Meta-Analytic Research on the Outcomes of Outdoor Education

James T. Neill
University of New Hampshire

Paper presented to the 6th Biennial Coalition for Education in the Outdoors Research Symposium, Bradford Woods, IN, 11-13 January

Author’s note.
Correspondence concerning this article should be addressed to the author at the Department of Kinesiology, University of New Hampshire, 160 Main Street, New Hampshire Hall, Durham, NH 03824, USA.
Phone: (603) 862 3047
Fax: (603) 862 0154
Email: james.neill@unh.edu

ABSTRACT

The impacts of outdoor education programs on participants remains a topic of much interest to practitioners and researchers. Hundreds of empirical pre-post studies have been conducted in an attempt to better understand the impacts. Typically, psychosocial constructs such as self-concept are measured before and after outdoor education programs. In order to summarize the results of such studies, five relevant meta-analyses have been conducted (Bunting & Donley, 2002; Cason & Gillis, 1994; Hans, 2000; Marsh, P. E., 1999; Marsh, H. W., Neill, & Richards, 1997). These research reviews show that education programs have small-moderate impacts on constructs such as self-concept, locus of control, and teamwork. Impressively, the effects appeared to be retained over time. The most effective programs were conducted by particular organizations, involved adult participants, and were longer. Comparisons to meta-analyses for other types of interventions revealed that outdoor education program impacts were similar to other self-concept change programs and other psychological interventions, but not as powerful as individual psychotherapy. Recommendations that emerged from this research review were that future research should seek to understand more about why different programs and different participants achieve different outcomes and for effect sizes to be used by outdoor education programs to monitor educational effectiveness.

WHAT ARE THE EFFECTS OF OUTDOOR EDUCATION?

Are people any different after participating in challenging adventure activities? Which programs or participants have the most dramatic effects? What are the effects of outdoor education (OE)? Such questions have attracted animated discussion since modern forms of OE started appearing over 50 years ago.

Typically, the effects of OE have been in the form of testimonial support, anecdotal examples, and passionate rhetoric. For example, the promontional material of OE organizations invariably communicates highly positive claims and examples, OE instructors’ verbiage with clients exhibits highly optimistic and positive hopes, and a surprising amount of so-called academic work has a clear tendency to uncritically promote the view that OE programs are good things for people.

For most OE advocates the ready-at-hand, ad hoc evidence is sufficient to support their enthusiastic claims about OE. Indeed, the smile when a participant conquers a challenge and the glowing words spoken during a debrief are often compelling. But is it enough to let such evidence speak for itself?

A list of claims about the effects of OE would be very long. Yet for most advocates of OE, what they have personally seen, heard and felt provides sufficient evidence for many wider claims.
Such a critique of OEs may seem harsh. But professional OEs should know more about research results, as should all stakeholders in OE programs and related fields of endeavor. Indeed, so should researchers of OE! To date, however, there has been a lack of high quality research reviews that have been written for both researchers and practitioners. Thus the current article undertakes to compare and summarize the findings of empirical outcome research on OE and related programs for both general audiences. A challenging aim for the article is to provide maximum information about the empirical outcome findings in OE in a minimum of space.

Initially this paper overviews two ways of measuring outcomes, then describes the use of meta-analysis as a method for objectively summarizing pre-post empirical research results, and finally provides recommendations for future OE research and evaluation studies.

WAYS OF RESEARCHING OUTCOMES
There have been two main objective ways that OE outcomes have been researched.

Post-program Surveys
One of the more common has been to directly ask participants at the end of a program something like “How valuable was this program for your personal growth?”. On face value this approach has appeal. Indeed, it appears to be a prevalent method. In a recent survey of the practices of over 100 Australian outdoor educators approximately 80% reported that their program used written end-of-program surveys, or ‘happy sheets’ as they are colloquially known (Richards, Neill, & Butters, 1997). However, such evaluations are notoriously problematic indicators of program effectiveness. For example, a study by Neill, Richards and Badenoch (1997) found little relationship between staff and participants’ ratings of amount of personal growth experienced by participants (correlations between .2 and .4). Similarly low correlations were found between participants’ end-of-program self-reports and their pre-post changes on reliable, psychometric questionnaires.

