy. Ideally, however, adventure therapy researchers and program evaluators should utilize the breakdowns of specific outcome categories reported by psychotherapy meta-analyses and the psychological intervention meta-analyses (Casey & Berman, 1985; Lipsey & Wilson, 1993; Smith et al., 1980) and then develop specific benchmarks appropriate to their particularly program and client group.

Limitations of Meta-analysis

While meta-analysis offers a more systematic way of reviewing research findings than traditional literature reviews, there are several limitations to be borne in mind. First, meta-analysis can only report on areas in which there have already been several primary research studies conducted. Thus, a meta-analysis reflects the existing scope of empirical research, but it cannot fill gaps where research is yet to be conducted. A second issue is that in order to be included in a meta-analysis, a study must report sufficient descriptive statistics (*M, SD* and *N*), which has not always been the case, particularly in low quality adventure therapy research studies.

A further important caution is that the published literature may not accurately represent actual adventure therapy programming. For example, a recent survey of outdoor education programs found that only 7% of programs used standardized outcome measures (Richards, Neill, & Butters, 1997). Of these 7%, an even smaller

percentage has their results published. Thus, the published data probably emerges from less than 1% of existing programs and may lack representativeness. It is also well known that journals are more likely to publish statistically significant results. A weakness of meta-analysis, then, is that it may unwittingly contribute to overestimating adventure therapy program effectiveness. Readers are cautioned against possible misuse of the available statistics. The available results should be used as guides for meaningful discussion and not for unjustified self-promotion.

Conclusions

Gillis (1992) made the recommendation that "someone needs to conduct a meta-analysis of therapeutic aspects of adventure-challenge-outdoor-wilderness that includes the criteria of clinical significance along with traditional methods of effect size" (p.7). Four years later, Gillis and Thomsen (1996) made two further key points —that the field of adventure therapy needs to create a collective document that addresses its effectiveness and that clinically significant events of adventure therapy need to be examined through a massive survey of consumers. In other words, Gillis has called for major quantitative and qualitative research studies to be conducted across the field of adventure therapy. Building on these suggestions, a proposed agenda for future adventure therapy research could be to conduct: (a) meta-analytic investigations of specific aspects of adventure therapy programming; (b) good quality and well disseminated program evaluations which utilize meta-analytic outcomes as benchmarks; (c) innovative, large-scale qualitative investigations which identify and seek to better understand clinically significant moments and processes which occur in adventure therapy programs; and (d) the possible development of an international adventure therapy outcomes online database.

Integral to conducting good quality research and program evaluation studies is the development of purpose-built, multidimensional evaluation tools which are rigorously developed using the best available psychometric techniques (e.g., see Ewert & Sibthorp, 2001). This is the current expectation in educational and psychological research and adventure therapy researchers should strive for no less. To date, the adventure therapy profession is yet to develop or incorporate into its work an appropriate suite of clinical outcome measurement tools. In the absence of specific adventure therapy tools, researchers should be considering well-developed and readily available mental health instruments, such as the *General Well-Being* instrument, which assesses positive and negative aspects of reported mental distress and well-being (Heubeck & Neill, 2000; Veit & Ware, 1983), and the *Youth Outcomes Questionnaire* used in Russell's investigations of outdoor behavioral healthcare program outcomes (Russell, 2001; 2002).

In the twenty-first century, the specter of accountability will continue to grow. Thus, in order for growth to continue in the adventure therapy profession, more sophisticated literature review methods and program evaluation practices are required. Benchmarks derived from meta-analysis offer a systematic way to categorize and evaluate outcomes. A comparison of a program's ESs with appropriate meta-analytic benchmarks can provide useful diagnostic information about the program's relative efficacy.

It is recommended that every adventure therapy program examine its evaluation methods to ensure that the most up-to-date and valid methods of comparison are being used. To be doing anything less borders on professional malpractice. Adventure education and therapy organizations have now generally accepted that financial and safety auditing is part of successful program management, but educational or therapeutic auditing is far from fully embraced.

The initial benchmarks suggested in this article are necessarily rough (~.4) and should be revised and expanded as new evidence becomes available. It is important that quantitative adventure therapy research and evaluation studies endeavor to report breakdowns of results in terms of programming factors such as client problem(s), intervention length, sequencing of activities, facilitation style, and so on. By doing so, we can better understand possible causal elements of the intervention and future meta-analyses can provide more insightful analyses of the extent to which adventure therapy outcomes are influenced by various programming and facilitation factors.

In the twenty-first century, it is possible to apply innovative research methods to fuel the professional growth of new fields such as adventure therapy. For example, a central electronic clearinghouse of adventure therapy research and evaluation studies, and outcome data could be developed. Such innovative efforts to integrate knowledge and practice could be the difference between adventure therapy remaining as a fringe method for psychological treatment or being developed into genuinely effective methods for the alleviation and treatment of psychological problems.

References

Bunting, C. J., & Donley, J. P. (2002, January). *Ten years of challenge course research: A review of affective outcome studies.* Poster session presented at the 6th Coalition for the Education in the Outdoors Research Symposium, Bradford Woods, IN.

Casey, R. J., & Berman, J. S. (1985). The outcome of psychotherapy with children. *Psychological Bulletin, 98,* 388-400.

