Journal of Experiential Education • 2008, Volume 31, No. 2 pp. 111–135

# Are Challenge (Ropes) Courses an Effective Tool? A Meta-Analysis

*H. Lee Gillis and Elizabeth Speelman*

*This study reports the results of a meta-analysis of 44 studies that examined the impacts of participation in challenge (ropes) course activities. Overall, a medium standardized mean difference effect size was found (*d = *0.43). Effect sizes were calculated for various study characteristics, including demographics and outcome. Higher effects were found for adult groups (*d = *0.80) and for studies measuring family functioning (*d = *0.67). Studies with therapeutic (*d = *0.53) or developmental foci (*d = *0.47) had higher effect sizes than those with educational foci (*d = *0.17). Higher effect sizes for group effectiveness (*d = *0.62) affirmed the use of challenge course experiences for team-building purposes. Implications for further research include the importance of recording detailed program design information, selecting appropriate instrumentation, and including follow-up data.*

Keywords: Challenge Course, Adventure Programming, Meta-Analysis, Outcomes, Evidence-Based

*H. Lee Gillis is a Professor in the Department of Psychology at Georgia College & State University, USA. E-mail:* lee.gillis@gcsu.edu

*Elizabeth Speelman is the Assistant Director of Program Development for the Youth Learning Institute at Clemson University, USA. E-Mail:* lizspeelman@gmail.com

Over the years there have been many requests from a variety of people and organizations for "the" study that "proves" challenge (ropes) courses are effective. Google searches for *challenge course* or *ropes course* revealed 398,000 and 837,000 entries respectively (retrieved May 31, 2007). Scanning these entries reveals numerous hits on *who* is using challenge courses, *what* challenge courses are, *where* they are used or being built, *how* they are being funded, and *why* they are beneficial. While much anecdotal evidence about the impact of challenge course experiences can be found in the form of testimonials (Martin, Cashel, Wagstaff, & Breunig, 2006), there is little statistical evidence published in peer-reviewed journals that can support the claim that challenge courses are effective over multiple settings with various client groups. The purpose of this paper is to report the results of a meta-analysis of 44 studies that examined the impacts of participation in challenge (ropes) course activities.

## Background

Rohnke, Rogers, Tait, and Wall (2007) trace the origin of challenge (ropes) courses (hereafter referred to as challenge courses) to 1941 and the use of obstacle courses by the military. Martin, Cashel, Wagstaff, and Breunig (2006) cite evidence of the British Outward Bound School using challenge courses in the 1940s and speculate that the courses were built to simulate the work sailors would do on riggings and climbing masts while working on large sailing ships. Challenge courses were introduced at the Colorado Outward Bound School in the 1960s.

Challenge courses have been categorized as low courses or high courses. Priest and Gass (2005) have defined low challenge course activities as requiring spotting and high challenge course activities as those requiring belaying. More specifically, low challenge course activities generally focus on group problem solving and team building and range in height from literally sitting on the ground to a height of 12 to 13 feet (Rohnke et al., 2007). Sometimes authors (Martin et al., 2006) have designated games, icebreakers, group initiatives, and trust activities as a category separate from low challenge course activities—especially when low elements are permanently constructed.

High challenge course activities have provided physical and mental challenges predominantly for individual development (Rohnke et al., 2007) and most often have involved safety systems similar to those used in rock climbing. Individual elements using a dynamic belay system have a

belayer who stands on the ground and manages a climber's rope, which is directed through a top anchor point attached to a cable above the participant. Other high challenge courses use a static belay system where participants are self-belayed using movable lanyards as they move from element to element, often not returning to the ground until they have finished a circuit (Ryan, 2005). The Association for Challenge Course Technology (ACCT), a professional trade organization that began in 1993, writes standards for constructing and managing challenge courses (ACCT, 2004).

Today challenge courses are used in recreational, educational, developmental, and therapeutic settings, as well as in camps, hospitals, and corporate training centers (Rohnke et al., 2007). Therapeutic uses of challenge courses date back to hospital settings in the early 1980s (Prouty, 1999). Schoel, Prouty, and Radcliffe (1988) have also documented therapeutic programming in schools and residential settings, as has Gass (1993, 1995).

Numerous authors (Rohnke et al., 2007, Martin et al., 2006) speak to the lack of research on challenge courses. Prouty (1999) and Rohnke (1977) documented an initial evaluation of Project Adventure's adaptation of Outward Bound principles into a high-school setting in the early 1970s. Positive changes were found in increased self-esteem and more internalized locus of control. The constructs used in that study have dominated research and evaluation in the field ever since (Gillis & Thompson, 1996). While bibliographies have been compiled (Attarian & Holden, 2005), to date there has not been a systematic and statistical review of research conducted on challenge courses. There have, however, been several meta-analyses conducted on outdoor and adventure education that are pertinent to this review.

## Review of Literature

### *Brief Overview of Meta-Analyses*

A meta-analysis is a statistical compilation of quantitative research. Lipsey and Wilson (2001) describe it as research in which studies rather than people are surveyed. Moher and others (1999) indicate that the strengths of a meta-analysis include: its potential as a systematic review with reduced bias; a foundation for evidence-based practice; and a guide for future research. Its limitations include incorrect use of statistical evidence caused by combining dissimilar studies and using studies with weak methodology or insufficient detail.

To compile a meta-analysis, research is coded into a standard format similar to interviewing individuals in a survey. Each study must contain the appropriate quantitative information to qualify for inclusion in the meta-analysis. The data compiled during the coding process are then computed, adjusted, and analyzed for effect sizes.

Effect sizes provide a means of statistical standardization for the meta-analysis by providing a comparable numeric value for interpretation of trends (Lipsey & Wilson, 2001). Effect sizes describe the relationship between two variables. According to the National Information Center on Health Services Research and Health Care Technology, an effect size is described as "a dimensionless measure of treatment effect" (2007). In the present study, effect size describes the relationship between outcome measures (e.g., self-efficacy or team effectiveness) of participants on a challenge course with those who did not participate. In order to interpret standardized mean difference effect sizes, Cohen (1988) defined widely accepted magnitudes for effect sizes as follows: small effect sizes are less than 0.20; medium effect sizes are 0.50; and large effect sizes are greater than 0.80. Wolf (1986) theorized that an effect size of 0.25 has educational significance while an effect size of 0.50 has practical or clinical significance (see authors' note).

### *Previous Meta-Analyses in Adventure Programming*

Neill (2003) summarized the data from meta-analyses related to outcomes in outdoor education, education, and psychotherapy and determined benchmarks for adventure therapy program outcomes. He identified three meta-analyses published in peer-reviewed journals that focused on outdoor or adventure education outcomes. Cason and Gillis (1993) compiled studies in adventure programming specifically for adolescents, reporting an overall nonstandardized mean difference effect size of 0.31. In a meta-analysis on adventure education and Outward Bound programs, Hattie, Marsh, Neill, and Richards (1997) calculated an overall standardized mean difference effect size of 0.34. Hans (2000) compiled studies for a meta-analysis on the specific outcome of locus of control in adventure programming that resulted in a nonstandardized mean difference effect size of 0.38. A fourth meta-analysis was published in conference proceedings and was conducted over a 10-year period by Bunting and Donley (2002) (as cited in Neill, 2003). Neill reported the findings of this study, focused specifically on challenge course outcomes, had an overall nonstandardized mean difference effect size of 0.55 from a sample of 15 studies.

The current study attempts to answer the following questions: Are challenge (ropes) courses an effective tool and, if so, for which populations and under what conditions are challenge courses most effective?