End-of-program reports of the impact of the value of an experience are one valuable source of evidence. However, post-hoc reports are vulnerable to many potential distortions. For example, scores can be inflated by post-group euphoria (Marsh, Richards, & Barnes, 1986), post-purchase rationalization (or post-commitment justification), group think, and other problems related to difficulties in validity remembering, understanding, and then accurately reporting on what happened during an experience and resultant effect on one’s psyche.

Disparities between different ways of assessing intervention programs is a broader phenomenon. For example, in a review of management training research, it was claimed that the “reaction of trainees bore no relationship to the effectiveness of the training” (Bright, cited in Lawson, 1997, p. 7).

To date very little research has explored the relationships between what stakeholders, such as staff and participants, say about a program’s value and other indicators of the program effectiveness. Thus, the validity of post-experience ratings remains an area for further investigation.

Pre-post Testing
A second major approach to examining the effectiveness of OE has been to gather participants’ self-perceptions (or observer perceptions) before and after a program and then test to see if there are any differences. Hundreds of empirical evaluation studies have utilized the basic pre-post design, or variations of it. Such studies have used a wide variety of dependent measures (e.g., self-esteem, locus of control, etc.) and have been conducted on several different types of programs. The validity of this approach depends on important factors such as the quality of measurement tool, the use of control or comparison groups, and whether followup testing is conducted (Campbell, & Stanley, 1963).

The variety in the methods and measures of OE pre-post outcome research makes the task of reviewing the literature formidable. There have been several notable reviews of the empirical research (e.g., Barrett & Greenaway, 1995; Burton, 1981; Crompton & Sellar, 1981; Ewert, 1982; Gillis, 1992; Gillis & Thomsen, 1996; Godfrey, 1974; Iida, 1975; Reddrop, 1997; Richards, 1977; Shore, 1977),
however they have all faced difficulties in presenting a systematic overview of the outcome research. Perhaps an indicator of this is that there is only one review which has been published in a peer-reviewed journal (Crompton & Sellar, 1981) and it is over 20 years old.

One of the main difficulties faced by these reviews has been to effectively synthesize and systematically summarize the empirical findings from the variety of empirical studies. Cason and Gillis (1994) described the problem that “the number of research endeavors in adventure programming has grown so large that qualitative literature reviews (like annotated bibliographies) are no longer sufficient tools for understanding the strengths and weaknesses of trying to measure change that results from adventure programming Cason and Gillis” (1994, p. 40).

Since the 1980’s, a new, systematic review method appeared in the form of meta-analysis. Meta-analysis provides a method for entering the empirical results reported by different studies into a central database which is then analyzed and summarized. Well conducted meta-analyses allows for more objective appraisal of the outcome evidence, which can lead to clearer understanding about the nature of OE’s impacts on participants. To date, five meta-analytic reviews of OE and related outcome literature have been conducted, and it is these five reviews which the current article summarizes.

UNDERSTANDING META-ANALYSIS

Meta-analysis is a statistical technique for accumulating and representing the research results reported in various studies. Meta-analysis pools the findings about a single research question from many different sources and analyzes the overall effects.

Meta-analysis results are reported in terms of effect sizes (ESs). An ES, as used in this article, is a measure of how much difference existed between ratings at two different points in time (e.g., before and after an OE program). In essence, the ES indicates the quantified amount of change.

For general interpretation, an ES of 0 means no change, a negative ES means a reduction for a measured outcome, and a positive ES means an increase in a measured outcome. It can also be noted that ESs are proportional. Thus, an ES of .40 represents twice as much change as an ES of .20.

Various experts have offered suggestions about the meaning of different ESs. For example, it has been suggested that an ES of .50 is a change of practical significance while .25 is educationally significant (Wolf, 1986). Cohen (1977) suggested that .2 was small, .5 moderate, and .8 large. Other authors have warned against global guides to interpreting ESs since who and what is being measured needs to be identified before a genuine assessment about the value of a particular ES is made.