Cason D., & Gillis, H. L. (1994). A meta-analysis of outdoor adventure programming with adolescents. *Journal of Experiential Education, 17*(1), 40-47.

Cohen, J. (1977). *Statistical power analysis for behavioral sciences* (revised ed.). New York: Academic Press.

Davis-Berman, J., & Berman, D. S. (1994). *Wilderness therapy: Foundations, theory and research.* Dubuque, IA: Kendall/Hunt.

Ewert, A., & Sibthorp, J. (2001). Multivariate analysis in experiential education: Exploring the possibilities. *Journal of Adventure Education, 23,* 108-117.

Eysenck, H. J. (1961). The effects of psychotherapy. In H. J. Eysenck (Ed.), *Handbook of abnormal psychology* (pp. 697-725). New York: Basic.

Eysenck, H. J. (1985). The battle over psychotherapeutic effects. In J. Harriman (Ed.), *Does psychotherapy really help people?* (pp. 52-61). Springfield, IL: Charles C. Thomas.

Gass, M. A. (Ed.). (1993). *Adventure therapy: Therapeutic applications of adventure programming.* Dubuque, IA: Kendall/Hunt.

Gillis, H. L. (1992, January). *Therapeutic uses of adventure-challenge-outdoor-wilderness: Theory and research.* Paper presented at the Coalition for Education in the Outdoors Symposium, Bradford Woods, Indiana University, Martinsville, IN.

Gillis, H. L. (2000, March). *Going around in circles: Research and program evaluation in therapeutic adventure and adventure therapy.* Paper presented to the Second International Adventure Therapy Conference, Augsburg, Germany.

Gillis, H. L., & Thomsen, D. (1996, January). *A research update (1992 –1995) of adventure therapy: Challenge activities and ropes courses, wilderness expeditions, and residential camping programs.* Paper presented at the Coalition for Education in the Outdoors Symposium, Bradford Woods, Indiana University, Martinsville, IN.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher, 5,* 3-8.

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in Education, 5,* 351-379.

Hans, T. (1997). *A meta-analysis of the effects of adventure programming on locus of control.* Unpublished master's thesis, Georgia College, Milledgeville, GA.

Hans, T. (2000). A meta-analysis of the effects of adventure programming on locus of control. *Journal of Contemporary Psychotherapy, 30*(1), 33-60.

Hattie, J. A. (1992). Enhancing self-concept. In J. M. Hattie (Ed.), *Self-concept* (pp. 221-240). New York: Lawrence Erlbaum.

Hattie, J. M., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and Outward Bound: Out-of-class experiences that have a lasting effect. *Review of Educational Research, 67,* 43-87.

Heubeck, B., & Neill, J. T. (2000). Confirmatory factor analysis and reliability of the Mental Health Inventory for Australian Adolescents. *Psychological Reports, 87,* 431-440.

Itin, C. M. (Ed.). (1998, July). *Proceedings of the First International Adventure Therapy Conference: Exploring the Boundaries.* Camping and Outdoor Education Association of Western Australia, Perth, Australia. Boulder, CO: The Association of Experiential Education.

Lipsey, M. W., & Wilson, D. B. (1993). The efficacy of psychological, educational, and behavioral treatment. *American Psychologist, 48,* 1181-1201.

Marsh, P. E. (1999). *What does camp do for kids? A meta-analysis of the influence of the organized camping experience on the self constructs of youth.* Unpublished master's thesis, Indiana University, Martinsville, IN.

Neill, J. T. (2002, January). *Meta-analytic research on the outcomes of outdoor education.* Paper presented at the 6th Biennial Coalition for Education in the Outdoors Research Symposium, Bradford Woods, IN.

Neill, J. T., & Richards, G. E. (1998). Does outdoor education really work? A summary of recent meta-analyses. *Australian Journal of Outdoor Education, 3*(1), 2-9.

Richards, K. (Ed.) (in press). *Therapy within adventure: Proceedings of the Second International Adventure Therapy Conference,* March 20-24, 2000, Augsburg, Germany.

Richards, G. E., Neill, J. T., & Butters, C. (1997, January). *Summary statistical report of attendees at the 10th National Outdoor Education Conference, Colloroy, NSW, Australia.* Canberra, Australia: National Outdoor Education & Leadership Services.

Russell, K. C. (2001). *Assessment of treatment outcomes in outdoor behavioral healthcare.* Technical Report 27, Idaho Forest, Wildlife, and Range Experiment Station, Moscow, ID. Retrieved December 1, 2002, from University of Idaho-Wilderness Research Center Web site: http://www. wrc@uidaho.edu.

Russell, K. C. (2002). *A longitudinal assessment of treatment outcomes in outdoor behavioral healthcare.* Technical Report 28, Idaho Forest, Wildlife, and Range Experiment Station, Moscow, ID. Retrieved December 1, 2002, from University of Idaho-Wilderness Research Center Web site: http://www. wrc@uidaho.edu.

Smith, M. L., Glass, G. V., & Miller, T. I. (1980). *The benefits of psychotherapy.* Baltimore: Johns Hopkins University Press.

Veit, C., & Ware, J. (1983). The structure of psychological distress and well-being in general populations. *Journal of Consulting and Clinical Psychology, 51,* 730-742.

Wolf, F. M. (1986). *Meta-analysis: Quantitative methods for research synthesis.* Beverly Hills, CA: Sage.