## Method

Selection criteria for the studies in this meta-analysis were based on information requested from challenge course practitioners (Gillis & Speelman, 2007). The eligibility criteria were as follows:

1. The study focused solely on challenge course activity outcomes. Studies in which participants were involved in *any* other activity (e.g., swimming, backpacking, etc.) during the period of study were excluded. As a result, several studies completed at camps were excluded due to the confounding factors created by other camp activities. One study that specifically compared campers who participated in challenge course activities with those who did not participate was included (Greene, 1992).
2. The study was conducted between the years 1986 and 2006. A 20-year period allowed trends in research to be followed and consistency in challenge course technology (Rohnke et al., 2007; Ryan, 2005) to be preserved, while keeping study information current.
3. Unlike earlier meta-analyses in adventure programming, the study had to include control groups. Control groups received no challenge course programming, were on a waiting list for programming at a future time, or received alternative programming. Single group pretest and posttest design studies were not included in this meta-analysis, since by definition these studies did not have control groups (Cook & Campbell, 1979).
4. Outcome measures had to be recorded quantitatively and with sufficient data to compute effect sizes. Studies in which the means were calculated at pretest and posttest had to include standard deviations to calculate effect sizes. Quantitative studies were included regardless of statistical significance. No purely qualitative studies were used in this analysis, nor were any qualitative data transformed to quantitative data.
5. The studies had to be accessible to the authors. As such, all studies used were available in English.

### *Identification and Retrieval of Eligible Studies*

Four sources were used to locate all relevant and eligible studies. First, an extensive search was conducted using two bibliographic databases: (a) EBSCOhost (www.ebscohost.com), which includes Academic Search Premier, ERIC, PsycINFO, and Psychology & Behavioral Sciences Collection; and (b) Dissertation Abstracts/Digital Dissertations at ProQuest (www.proquest.com). Second, an existing bibliography on challenge course-related articles (Attarian & Holden, 2005) was reviewed. Third, as studies were collected, their references were reviewed. Finally, at a presentation during the annual conference of the ACCT (Gillis & Speelman, 2007), participants were asked to submit sources that might have been overlooked. A total of 137 potential studies were identified.

An attempt was made to retrieve complete copies of all 137 studies. Studies were obtained through university interlibrary loan service, directly from the author, or downloaded from the Internet. Five studies (Bocarro, 1998; Harper, 2006; Knott, 2004; Terry, 2002; and Wick, 1997) were inaccessible through the means described.

***Coding of Eligible Studies***

A total of 44 studies matched the criteria for inclusion in this meta-analysis. All of the studies were conducted in the United States. The 44 studies were coded using a coding manual modified from the one described in Lipsey and Wilson (2001). As their coding manual was created for a meta-analysis involving outcomes on wilderness challenge programs for delinquent youth (Wilson & Lipsey, 2000), only slight changes were needed to make it applicable to the specific outcomes and populations of this study. Each study was given an identification number and coded based on demographic and effect size information. Each study was reviewed at least three times to ensure accuracy of information.

*Study descriptors.* Information regarding the date of the research, the population studied, the activities conducted, the purpose of the activities, and whether the research was published was coded for each study. Populations were coded according to age, race, gender, and type of group (e.g., school, therapeutic, etc.). Challenge course activities were coded as low activities only or a combination of low and high activities (as previously described), and by duration (hours). The focus of the activities was coded according to the intended purpose of the program that was being researched: recreational; educational; developmental; or therapeutic (Priest, Gass, & Gillis, 2000/2003). There were no programs in the sample whose intentional focus was solely recreational. Educational programs were intended to improve the overall performance of students at school. Developmental programs were designed to improve behaviors of the participants. The intention of therapeutic programs was to change patterns of behavior of the participants. Effect sizes were also coded according to whether they were calculated from posttest or follow-up results.

*Outcome measures.* Each study was coded by the specific outcome measures being researched. Outcome measures were categorized into 10 groups: self-esteem or self-concept; self-efficacy; personality measures (e.g., mood, personality characteristics); behavioral observations; academic measures (e.g., grade point average, attendance); atmosphere; attitudes about physical condition (e.g., diabetes, sports); family; physical variables (e.g., weight, body mass index); and group dynamics (e.g., interpersonal, cohesion, group effectiveness).

### *Effect Size*

*Calculation.* Effect sizes were calculated using an effect size computation program (Wilson, 2001). These effect sizes were calculated using both mean and standard deviation, *F*-value or *t*-value statistics provided in the study. The resulting effect size was a mean difference (Cohen's *d*). For the purpose of this study, reported effect sizes represent the difference, positive or negative, between results on outcome measures for participants in the challenge course group and the control group. A positive result indicates the challenge course group performed better than the control group; a negative result indicates the opposite.

*Adjustment.* Studies varied with regards to sample size. Because a small sample size could create less precision, Lipsey and Wilson (2001) recommend weighting procedures to increase validity. The weighting procedures for this study included calculating standard errors and computing the inverse variance weights for every effect size. This adjustment created standardized mean difference effect sizes.

*Analysis.* Lipsey and Wilson (2001) also recommend calculating one effect size per outcome. They provide mathematical procedures for obtaining standardized mean difference effect sizes (MEANES.SPS) and performing analysis of variance (METAF.SPS) for each outcome. Using the provided SPSS macros, the current study followed these recommendations.

*Interpretation.* Lipsey and Wilson (2001) note that the interpretation of effect sizes does not account for the situation in which the results were derived (e.g., facilitator competency or environmental variables such as weather). There may be circumstances where a small effect size change is highly meaningful due to the nature of the situation.

## Results

A combined total of 2,796 participants were involved in the studies included in this meta-analysis either as challenge course participants ($n$ = 1,417) or as members of a control group ($n$ = 1,379). Standardized mean difference effect size and additional information are presented for each of the 44 studies including general characteristics, specific outcomes, and posttest studies compared with follow-up studies. All effect sizes reported in this section are standardized mean difference effect sizes.

### *Standardized Mean Difference Effect Size of Each Study*

Table 1 reports the standardized mean difference effect size, number of effect sizes, mean age, number of participants, and duration (hours) for each of the 44 studies. An overall effect size of 0.43 was calculated from the 44 studies. The highest effect size was 2.83 (Bisson, 1998) in a study

**Table 1**
***Mean Effect Size* (d), *Number of Effect Sizes, Age, Number of Participants, and Duration in Hours of Contributing Studies***