In practice, a small ES can be very impressive if, for example, the outcome is difficult to change (e.g., a personality construct) or if the outcome is very valuable (e.g., an increase in life expectancy). On the other hand, a large ES does not necessarily mean that there is any practical value in the change, particularly if it isn’t related to the aims of the intervention (e.g., political orientation).

Meta-analysis as a review method is not without its critics. Criticisms tend to fall into two categories (Bangert-Drowns & Rudner, 1991). Some researchers complain that meta-analysis obscures important information by averaging numerical representations across studies. Other researchers argue that research is best reviewed by a reflective expert who sifts through research and develops kernels of insight from the array of argumentation in a field. Thus, there are advantages and disadvantages to both traditional style and meta-analytic reviewing. An ideal reviewing method, then might involve making use of both meta-analytic methods and traditional review methods.

META-ANALYTIC OUTDOOR EDUCATION RESEARCH

Five meta-analyses of the effects of OE and related programs have been conducted (Bunting & Donley, 2002; Cason & Gillis, 1994; Hans, 2000; Marsh, P.E., 1999; Marsh, H.W., Neill, & Richards, 1997; see Table 1). The studies vary considerably in their scope and focus. The major study (Hattie, Marsh, H.W., Neill, & Richards, 1997) included 96 studies, representing over 12,000 participants, included
OE programs results for both youth and adults, coded for pre-program effects, program effects and post-program effects. Cason and Gillis (1994) conducted the first OE meta-analysis, based on pre-post outcomes reported in OE research with adolescents (43 studies). Most of the studies in the Cason and Gillis (1994) study were included in the Hattie et al. (1997), hence the results of the two studies are generally consonant.


Overall Findings

The five OE meta-analyses reported that OE programs have small-moderate effects (see Table 1). The lowest average ES was reported for American camp programs (.20; Marsh, P. E., 1999) and the highest average ES was reported for ropes challenge courses (.55; Bunting & Donley, 2002). For the other studies, Hans (2000) reported an ES of .38, Hattie et al. (1997) reported .34 and Cason and Gillis (1994) reported .31. The findings for the larger studies and the ones focused on OE programs (i.e, Cason & Gillis; Hattie et al., 1997) suggest overall ESs between .3 and .4. By broader educational and psychological standards this is be considered a small-moderate amount of change (Cohen, 1977).

Camping programs’ seemingly low ES of .20 may not be unexpected, given that not all camps aim to enhance campers’ personal and/or social development. Indeed, camping programs that had a personal development philosophy had an ES of .41. It should also be noted that Bunting and Donley’s (2002) relatively high overall ES of .55 for ropes challenge courses was based on only 15 studies, and these studies exhibited a wide range in ESs. Further undermining the reliability of this study is issue that a strong negative relationship was found between quality of study and size of outcome; in other words, poorer quality studies tended to have higher ES, a finding also reported by Cason and Gillis (1994).

In summary, a reasonable consensus amongst the five OE meta-analyses is that OE programs have a small-moderate impact for typically measured outcomes such as self-esteem, behavior problems, and teamwork. In an attempt to be more precise, Hattie et al. (1997) reported that their results suggested that 65% of participants were better off for having participated in OE programs.

Do the Effects of Outdoor Education Programs Last?

Hattie et al., (1997) examined the long-term effects of OE programs and reported an additional ES of .17 for assessments up to 18 months after the end of the program. This finding means that the average program ES of .34 is followed by an average ongoing improvement of .17, suggesting a moderate overall ES of .51. This evidence suggests that, on average, OE participants experience additional growth on returning to their home environments. Despite wide ES variations between different studies, Hattie et al. (1997) noted that the additional .17 is impressive because longitudinal education and training research generally shows a loss over time of immediate program benefits. However, it should also be noted that much of the research on which these long-term results were based was Outward Bound expedition-style programs, thus the generalizability of the finding is limited.

What are the Main Outcomes of Outdoor Education Programs?