| Study | Mean *d* | Number | Age | *n* (Exp) | *n* (Cnt) | Hours |
|---|---|---|---|---|---|---|
| Aghazarian, T. L. (1997) | 0.02 | 18 | 16.5 | 46 | 49 | 8.0 |
| Bisson, C. (1998) | 2.83 | 8 | 11.0 | 54 | 56 | |
| Blanchard, C. W. (1994) | 0.53 | 5 | 14.8 | 20 | 20 | 27.0 |
| Boyle, S. (1986) | 0.02 | 10 | 24.7 | 30 | 30 | 7.0 |
| Breheny, M. S. (2000) | 0.11 | 24 | 19.2 | 21 | 18 | 4.0 |
| Bronson, J., Gibson, S., Kishar, R., & Priest, S. (1992) | 0.68 | 10 | | 17 | 11 | 24.0 |
| Daheim, T. J. (1999) | 0.10 | 6 | 12.7 | 110 | 86 | 8.0 |
| Eagle, H. A. (2000) | 0.13 | 8 | 13.0 | 74 | 64 | 37.0 |
| Faubel, G. (1998) | 0.74 | 6 | 13.0 | 41 | 27 | |
| Faulkner, S. S. (2002) | 0.89 | 6 | | 54 | 66 | 3.0 |
| Finkenburg, M. E., Shows, D., & DiNucci, J. M. (1994) | 0.62 | 20 | 21.3 | 18 | 32 | |
| Greene, S. S. (1992) | -0.09 | 2 | 15.3 | 4 | 4 | 16.0 |
| Hatch, K. D. (2006) | 0.24 | 8 | 21.1 | 48 | 22 | 4.0 |
| Horak, R. A., Jr. (2003) | 0.10 | 4 | 11.0 | 51 | 54 | |
| Hughes, J. S. (1994) | 1.20 | 1 | | 30 | 30 | |
| Jacobson, S. E. (1993) | 0.72 | 6 | | 16 | 4 | 23.5 |
| Jelalian, E., Mehlenbeck, R., Lloyd-Richardson, E. E., Birmaher, V., & Wing, R. R. (2006) | 0.26 | 11 | 14.5 | 37 | 39 | 16.0 |
| Kanters, M. A., Bristol, D. G., & Attarian, A. (2002) | 0.25 | 12 | | 57 | 78 | |
| Kelly, S. A. (1996) | -0.13 | 33 | | 34 | 12 | 16.0 |
| McDaniel, P. E. (1999) | 0.37 | 6 | 16.5 | 22 | 20 | 6.0 |
| McDonald, R.G., & Howe, C.Z. (1989) | 0.52 | 6 | | 18 | 20 | 28.0 |
| McGarvey, A. L. (2004) | -0.04 | 44 | 13.0 | 29 | 15 | |
| Meyer, B. B. (2000) | 0.53 | 7 | 16.0 | 16 | 19 | 8.0 |
| Miller, D. J. (1998) | 0.26 | 2 | | 59 | 68 | 3.0 |
| Moorefield, D. L. (1994) | 0.37 | 5 | | 16 | 25 | 4.0 |
| Moreau de la Meuse, N. (1993) | 0.04 | 5 | 12.0 | 78 | 63 | 20.0 |

**Table 1 (continued)**

| Study | Mean *d* | Number | Age | *n* (Exp) | *n* (Cnt) | Hours |
|---|---|---|---|---|---|---|
| Munns, K. M. (1996) | 1.02 | 18 | 43.0 | 24 | 53 | 24.0 |
| O'Bannon, P. M. (2000) | 1.35 | 7 | | 37 | 31 | 16.0 |
| Parker, M. W. (1993) | 0.10 | 20 | 15.0 | 37 | 47 | 16.0 |
| Pazmino-Cevallos, M. (2004) | 0.85 | 2 | | 32 | 29 | 8.0 |
| Shoemaker, J. A. (1997) | 0.00 | 1 | 22.2 | 26 | 11 | |
| Smith, S. R. (1994) | 0.31 | 9 | 21.2 | 20 | 21 | |
| Sturdivant, V. A. (1991) | 0.17 | 20 | | 14 | 12 | 8.0 |
| Sung, Y. K. (2005) | 0.15 | 2 | | 56 | 56 | 8.0 |
| Talbot, P. A. (2001) | -0.10 | 2 | | 30 | 26 | |
| Thompson-Grim, L. J. (1999) | 0.18 | 2 | 15.6 | 17 | 17 | 48.2 |
| Ulrich, J. S. (1993) | 0.31 | 3 | | 7 | 13 | 16.0 |
| Vasquez, L. M. (2001) | 1.35 | 2 | 14.5 | 10 | 11 | 25.0 |
| Voight, A. (1988) | 0.54 | 8 | 15.0 | 11 | 10 | 5.0 |
| White, F. E., Jr. (1998) | 0.62 | 4 | 16.1 | 12 | 12 | 45.0 |
| Wisnyai, M. (1989) | 0.82 | 4 | | 27 | 34 | |
| Witman, J. (1987) | 1.04 | 1 | 17.0 | 5 | 5 | 5.0 |
| Witter, B. E. (2005) | -0.24 | 3 | 20.2 | 16 | 11 | 3.0 |
| Ziven, H. S. (1989) | 0.09 | 9 | 15.3 | 36 | 48 | |

of the impact of sequencing of challenge course activities on group cohesion, and the lowest effect size was -0.24 (Witter, 2005) in a study of the impact of challenge course initiatives on the self-efficacy of college students. The effect size of 2.83 (Bisson) indicates the positive gain between the challenge course group over the control group. The negative effect size ($d$ = -0.24, Witter) indicates that the challenge course group scored lower than the control group.

Calculating an effect size without Bisson's study reduced the overall effect by 0.04 to $d$ = 0.39. Although Bisson's study appears to be an outlier, the study was included in calculating the effect sizes in this meta-analysis.

***General Characteristics***

*Number of studies.* Table 2 reports the number of studies, the percentage of total studies, the standardized mean difference effect size, the number of effect sizes, and percentage of the total effect sizes for each characteristic. A majority of the studies were dissertations ($n$ = 36; 81.8%). The highest number of studies by age included high-school participants ($n$ = 15, 34.1%). The intended focus for half of the studies was developmental ($n$ = 22; 50%). With regard to time (duration in hours) on the challenge course, there were eight studies (18.2%) where challenge course activities were conducted for five hours or less, eight studies (18.2%) where the activities were conducted for 6–10 hours, seven (15.9%) conducted for 11–20 hours, six (13.6%) conducted for 21–30 hours, and three (6.8%) conducted for more than 31 hours. The time category represented total hours on task, regardless of whether activities were conducted continuously over several days or over an extended period of time. When reported, hours did not separate time on task into categories that allowed for standardized coding. In fact, 12 (27.3%) of the studies did not report time on task at all.

Half of the studies were conducted with a combination of low and high challenge course activities ($n$ = 22, 50%). However, 10 (22.7%) of the studies did not indicate the type of challenge course activities that were used in the research. The control group, for more than half of the studies ($n$ = 25; 56.8%), either received no challenge course experience or was put on a waiting list to receive the challenge course experience following the collection of data for the study.

*Effect size.* A total of 390 effect sizes were obtained from the 44 studies and recorded in Table 2. These effect sizes were calculated *without* regard to the outcome measure. For example, effect sizes calculated for middle-school students did not differentiate between measures of self-efficacy and measures of grade point average; only one value was reported for *all* effect sizes involving middle-school students.

The 75 (19.2%) effect sizes calculated from studies published as journal articles produced a higher mean effect size ($d$ = 0.49) than the

**Table 2**
***General Characteristics and Effect Sizes of Studies Included in the Meta-Analysis***

| General Characteristics | Number of Studies | Percentage | Mean *d* | Number of *d* | Percentage |
|---|---|---|---|---|---|
| **Type of publication** | | | | | |
| Journal Article | 8 | 18.2 | 0.49 | 75 | 19.2 |
| Thesis or Dissertation | 36 | 81.8 | 0.42 | 315 | 80.8 |
| **Age** | | | | | |
| Middle-School Age | 8 | 18.2 | 0.46 | 87 | 22.3 |
| High-School Age | 15 | 34.1 | 0.38 | 102 | 26.2 |
| University Age | 11 | 25.0 | 0.18 | 111 | 28.5 |
| Other Adults | 7 | 15.9 | 0.80 | 76 | 19.5 |
| Family | 3 | 6.8 | 0.67 | 14 | 3.6 |
| **Focus** | | | | | |
| Educational | 7 | 15.9 | 0.17 | 39 | 10.0 |
| Developmental | 22 | 50.0 | 0.47 | 218 | 55.9 |
| Therapeutic | 15 | 34.1 | 0.53 | 133 | 34.1 |
| **Activity** | | | | | |
| Only Low Challenge Course | 12 | 27.3 | 0.32 | 82 | 21.0 |
| Low and High Challenge Course | 22 | 50.0 | 0.54 | 202 | 51.8 |
| Could Not Determine | 10 | 22.7 | 0.35 | 106 | 27.2 |
| **Number of hours** | | | | | |
| < 5 hours | 8 | 18.2 | 0.37 | 57 | 14.6 |
| 6–10 hours | 8 | 18.2 | 0.28 | 71 | 18.2 |
| 11–20 hours | 7 | 15.9 | 0.32 | 81 | 20.8 |
| 21–30 hours | 6 | 13.6 | 0.79 | 47 | 12.1 |
| > 31 hours | 3 | 6.8 | 0.24 | 14 | 3.6 |
| Was Not Reported | 12 | 27.3 | 0.52 | 120 | 30.8 |
| **Controlled** | | | | | |
| No Challenge Course Experience or "On a Waiting List" | 25 | 56.8 | 0.41 | 241 | 61.8 |
| Received Regular or Alternative Programming | 18 | 40.9 | 0.48 | 141 | 36.2 |
| Could Not Determine | 1 | 2.3 | 0.24 | 8 | 2.1 |

315 (80.8%) effect sizes that were calculated from theses and dissertations ($d$ = 0.42). This finding is different from Cason and Gillis (1993), who found higher effect sizes for dissertations and theses. Both types of publications produced medium (Cohen, 1988) and educationally significant (Wolf, 1986) effect sizes.

**Table 3**
***Mean Effect Size, Number of Effect Sizes, and Percentage of Total Effect Sizes For Each Age by Focus, Activity, and Duration***

| | Number of Studies | Percentage | Mean *d* | Number of *d* | Percentage |
|---|---|---|---|---|---|
| **Middle School** | | | | | |
| Focus | | | | | |
| Educational | 3 | 37.5 | 0.11 | 18 | 20.7 |
| Developmental | 2 | 25.0 | 1.33 | 13 | 14.9 |
| Therapeutic | 3 | 37.5 | 0.41 | 56 | 64.4 |
| Activity | | | | | |
| Low Challenge Course Only | 1 | 12.5 | 0.52 | 6 | 6.9 |
| Low and High Challenge Course | 6 | 75.0 | 0.61 | 76 | 87.4 |
| Could Not Determine | 1 | 12.5 | 0.04 | 5 | 5.7 |
| Number of hours | | | | | |
| < 5 hours | 0.0 | 0.0 | | 0.0 | 0.0 |
| 6–10 hours | 1 | 12.5 | 0.10 | 6 | 6.9 |
| 11–20 hours | 1 | 12.5 | 0.04 | 5 | 5.7 |
| 21–30 hours | 1 | 12.5 | 0.52 | 6 | 6.9 |
| > 31 hours | 1 | 12.5 | 0.13 | 8 | 9.2 |
| Was Not Reported | 4 | 50.0 | | 62 | 71.3 |
| **High School** | | | | | |
| Focus | | | | | |
| Educational | 3 | 20.0 | 0.15 | 9 | 8.8 |
| Developmental | 3 | 20.0 | 0.41 | 29 | 28.4 |
| Therapeutic | 9 | 60.0 | 0.36 | 64 | 62.7 |
| Activity | | | | | |
| Only Low Challenge Course | 4 | 26.7 | 0.14 | 18 | 17.6 |
| Low and High Challenge Course | 8 | 53.3 | 0.48 | 42 | 41.2 |
| Could Not Determine | 3 | 20.0 | 0.28 | 42 | 41.2 |
| Number of hours | | | | | |
| < 5 hours | 2 | 13.3 | 0.69 | 9 | 10.3 |
| 6–10 hours | 3 | 20.0 | 0.30 | 31 | 35.6 |
| 11–20 hours | 3 | 20.0 | 0.23 | 34 | 39.1 |
| 21–30 hours | 2 | 13.3 | 0.78 | 7 | 8.0 |
| > 31 hours | 2 | 13.3 | 0.38 | 6 | 6.9 |
| Was Not Reported | 3 | 20.0 | | 0 | 0.0 |
| **University** | | | | | |
| Focus | | | | | |
| Educational | 1 | 9.1 | 0.25 | 12 | 10.8 |
| Developmental | 10 | 90.9 | 0.19 | 99 | 89.2 |
| Therapeutic | 0 | 0.0 | | 0 | 0.0 |
| Activity | | | | | |
| Only Low Challenge Course | 1 | 9.1 | 0.25 | 37 | 33.3 |
| Low and High Challenge Course | 10 | 90.9 | 0.19 | 22 | 19.8 |
| Could Not Determine | 0 | 0.0 | | 52 | 46.8 |

**Table 3 (continued)**

| | Number of Studies | Percentage | Mean *d* | Number of *d* | Percentage |
|---|---|---|---|---|---|
| **Number of hours** | | | | | |
| < 5 hours | 4 | 36.4 | 0.19 | 37 | 33.3 |
| 6–10 hours | 3 | 27.3 | 0.11 | 32 | 28.8 |
| 11–20 hours | 0 | 0.0 | | 0 | 0.0 |
| 21–30 hours | 0 | 0.0 | | 0 | 0.0 |
| > 31 higher | 0 | 0.0 | | 0 | 0.0 |
| Was Not Reported | 4 | 36.4 | | 42 | 37.8 |
| **Other Adult** | | | | | |
| Focus | | | | | |
| Educational | 0 | 0.0 | | 0 | 0.0 |
| Developmental | 6 | 85.7 | 0.72 | 75 | 98.7 |
| Therapeutic | 1 | 14.3 | 1.20 | 1 | 1.3 |
| Activity | | | | | |
| Only Low Challenge Course | 2 | 28.6 | 0.51 | 15 | 19.7 |
| Low and High Challenge Course | 4 | 57.1 | 0.77 | 54 | 71.1 |
| Could Not Determine | 1 | 14.3 | 1.35 | 7 | 9.2 |
| Number of hours | | | | | |
| < 5 hours | 1 | 14.3 | 0.37 | 5 | 6.6 |
| 6–10 hours | 1 | 14.3 | 0.85 | 2 | 2.6 |
| 11–20 hours | 2 | 28.6 | 0.64 | 40 | 52.6 |
| 21–30 hours | 2 | 28.6 | 0.87 | 28 | 36.8 |
| > 31 hours | 0 | 0.0 | | 0 | 0.0 |
| Was Not Reported | 1 | 14.3 | | 1 | 1.3 |
| **Family** | | | | | |
| Focus | | | | | |
| Educational | 0 | 0.0 | | 0 | 0.0 |
| Developmental | 1 | 33.3 | -0.09 | 2 | 14.3 |
| Therapeutic | 2 | 66.7 | 0.87 | 12 | 85.7 |
| Activity | | | | | |
| Only Low Challenge Course | 1 | 33.3 | 0.89 | 6 | 42.9 |
| Low and High Challenge Course | 2 | 66.7 | 0.33 | 8 | 57.1 |
| Could Not Determine | 0 | 0.0 | | 0 | 0.0 |
| Number of hours | | | | | |
| < 5 hours | 1 | 33.3 | 0.89 | 6 | 42.8 |
| 6–10 hours | 0 | 0.0 | | 0 | 0.0 |
| 11–20 hours | 1 | 33.3 | -0.09 | 2 | 14.3 |
| 21–30 hours | 1 | 33.3 | 0.72 | 6 | 42.9 |
| > 31 hours | 0 | 0.0 | | 0 | 0.0 |

Most of the effect sizes ($n$ = 111, 28.5%) categorized by age group were calculated from studies that involved university-age participants. The lowest number of effect sizes ($n$ = 14, 3.6%) was calculated from studies involving families. There were more effect sizes for high school age participants ($n$ = 102, 26.2%) than for middle school age participants ($n$ = 87, 22.3%). Studies involving adults who were not involved through a college or university provided 76 effect sizes (19.5%). The adult group had the highest effect size ($d$ = 0.80), followed by family groups ($d$ = 0.67). Both of these values were large (Cohen, 1988) and practically significant (Wolf, 1986). Middle school ($d$ = 0.46) and high school ($d$ = 0.38) age groups produced medium (Cohen) and educationally significant (Wolf) effect sizes. The lowest effect size ($d$ = 0.18) represented the outcomes of university age participants.