While overall outcomes are useful as a general guide, detailed information about the effects is needed. Hattie et al. (1997) grouped the results into six broad categories – leadership, self-concept, academic, personality, interpersonal, and adventuresome (see Table 2).

The pre-post program ESs for each of Hattie et al.’s (1997) six categories were statistically significant. Interestingly, the smallest program ES was for self-concept (.28), however the largest follow-up ES was for self-
This might indicate a sleeper effect whereby self-concept changes in particular are begun during a program and then continue to unfold afterwards.

The ESs reported by Cason and Gillis (1994) for programs with adolescents were placed into seven categories – behavioral assessment by others, self-concept, academic, attitude, locus of control, clinical scales, and school attendance (see Table 3). There was a very high ES (1.05) for studies using clinical measurement scales. This finding could be due to the fact that clinical scales are mainly used for researching specialized and intensive OE programs for participants who have emotional, behavioral or psychological difficulties. Such programs tend to utilize a more highly disciplined and individualized approach than is usually the case in programs with an personal development, educational or recreational emphasis. High ES for therapeutic and/or behavioral problem clients may also be due to the clients having a greater room for growth than average OE program participants. The results for the other six categories of outcomes reported by Cason and Gillis (1994) ranged between .30 (locus of control) and .61 (academic).


Roughly, the small-moderate average ES finding applies to different outcomes categories, with the notable exception of clinical scales in the Cason and Gillis (1994) study. The largest followup effects appear to be in the area of self-concept. The overall effects seem to suggest enhancement of self-related constructs, action-orientation, and coping behaviors.

### Are All Outdoor Education Programs Effective?

According to Hattie et al. (1997, p.77), OE programs “can obtain notable outcomes and have particularly strong, lasting effects. It is clear, however, that adventure programs are not inherently good. There is a great deal of variability in outcomes between different studies, different programs, and different individuals”.

This theme of wide variability in the ESs was also evident in the Cason and Gillis (1994, p. 46) meta-analysis which stated that “when one considers the large range of effect sizes collected, -1.48 to 4.26, and the wide variation in the findings, a singular summary effect size...cannot adequately reflect the effectiveness of adventure programming”.

Other research reviews have come to similar conclusions about variability in outcomes. Burton (1981), for example, reviewed 161 OE studies and reported that 38% had “mostly positive” outcomes, 26% had “some positive” outcomes, 34% had “non-significant” outcomes and 2% had “negative” outcomes.

This variability in outcomes between programs is a striking feature in the OE research and warns against accepting false comfort from the small-moderate average ESs. Hattie et al. (1997, p.70) emphasized that “only some adventure programs are effective, and then on only some outcomes, and it is probable that only parts of the programs are influencing these outcomes”.

### Moderators of Outdoor Education Outcomes?

Meta-analytic methods of summarizing research can be used to consider the effect of moderator (or process) variables on outcomes. The meta-analytic studies have examined the influence of program length, participant gender, participant age, participant type (e.g., delinquent, management, gender mix of group (all male, all female, co-educational), nature of program (e.g. sailing, wilderness), organization running the program and quality of the study.

Hattie et al. (1997) found that 36% of the variance in outcomes could be explained by the abovementioned factors. Outcomes were most influenced by the organization running the program (Outward Bound Australia, ES = .47; all other OE program, ES = .17), program length (programs greater than 20 days, ES = .41; less than 20 days, ES = .23), and age of
participants (adults, ES = .38; students, ES = .21).

Cason and Gillis (1994) found that the quality of the study had an impact on the outcomes, with lower quality studies reporting higher ESs, as did Bunting and Donley (2002). This should warn researchers to be wary of using instrumentation and designs which are below par since they may inflate results. This concern becomes even more important in the light of how few programs use high quality assessment tools. According to a recent survey only 7% of outdoor educators reported that their programs utilized standard tests as part of their evaluation strategy (Richards, Neill, & Butters, 1997).