In terms of program focus, the largest number of effect sizes was calculated from studies of programs with a developmental purpose ($n$ = 218, 55.9%) and produced a medium (Cohen, 1988) effect size ($d$ = 0.47). This was followed by programs with a therapeutic focus ($n$ = 133, 34.1%) with a practical or clinically significant (Wolf, 1986) effect size ($d$ = 0.53). The smallest number of effect sizes was related to educational programs ($n$ = 39, 10.0%) and showed only a small effect ($d$ = 0.17) (Cohen).

A total of 202 (51.8%) effect sizes were calculated from studies primarily involving a combination of low and high challenge course activities. Low and high challenge course activities demonstrated a higher effect size ($d$ = 0.54) than only low challenge course activities ($d$ = 0.32). The low and high challenge course effect size is practically significant while the low challenge course effect size is educationally significant (Wolf, 1986). There were 106 effect sizes (27.2%) from 10 studies that did not indicate the type of challenge course activity. These studies could change the mean effect sizes for either category of activity.

The largest effect sizes ($d$ = 0.79) (Cohen, 1988) for time occurred for challenge courses conducted between 21–30 hours in duration ($n$ = 47, 12.1%) and produced practical or clinically significant effects (Wolf, 1986). Effect sizes for less than five hours ($n$ = 57, 14.6%), 6–10 hours ($n$ = 71, 18.2%), 11–20 hours ($n$ = 81, 20.8%), and more than 31 hours ($n$ = 14, 3.6%) all had small effect sizes (Cohen) ranging from 0.24 to 0.37, thus producing effects that would be categorized as educationally significant (Wolf).

Effect sizes ($n$ = 241, 61.8%) were calculated from studies involving control groups who either received no challenge course experience or experienced the challenge course program after the study had been completed (e.g., were on a waiting list). One study, although mentioning that a control group was used, did not record the nature of the control group's activities (Hatch, 2006). The highest effect size for the control groups was for those who received regular or alternative programming without any

challenge course experience ($d$ = 0.48). The effect size for the groups who either received no challenge course experience or experienced the challenge course after the study was completed is a medium effect size ($d$ = 0.41) (Cohen). There was a greater positive difference between the challenge course group and control group in the studies where control groups received regular or alternative programming.

***Matrices of Age by Focus, Activity, and Time***

Practitioners who have seen presentations of preliminary data from this meta-analysis have repeatedly and consistently asked for details of specific demographics by age group. The authors have been reluctant to provide this level of detail since many of the categories contain only one study creating a potential for misunderstanding. With this caution in mind, Table 3 reports the number of studies, standardized mean difference effect sizes and number of effect sizes for the program's focus, the type of activity, and the time spent on the activity for each age group.

*Middle school.* The highest effect size was recorded for programs with a developmental focus ($d$ = 1.33). Most studies occurred on challenge courses that presented a combination of low and high activities ($n$ = 6) although effect sizes were similar for only low activities and the combination of high and low. Each time period had only one study and four studies did not report time length. One study in the time period of 21–30 hours had an effect size of 0.52.

*High school.* Developmental programs had the highest effect size ($d$ = 0.41) while therapeutic programs had the largest number of studies ($n$ = 9). Highest effect size ($d$ = 0.48) and the most studies ($n$ = 8) were recorded for a combination of low and high activities. Time periods of less than five hours and 21–30 hours, with two studies each, had the highest effect sizes (0.69 and 0.78 respectively).

*University.* There were no therapeutic programs in this age group. They were predominately developmental ($n$ = 10) but with a low effect size ($d$ = 0.19). Ten of the eleven studies used a combination of low and high activities with the same low effect size ($d$ = 0.19). No reported study was longer than 10 hours and effect sizes were low.

*Other adults.* This category accounts for the highest effect sizes across all three categories. Six of the studies in this age group were developmental with an effect size of 0.72; the one therapeutic study had an effect size of 1.20. Four studies using a combination of low and high challenge course activities had an effect size of 0.77; the one study for which the type of activity could not be determined had an effect size of 1.35. The two highest effect sizes occurred in the 6–10 hour timeframe ($n$ = 1; $d$ = 0.85) and in the 21–30 hour timeframe ($n$ = 2; $d$ = 0.87).

*Families.* There were only three studies in this category. The one developmental study had a negative effect size, while the two therapeutic studies recorded an effect size of 0.87. The one low challenge course study recorded an effect size of 0.89. The two highest effect sizes occurred in the less-than-five-hour timeframe ($n$ = 1; $d$ = 0.89) and in the 21–30 hour timeframe ($n$ = 1; $d$ = 0.72); the 11–20 hour timeframe had a negative effect size indicating the control group scored higher than the challenge course group.

### *Outcome Measures*

*Number of studies.* Table 4 indicates that the four most frequent outcomes used in challenge course studies were self-esteem or self-concept ($n$ = 20, 45.5%), group dynamics ($n$ = 14, 31.8%), personality measures ($n$ = 10, 22.7%), and self-efficacy ($n$ = 7, 15.9%). A small number of studies measured behavioral observations ($n$ = 4, 9.1%), classroom environment ($n$ = 4, 9.1%), academic measures ($n$ = 3, 6.8%), attitudes about physical condition ($n$ = 3, 6.8%), family ($n$ = 2, 4.6%), and physical variables ($n$ = 1, 2.3%).

**Table 4**
***Mean effect size, number of effect sizes, and percentage of total effect sizes for each outcome construct***

| Outcome | Number | Percentage | Mean *d* | Number of *d* | Percentage |
|---|---|---|---|---|---|
| Self-Esteem or Self-Concept | 20 | 29.4 | 0.26 | 125 | 32.1 |
| Self-Efficacy | 7 | 10.3 | 0.48 | 34 | 8.7 |
| Personality Measures | 10 | 14.7 | 0.29 | 45 | 11.5 |
| Behavioral Observations | 4 | 5.9 | 0.37 | 42 | 10.8 |
| Academic Measures | 3 | 4.4 | 0.26 | 8 | 2.1 |
| Environmental (atmosphere) | 4 | 5.9 | 0.01 | 21 | 5.4 |
| Attitudes about Physical Condition | 3 | 4.4 | 0.52 | 11 | 2.8 |
| Family | 2 | 2.9 | 0.86 | 10 | 2.6 |
| Physical Variable (e.g., weight, BMI, blood) | 1 | 1.5 | 0.00 | 4 | 1.0 |
| Group Dynamics (Interpersonal, Cohesion, Effectiveness) | 14 | 20.6 | 0.62 | 90 | 23.1 |

*Effect size.* The highest effect size for outcomes was calculated from studies based on family measures ($d$ = 0.86). This is a large effect size (Cohen, 1988) with practical significance (Wolf, 1986). Outcome measures that exhibited a medium effect size (Cohen) with practical significance

(Wolf) were group dynamics ($d$ = 0.62) and attitudes about physical condition ($d$ = 0.52). Medium effect sizes (Cohen) with educational significance (Wolf) were reported for self-efficacy ($d$ = 0.48), behavioral observations ($d$ = 0.37), personality measures ($d$ = 0.29), self-esteem or self-concept ($d$ = 0.26), and academic measures ($d$ = 0.26). Outcome measures related to classroom environment were small ($d$ = 0.01). Physical variables that were studied resulted in a standardized mean difference effect size of 0.00.

***Follow-Up Effects***

*Number of studies.* Table 5 reports the number of studies, percentage of total studies, standardized mean difference effect sizes, number of effect sizes, and percentage of total effect sizes for both posttest studies and follow-up studies. From the 44 studies, one study did not involve a posttest outcome. The research for this study was focused on long-term outcomes and only the follow-up data was recorded (Eagle, 2000). Only 12 (27.3%) studies contained follow-up data.

*Effect size.* The effect size for posttest outcomes is 0.45, a medium effect (Cohen, 1988) that is educationally significant (Wolf, 1986). The effect size for outcomes in follow-up tests resulted in a small effect size ($d$ = 0.23).

**Table 5**
***Mean effect size, number of effect sizes, and percentage of total effect sizes for posttest and follow-up studies***

| Study Type | Number | Percentage | Mean *d* | Number of *d* | Percentage |
|---|---|---|---|---|---|
| Posttest | 43 | 78.2 | 0.45 | 309 | 79.2 |
| Follow-up | 12 | 21.8 | 0.23 | 81 | 20.8 |

## Discussion

Challenge courses are an effective tool for impacting a variety of educational and psychological constructs with a variety of participants. These effects are small to medium in magnitude. Because only controlled studies were included in this analysis, however, these effects have practical significance (Wolf, 1986) for participants in a challenge course experience when compared to those who were assessed on similar constructs but did not have a challenge course experience.

Challenge course research published in peer-reviewed journals reported higher effect sizes. Highest effect sizes occurred for studies conducted in therapeutic settings, perhaps due to the nature of the populations studied and the assessment of this population. Research with therapeutic outcomes had higher effect sizes than developmental or educational outcomes. Family outcome measures had the highest effect sizes and physical variables had the lowest. Self-efficacy had higher effect sizes than measures of self-esteem.

The results from this meta-analysis provide a big-picture view of research on challenge courses. Limitations of this meta-analysis are significant and highlight shortcomings in the available studies.

### *Limitations*

There are a variety of limitations in this study. Most notably, the type of facilitation, competence of the facilitator, and factors related to the delivery of the challenge course experience were not coded in this research because it was not included in the studies examined for this project. Although no empirical studies can be found, a commonly held belief among many practitioners of challenge course programming is that the facilitator can impact the outcomes of a challenge course experience both positively and negatively. Regretfully, facilitator data is not present, thus its impact does not enter into the discussion. However, the positive outcomes of this study suggest that challenge courses are effective tools regardless of facilitation type or competence.

Secondly, many of the included studies failed to provide adequate detail concerning the types and duration of activities, participant demographics, and control group information. Although the lack of detail caused limitations for the current meta-analysis, of greater concern is the lack of generalizability of challenge course research when studies cannot be accurately compared with each other or replicated.

There are a wide variety of outcome measures included in this study, including physical attributes (e.g. weight), attitudes, and observational measures. Most of the outcome measures are self-reported surveys of attitudes. One can always call into question the validity of self-report data, especially when assessing attitudes about a challenge course experience. Validity of measurement is increased with observational measures of behavior or direct measures of changes in physical variables such as weight or heart rate. Such studies represented a very small portion of the current research.

### *Implications for Practitioners*

Using the compiled research on challenge courses from a broad spectrum of outcome measures, practitioners can evaluate which activities and outcome measures are most appropriate. Self-esteem, for example, has long been a common outcome measure in challenge course research (Hattie, et al., 1997). The effect size for self-esteem as an outcome, how-

ever, is half that of self-efficacy, which is used far less often in challenge course research. Organizations and researchers should learn from this data to promote challenge courses for their more significant outcome qualities rather than continue to attempt to prove the elusive impact on self-esteem.

The highest outcome measures were family and group interactions. Challenge courses are most often marketed for their team-building qualities. This data corroborates the common qualitative assertions regarding the importance of the relationships that are positively impacted through the use of challenge courses.

A more recent category of outcome measures in relation to challenge course outcomes is physical variables. Of importance is the diversity of this category. While the effect size of the physical variable (weight loss) was 0.00, one study (Jelalian, Mehlenbeck, Lloyd-Richardson, Birmaher, & Wing, 2006) concerned with the impact of weight loss and challenge course activities recorded statistically significant findings at follow-up. Future studies using physical outcomes, such as weight gain and loss, should be encouraged for a better understanding of these results.

### *Future Research*

A key component of a meta-analysis is its heuristic value, in that it can promote more evidence-based research. As effect sizes are compiled and categorized, gaps in research become apparent and researchers and practitioners begin to formulate questions concerning outcome measures. A meta-analysis does not provide answers to the question "why" but rather offers a guide for further research.

Future research considerations from the data in this meta-analysis are extensive. Why did research with university students result in the lowest effect sizes? It could be that university students are less receptive or more resistant to these experiences, or it could be that the instruments used or type of activity that was conducted may not have been effective. In terms of activity, a combination of low and high challenge course experiences resulted in the highest effect size. Do the activities themselves provide a greater impact or is this a result of how the activities are conducted?

Only 27.3% of studies contained follow-up data. Where that data did exist, effect sizes were consistently lower for follow-up data when compared to posttest data. The lack in number of studies with follow-up data, however, reduced the quality of information that could be provided by this type of grouping. The value of increased quality and quantity of follow-up research is considerable.

Some of the highest effect sizes were recorded from research focused on therapeutic outcomes. Due to the circumstances of the participants, are they more receptive to the experiences? Are they looking for meaning in the experiences simply because they are aware of the deliberate nature of

therapy? Is therapy conducted on challenge courses more effective than traditional forms of therapy? Is the method of facilitation more effective because the facilitators are more focused on the elements of change? Is it simply that the constructs used to measure change in therapeutic outcomes are more appropriate or sensitive to the nature of change that is taking place? Or is it simply regression to the mean?

It is clear that challenge course experiences are beneficial tools for participants. The quantitative outcome, however, depends on the type of activity, the participants and, most importantly, on what outcome is being measured. A valuable experience measured qualitatively may produce ineffective quantitative results due to poor selection of data collection instruments. Similarly, poor qualitative information makes it difficult to explain and reproduce quantitative findings. Rather than continuing to pursue outcomes that are not producing significant results, researchers and practitioners alike should focus on optimizing the experience for participants in relation to their most effective and long-lasting quantitative and qualitative outcomes.

## References

Aghazarian, T. L. (1997). Use of a challenge course as an intervention tool to adolescent self-esteem (Masters thesis, San Jose State University, 1996). *Masters Abstracts International*, *35*(03), 908.

Association for Challenge Course Technology. (2004). *Challenge course standards.* (6th ed.). Deerfield, IL: Association for Challenge Course Technology.