Hattie et al (1997) and Cason and Gillis (1994) reported a weak but positive relationship between program length and the size of outcome, and Bunting and Donley (2002) reported a correlation of .28 between the length of ropes course programs and outcomes. Hans (2000) did not report any effects for program length, and length was not used by Marsh, P. E. (1999). Overall, OE program length explains a small but consistent amount of variance in participant outcomes.

With regard to the effect of participant age, stronger outcomes were found by Cason and Gillis (1994) for younger rather than older adolescents. This negative correlation between age and outcome was also found for the camping studies (Marsh, P. E., 1999). This is consistent with other research suggesting that older adolescence is a period during which self-concept is somewhat resilient to change (e.g. see Hattie, 1992). However, the strongest outcomes of all were found for adult-aged participants (Hattie et al., 1997), although Bunting and Donley (2002) found linear effects of age on ES, from school-age, through college-age to corporate client-age.

Hans (2000) found two significant influences on locus of control outcomes. Firstly, programs that had therapeutic goals as their primary purpose had higher ESs than programs with goals such as recreation. This finding seemed to be consistent with Marsh, P. E.’s (1999) finding that camps with a self-development philosophy achieved a reasonably high ES, whereas for camps without such a philosophy, participant changes were negligible. These findings lend credibility to the possibility that organizational mission and culture is a critical factor underlying program effects and this could explain the strong ESs that have been reported for Outward Bound programs in Australia (Hattie et al, 1997).

The second significant moderator from the Hans (2000) study was that residential and semi-residential programs were more effective than sessional programs which utilized adventure activities but did not take participants away from their familiar environments overnight.

In summary, the main influences on empirical OE research outcomes are the OE organization running the program, the age of participants, and the length of program. Other moderators of note were quality of study and whether the program was residential. Overall, it appears that the results of OE program research show that there are small-moderate average effects which vary considerably from participant to participant and from program to program.

OUTDOOR EDUCATION OUTCOMES COMPARED TO OTHER INTERVENTIONS PROGRAMS

Research on many educational, training and therapeutic intervention programs can provide useful comparison with OE outcomes. The most famous meta-analytic study (based on 475 studies) reported an average ES of .70 for psychotherapy (Smith, Glass & Miller, 1980), noticeably greater than the average OE ESs, although it should be noted that psychotherapy tends to be conducted with a different clientele than is normally the case for OE programs.

Another useful comparison is Lipsey and Wilson’s (1993) meta-analyses on psychological treatments, a study which represented over one million participants. The overall ES was moderate (.47), with 83% of the reviewed meta-analyses finding ESs over .20. Lipsey and Wilson stated that “little in conventional reviews and past discussions of these treatment areas, either individually or collectively…prepares a reviewer for the rather stunning discovery that meta-analysis shows nearly every treatment examined to have positive effects” (p. 1192). The Lipsey and Wilson (1993) study provides OE programs with a useful comparative
benchmark, particularly for OE programs that seek recognition for effecting psychological changes.

Another useful comparison can be made with non-OE programs designed to enhance self-concept. Hattie, J. A. (1992) conducted a meta-analysis of 89 self-concept enhancement studies, and found an average ES of .37. This ES represents an average of outcomes for self-concept change programs such as psychology courses, behavior therapy, cognitive therapy, and client-centered counseling. An ES .37 for non-OE self-concept program changes is similar to OE programs. As with the OE meta-analyses, however, the Hattie, J. A. (1992) study of non-OE program effects found large variations depending on the type of intervention program.

RECOMMENDATIONS
On the basis of the research reviewed, three main recommendations emerge for future research and evaluation in OE. Firstly, future research should describe the program’s methods in more detail. Secondly, investigation is needed of the influence of individual differences on outcomes. Thirdly, ES benchmarks should be used in program evaluation studies.

FUTURE RESEARCH SHOULD DESCRIBE THE PROGRAM’S METHODS IN MORE DETAIL
All five OE meta-analyses were consistent in calling for future research to provide more details about the educational methods used in the program being investigated. The lack of methods information in the primary empirical studies limited the extent to which meta-analysis could analyze process variables. A lack of more detailed information about program features such as program philosophy, type of activities, instructional techniques and facilitation style, instructor experience, program difficulty, weather, environment, group sizes, activity sequencing, and methods for dealing with behavior problems was noted. Given the wide variation of outcomes between different types of programs it is critical that future research must undertake to document in much greater detail the nature of the programs being conducted.

An argument can be made that future empirical studies be treated as mixed mode studies with indepth qualitative descriptions of the program process. Other intervention research areas, such as psychology, provide limited information because they use standardized treatment regimens. However, OE programs often lack reliable documentation. OE programming methods needs to receive greater attention in future research.

INVESTIGATION IS NEEDED OF THE INFLUENCE OF INDIVIDUAL DIFFERENCES ON OUTCOMES
Meta-analytic findings to date have provided little insight into the phenomenon of individual differences influencing OE participant outcomes. Research reviews have struggled to identify and provide models of relevant participant characteristics and processes that significantly influence the nature of their experience and learning. The lack of definitive evidence about causes of individual variations in outcomes may signal a need for different lines of investigation. Two recent studies have illustrated some possible approaches.

Fry and Heubeck (1998) examined the effects of personality on the mood states of participants in Outward Bound programs. In this study, extroverts and introverts had noticeably different mood patterns during the different phases of the program.

Neill and Heubeck (1998) found that school students who reported using productive coping skills (such as ‘focusing on the positive’, ‘working hard’ and ‘solving the problem) experienced greater mental health benefits from OE programs than those who reported using non-productive coping skills (such as ‘self-blame’, ‘worrying’ and ‘ignoring the problem’).

These two studies suggest that research on individual differences, such as personality or coping styles, may enlighten a path towards developing more effective, individualized programming in OE.

USE EFFECT SIZE BENCHMARKS IN PROGRAM EVALUATION
The reporting of ESs has utility beyond meta-analysis. Reporting of ESs is now considered a standard part of empirical reporting (American Psychological Association, 2001). ESs allow for comparative quantification of
program effectiveness. Program evaluation results can be compared with internal benchmarks (e.g., compare this program to last year’s program) or external benchmarks (e.g., compare this program with typical effects for OE programs or for other types of intervention programs).

An important contribution of such comparisons can be to provide feedback about a program’s relative strengths and weaknesses. This can be valuable diagnostic information to help identify areas for possible improvement, and to identify particular strengths of the program.

CONCLUSION

After five decades of modern day OE, the empirical outcome research has been synthesized. On average, OE programs appeared to have small to moderate effects on participants’ self-perceptions of personal qualities and capabilities. This finding is similar to the average outcome for psychological training and other types of educational self-concept change programs. These results provide scientific support that OE programs, on average, provide legitimate and effective educational training. It would seem that OE methods have something genuine to offer education, training, and psychology. A particularly impressive strength is that OE programs seem capable of triggering an ongoing cycle of positive change within participants.

However, readers should be warned against the comfort of average results. Some OE programs are much less effective and some OE programs are more effective. OE is not a panacea. Many programs which claim to deliver personal development through the outdoors may be ineffective and almost all programs could be substantially improved if the outcomes of psychological interventions and psychotherapy are accepted as reasonable benchmarks.

Unfortunately, most OE programs do not conduct systematic, rigorous evaluations of program effectiveness, nor are the methodologies of various OE programs well described and available for public dissemination. For the programs that conduct some evaluation, there are disturbingly few that use standardized testing instruments and too many rely on end-of-program written surveys as indicators of their effectiveness.

Further OE meta-analyses would also be useful. The Hattie et al. (1997) study could be updated. In addition, meta-analytic studies of specific aspects of OE programming, such as adventure therapy, would be useful. For example, a meta-analysis of adventure therapy programs could be conducted, as could a meta-analysis examining the influence of different facilitation techniques on participant outcomes.

OE scholars might also consider other innovative research reviewing methods. For example, a central electronic database of primary OE research data could be developed to allow powerful secondary and cumulative analyses.