Attarian, A., & Holden, G. T. (2005). *The literature and research on challenge courses: An annotated bibliography* (2nd ed.). Raleigh, NC: North Carolina State University and Alpine Towers International.

Bisson, C. (1998). The effects of varying the sequence of categories of adventure activities on the development of group cohesion (Doctoral dissertation, University of Northern Colorado, 1997). *Dissertation Abstracts International, 5*(03), 768A.

Blanchard, C. W. (1994). Effects of ropes course therapy on interpersonal behavior and self-esteem of adolescent psychiatric inpatients (Doctoral dissertation, New Mexico State University, 1993). *Dissertation Abstracts International, 55*(2), 584B.

Bocarro, J. (1998). An examination of the efficacy of an adventure-based counseling program on at-risk youth (Masters thesis, Dalhousie University, 1997). *Masters Abstracts International, 37*(04), 1119.

Boyle, S. (1986). The effects of a ropes course experience on locus of control and self-esteem (Doctoral dissertation, Oklahoma State University, 1985). *Dissertation Abstracts International, 46*(12), 4391B.

Breheny, M. S. (2000). Investigating the effects of a low ropes course experience vs. classroom instruction on the problem-solving appraisal of college freshmen (Doctoral dissertation, Florida Institute of Technology, 2000). *Dissertion Abstracts International, 61*(05), 1738A.

Bronson, J., Gibson, S., Kishar, R., & Priest, S. (1992). Evaluation of team development in a corporate adventure training program. *Journal of Experiential Education, 15*(2), 50–53.

Cason, D. R., & Gillis, H. L. (1993). A meta-analysis of adventure programming with adolescents. *Journal of Experiential Education, 17*(1), 40–47.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis for field settings.* Chicago: Rand McNally.

Daheim, T. J. (1999). Effects of ropes course therapy on individual perceptions of the classroom environment (Doctoral dissertation, Oklahoma State University, 1998). *Dissertation Abstracts International, 60*(02), 854B.

Eagle, H. A. (2000). Long-term differences between participants and non-participants in "Beyond the Limits" adventure education program (Doctoral dissertation, University of Maryland College Park, 1999). *Dissertation Abstracts International, 60*(12), 4363A.

Faubel, G. (1998). An efficacy assessment of a school-based intervention program for emotionally handicapped students (Doctoral dissertation, Miami Institute of Psychology of the Caribbean Center for Advanced Studies, 1998). *Dissertation Abstracts International, 58*(11), 4183A.

Faulkner, S. S. (2002). Ropes course as an intervention: The impact on family cohesion and self-esteem for adolescents in therapeutic foster care and their foster families (Doctoral dissertation, The University of Texas at Arlington, 2001). *Dissertation Abstracts International, 62*(07), 2570A.

Finkenburg, M. E., Shows, D., & DiNucci, J. M. (1994). Participation in adventure-based activities and self-concepts of college men and women. *Perceptual and Motor Skills, 78*, 1119–1122.

Gass, M. A. (1993). (Ed) *Adventure therapy.* Dubuque, IA: Kendall/Hunt.

Gass, M. A. (1995). *Book of metaphors, Volume 2*. Dubuque, IA: Kendall/Hunt.

Gillis, H. L. & Speelman, E. (2007). *Challenge course outcome research: A rough cut.* Paper presented at the 17th Annual International Challenge Course Symposium and Conference, Boston, MA.

Gillis, H. L. & Thomsen, D. (1996). *A research update (1992–1995) of adventure therapy: Challenge activities and ropes courses, wilderness expeditions, & residential camping programs.* In: Coalition for Education in the Outdoors Research Symposium Proceedings (3rd, Bradford Woods, Indiana, January 12–14, 1996). (ERIC Document Reproduction Service No. ED413128)

Greene, S. S. (1992). The use of a family ropes course experience in conjunction with diabetes camp: Can the family and the adolescent benefit? (Doctoral dissertation, The Union Institute, 1991). *Dissertation Abstracts International, 52*(09), 4975B.

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

Harper, S. L. (2006). A high ropes course and the self-concept of women in a welfare to work program (Doctoral dissertation, University of New Mexico, 2005). *Dissertation Abstracts International, 67*(01)A.

Hatch, K. D. (2006). An evaluation of a supplemental procedure geared toward prolonging challenge course benefits (Doctoral dissertation, University of Texas at Austin, 2005). *Dissertation Abstracts International,* 6923B.

Hattie, J. A., Marsh, H. W., Neill, J. T., & Richards, G. E. (1997). Adventure education and Outward Bound: Out-of-class experiences that make a lasting difference. *Review of Educational Research. 67*(1), 43–87.

Horak, R. A., Jr. (2003). The impact of adventure-based counseling on middle school climate (Doctoral dissertation, University of South Carolina, 2003). *Dissertation Abstracts International, 64*(03), 796A.

Hughes, J. S. (1994). The effects of adventure based counseling and levels of sensation seeking on the self-efficacy of chemically dependent males (Doctoral dissertation, Mississippi State University, 1993). *Dissertation Abstracts International, 54*(07), 2467A.

Jacobson, S. E. (1993). Family strengths: Effects of participation in an experiential/adventure-based program for clinically presenting families (Masters thesis, The University of Texas at Arlington, 1992). *Masters Abstracts International, 31*(03), 1088.

Jelalian, E., Mehlenbeck, R., Lloyd-Richardson, E. E., Birmaher, V., & Wing, R. R. (2006). Adventure therapy combined with cognitive behavioral treatment for overweight adolescents. *International Journal of Obesity, 30*, 31–39.

Kanters, M. A., Bristol, D. G., & Attarian, A. (2002). The effects of outdoor experiential training on perceptions of college stress. *Journal of Experiential Education, 25*(2), 257–267.

Kelly, S. A. (1996). The effects of a corporate adventure training experience on group dynamics and individual self-actualization of middle-management professionals (Doctoral dissertation, Western Michigan University, 1996). *Dissertation Abstracts International, 57*(4), 2921B.

Knott, J. M. (2004). Self-efficacy and motivation to change among chronic youth offenders: An exploratory examination of the efficacy of an experiential learning motivation enhancement intervention (Doctoral dissertation, University of Oregon, 2004). *Dissertation Abstracts International, 65*(02), 411A.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis.* Thousand Oaks, CA: Sage.

Martin, B., Cashel, C., Wagstaff, M., & Breunig, M. (2006). *Outdoor leadership: Theory and practice.* Champaign, IL: Human Kinetics.

McDaniel, P. E. (1999). Benefits of a ROPES program in working with at-risk youth's self-concept (Doctoral dissertation, St. Mary's University of San Antonio, 1999). *Dissertation Abstracts International, 59*(12), 4370A.

McDonald, R. G., & Howe, C. Z. (1989). Challenge/initiative recreation programs as a treatment for low self-concept children. *Journal of Leisure Research, 21*(3), 242–253.

McGarvey, A. L. (2004). An evaluation of a ropes course: Efficacy for at-risk youth with externalizing versus internalizing symptoms (Doctoral dissertation, The University of Connecticut, 2003). *Dissertation Abstracts International, 64*(12), 6334B.

Meyer, B. B. (2000). The ropes and challenge course: A quasi-experimental examination. *Perceptual and Motor Skills, 90*(3, pt 2), 1249–1257.

Miller, D. J. (1998). The impact of low ropes course training on team development (Doctoral dissertation, Colorado Technical University, 1997). *Dissertation Abstracts International, 58*(10), 3993A.