It is no longer sufficient for OE advocates to simply believe in the benefits of OE, or to accept anecdotal evidence at face value. Instead, all available tools be used in order to understand how what effects OE programs have on participants and how the best possible learning experiences can be provided.

REFERENCES


### Table 1
**Overall Results from Five Meta-analyses related to Outdoor Education**

<table>
<thead>
<tr>
<th>Study</th>
<th>Focus</th>
<th>$d$</th>
<th>$N_{studies}$</th>
<th>$N_{effects}$</th>
<th>$N_{participants}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cason &amp; Gillis (1994)</td>
<td>Adventure programming for adolescents</td>
<td>.31</td>
<td>43</td>
<td>147</td>
<td>~7,030</td>
</tr>
<tr>
<td>Hattie et al. (1997)</td>
<td>Adventure education and Outward Bound programs</td>
<td>.34</td>
<td>96</td>
<td>1,728</td>
<td>12,057</td>
</tr>
<tr>
<td>Hans (2000)</td>
<td>Adventure programming locus of control outcomes</td>
<td>.38</td>
<td>24</td>
<td>30</td>
<td>1,632</td>
</tr>
<tr>
<td>Marsh, P. E. (1999)</td>
<td>Camping programs</td>
<td>.20</td>
<td>22</td>
<td>37</td>
<td>na</td>
</tr>
<tr>
<td>Bunting &amp; Donley (2002)</td>
<td>Ropes Challenge Courses</td>
<td>.55</td>
<td>15</td>
<td>na</td>
<td>na</td>
</tr>
</tbody>
</table>

### Table 2
**Effect Sizes (ESs) by Outcome Category reported by Hattie et al (1997)**

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
<th>Program ES ($N$)</th>
<th>Follow-up ES ($N$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leadership</td>
<td>Conscientiousness, Decision making, Leadership, Organisational ability, Time management, Values, Goals</td>
<td>.38* (222)</td>
<td>.15 (58)</td>
</tr>
<tr>
<td>Self-concept</td>
<td>Physical ability/appearance, Peer relations, General, Academic, Confidence, Self-efficacy, Family, Self-understanding, Well-being, Independence</td>
<td>.28* (271)</td>
<td>.23* (149)</td>
</tr>
<tr>
<td>Academic</td>
<td>Mathematics, Reading, Grade Point Average (GPA), Problem solving</td>
<td>.46* (30)</td>
<td>.21 (9)</td>
</tr>
<tr>
<td>Personality</td>
<td>Femininity/Masculinity, Achievement motivation, Emotional stability, Aggression, Assertiveness, Locus of control, Maturity, Neurosis reduction</td>
<td>.37* (235)</td>
<td>.14 (76)</td>
</tr>
<tr>
<td>Interpersonal</td>
<td>Cooperation, Interpersonal communication, Social competence, Behavior, Relating skills, Recidivism</td>
<td>.32* (176)</td>
<td>.17 (36)</td>
</tr>
<tr>
<td>Adventure-some</td>
<td>Challengeness, Flexibility, Physical fitness, Environmental awareness</td>
<td>.38* (69)</td>
<td>-.06 (19)</td>
</tr>
</tbody>
</table>

*Note. * $p < .05$

### Table 3
**Effect Sizes (ESs) by Outcome Category reported by Cason and Gillis (1994)**

<table>
<thead>
<tr>
<th>Category</th>
<th>Program Effect ES ($N$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral assessment by others</td>
<td>.40 (23)</td>
</tr>
<tr>
<td>Self-concept</td>
<td>.34 (23)</td>
</tr>
<tr>
<td>Academic</td>
<td>.61 (10)</td>
</tr>
<tr>
<td>Attitude surveys</td>
<td>.46 (19)</td>
</tr>
<tr>
<td>Locus of control</td>
<td>.30 (13)</td>
</tr>
<tr>
<td>Clinical scales</td>
<td>1.05 (12)</td>
</tr>
<tr>
<td>School attendance</td>
<td>.47 (9)</td>
</tr>
</tbody>
</table>