Moher, D., Cook, D. J., Eastwood, S., Olkin, I., Rennie, D., & Stroup, D. F. (1999). Improving the quality of reports of meta-analyses of randomised controlled trials: The QUOROM statement. *The Lancet, 354,* 1896–1900.

Moorefield, D. L. (1994). A comparative study of experiential learning utilizing indoor-centered training and outdoor-centered training (Doctoral dissertation, Texas Women's University, 1994). *Dissertation Abstracts International, 55*(06), 1479A.

Moreau de la Meuse, N. (1993). Selected psychological effects of Tae Kwon Do and a ROPES course on middle school students (Doctoral dissertation, The University of Texas at Austin, 1992). *Dissertation Abstracts International, 53*(7), 2302A.

Munns, K. M. (1996). The effects of team-building interventions (Doctoral dissertation, University of Wisconsin, 1995). *Dissertation Abstracts International, 56*(11), 6436B.

National Information Center on Health Services Research and Health Care Technology (n.d.). Retrieved December 16, 2007 from http://www.nlm.nih.gov/nichsr/hta101/ta101014.html.

Neill, J. T. (2003). Reviewing and benchmarking adventure therapy outcomes: Applications of meta-analysis. *Journal of Experiential Education, 25*(3), 316–321.

O'Bannon, P. M. (2000). The effects of the outdoor experiential ropes course on perceived team performance (Doctoral dissertation, Florida International University, 2000). *Dissertation Abstracts International, 61*(03), 852A.

Parker, M. W. (1993). Impact of adventure interventions on traditional counseling interventions (Doctoral dissertation, The University of Oklahoma, 1992). *Dissertation Abstracts International, 53*(09), 4964B.

Pazmino-Cevallos, M. (2004). The effects of an adventure education experience on team development of K–12 teachers (Doctoral dissertation, University of Southern Mississippi, 2003). *Dissertation Abstracts International, 64*(09), 3232A.

Priest, S., & Gass, M. A. (2005). *Effective leadership in adventure programming* (2nd ed.). Champaign, IL: Human Kinetics.

Priest, S., Gass, M., & Gillis, L. (2000/2003). *The essential elements of facilitation.* Seattle, WA: Tarrak Publications.

Prouty, D. (1999). Project Adventure: A brief history. In J. C. Miles & S. Priest (Eds.), *Adventure programming* (pp. 93–103). State College, PA: Venture.

Rohnke, K. (1977). *Cowstails & cobras.* Hamilton, MA: Project Adventure.

Rohnke, K., Rogers, D., Tait, C. M., & Wall, J. B. (2007). *The complete ropes course manual*. Dubuque, IA: Kendall/Hunt.

Ryan, B. (2005). *The guide for challenge course operations: An essential reference for challenge course practitioners.* Beverly, MA: Project Adventure.

Schoel, J., Prouty, D., & Radcliffe, P. (1988). *Island of healing: A guide to adventure-based counseling.* Hamilton, MA: Project Adventure.

Shoemaker, J. A. (1997). *The relationship between client introspection and challenge course efficacy.* Unpublished masters thesis, Georgia College & State University.

Smith, S. R. (1994). *Generalization of self-efficacy and self-concept following a challenge ropes course, fitness walking, and introductory psychology class*. Unpublished masters thesis, Georgia College & State University.

Sturdivant, V. A. (1991). The effect of ropes course elements on self-concept and affective behavior (Doctoral dissertation, University of North Texas, 1990). *Dissertation Abstracts International, 52*(02), 468A.

Sung, Y. K. (2005). A determination of the effects of a ropes course on the self-esteem of undergraduate and graduate students (Doctoral dissertation, Florida International University, 2004). *Dissertation Abstracts International, 65*(11), 4073A.

Talbot, P. A. (2001). An evaluation on the impact of adventure-based educational program on the self-esteem and academic achievement of adolescents at risk (Master's thesis, California State University, 2000). *Masters Abstracts International, 39*(02), 331.

Terry, L. (2002). *A quantitative and qualitative analysis on the effects of an adventure course: An intervention with female adolescents.* (Unpublished bachelors thesis, Scripps College, 2002).

Thompson, B.(2007). Effect sizes confidence intervals, and confidence intervals for effect sizes. *Psychology in the Schools, 44*, 423-432.

Thompson-Grim, L. J. (1999). Ropes programming as a means toward improving moral reasoning in at-risk youth (Doctoral dissertation, University of Houston, 1999). *Dissertation Abstracts International, 60*(05), 1465A.

Ulrich, J. S. (1993). The effects of alternative school students' participation in a ropes course experience (Master's thesis, The University of Texas at Arlington, 1992). *Masters Abstracts International, 3*(03), 1090.

Vasquez, L. M. (2001). The effects of an experiential-based prevention education program on alcohol, tobacco, and other drugs knowledge, and social attitudes and skills on first-time offender, non-adjudicated youth (Doctoral dissertation, University of Southern Mississippi, 2000). *Dissertation Abstracts International, 62*(01), 114A.

Voight, A. (1988). The use of ropes courses as a treatment modality for emotionally disturbed adolescents in hospitals. *Therapeutic Recreation Journal*, 2nd Quarter, 57–64.

White, F. E., Jr. (1998). The long-term effects of a nine-week challenge initiative program on locus of control and self-esteem of fourteen- to eighteen-year-old youth (Doctoral dissertations, The University of Utah, 1997). *Dissertation Abstracts International, 58*(08), 4178B.

Wick, D. T. (1997). Improving self-esteem of elementary school children with adlerian adventure therapy (Doctoral disseration, Our Lady of the Lake University, 1997). *Dissertation Abstracts International, 58*(03), 1553B.

Wilson, D. P. (2001). *Effect size determination program.* College Park: University of Maryland.

Wilson, S. J., & Lipsey, M. W. (2000). Wilderness challenge programs for delinquent youth: A meta-analysis of outcome evaluations. *Evaluation & Program Planning, 23*(1), 1–12.

Wisnyai, M. (1989). The effects of an integrated outdoor adventure experience on the risk-taking attitudes of high school students (Doctoral dissertation, The Ohio State University, 1988). *Dissertation Abstracts International, 50*(11), 5306B.

Witman, J. (1987). The efficacy of adventure programming in the development of cooperation and trust with adolescents in treatment. *Therapeutic Recreation Journal,* 3rd Quarter, 22–29.

Witter, B. E. (2005). The impact of facilitation of challenge initiatives on college social self-efficacy generality (Doctoral dissertation, The University of Utah, 2004). *Dissertation Abstracts International, 65*(09), 3566A.

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

Ziven, H. S. (1989). The effects of the challenge group treatment program on psychiatrically hospitalized adolescents (Doctoral dissertation, Massachusetts School of Professional Psychology, 1988). *Dissertation Abstracts International, 49*(10), 4567B.

### *Authors' Note*

Subsequent to submission, the authors became aware that Cohen's (1988) magnitudes for effect sizes were not to be employed as absolutes (Thompson, 2007) to place value on effect sizes (as is done in much of the literature). Effect sizes are to be used for "direct and explicit comparison against the effects in the related prior literature" (p. 430). In discussion with editors, a decision was made to keep Cohen's references to small. medium, and large effects in the current article. However, readers are encouraged to use this article to directly compare other challenge course research on the specific populations and specific outcomes measured. Comparing future research against these effect sizes will help establish their usefulness and accuracy. More importantly, reporting effect sizes in subsequent research will help establish a valid metric.

The authors are indebted to James Neill, not only for his excellent website at www.wilderdom.com, but for his peer review of this article prior to submission.