

Pushing Beyond Outcome: What Else Changes in Wilderness Therapy

Matthew J. Hoag

Evoke Therapy Programs

Katie Massey Combs

Evoke Therapy Programs

Sean Roberts

Cascade Crest Transitions

Patrick Logan

The University of New Hampshire

Abstract

The field of wilderness therapy has placed increasing importance on measuring effectiveness. Though studies demonstrate positive outcomes for adolescents, we lack representative samples, post-discharge data, and replication of positive results. This three-year study sought to measure outcomes and to identify mechanisms of change in wilderness therapy. We found statistically significant change from intake to discharge on the Youth Outcome Questionnaire® and on measures of hope, life effectiveness, and treatment expectancy. We discuss when change occurs as well as mechanisms of change, demographic differences, and critical lessons we learned about conducting research in a clinical setting.

Keywords: wilderness therapy, adolescents, residential treatment, outcomes

Pushing Beyond Outcome: What Else Changes in Wilderness Therapy?

The practice of wilderness therapy has grown considerably over the last twenty years. As it has changed, the importance of measuring the effectiveness of this relatively new intervention has become apparent. The evolving industry of wilderness programs has come to be identified as Outdoor Behavioral Healthcare (OBH; Gass, Gillis, & Russell, 2012; Russell, Gillis, & Lewis, 2008). The eagerness of these programs to evaluate outcome as well as to collaborate and share best practices led to the formation in 1997 of the Outdoor Behavioral Healthcare Council and member programs now do outcome research as a way to demonstrate effectiveness (OBHC, 2013a). Additionally, the National Association of Therapeutic Schools and Programs (NATSAP) created a Research and Evaluation Network with the goal of evaluating the effectiveness of member programs. Participating programs contribute to a database that collects outcome data at intake, discharge and one year post-discharge. All members of the Outdoor Behavioral Healthcare Council contribute to this database as well. More recently OBHC has begun sponsoring a yearly symposium, the Wilderness Therapy Symposium. This symposium brings together a diverse group of clinicians, administrators, field guides, and researchers associated with wilderness therapy to focus on improving the practice of wilderness therapy (OBHC, 2013b).

The present study, which we started over six years ago, represents our first foray into outcome research with adolescents in a wilderness setting. We learned a great deal about designing and implementing a research agenda while simultaneously working clinically with a group of adolescents in the wilderness (Massey, Hoag, & Roberts, 2013). Client care was our first priority and as such our caseloads often took priority over conducting research. We experienced low response rates due to our failure to monitor or assist support staff in following through with administering the protocol. As busy clinicians, we did not invest in the day-to-day execution of the study. This study has several weaknesses, though it does represent a start in our outcome research program that has developed significantly over the last six to seven years. The intention of this paper is too both disseminate our results and to educate others about our process to hopefully continue raising the quality of research within OBH.

PUSHING BEYOND OUTCOME

A variety of wilderness therapy outcome studies have provided initial evidence of efficacy (Behrens, Santa, & Gass, 2010; Behrens & Satterfield, 2007; Hoag, Burlingame, Reedy, Parsons & Hallows, 1999; Hoag, Savicki & Burlingame, 2001; Lewis, 2007; Russell, 2003, 2005, 2007; Tucker, Zelow, & Young, 2011; Young & Gass, 2010). These studies have shown a sharp reduction in symptoms for adolescents during the time they are in wilderness therapy, as well as continued improvement in mental health over the course of the year following therapy. While symptoms fluctuate post-treatment, the mean scores do not return to pre-treatment levels and may even continue to show improvement two years after clients complete wilderness programs (Russell, 2005).

Despite these gains, OBH research lacks post-discharge data, methodological sophistication, and representative samples (Russell, 2007; Scott & Duerson, 2009; Tucker et al., 2011). Additionally, most studies can be simplistically summarized as assessing whether change occurred over the course of wilderness treatment. Given the paucity of outcome studies in wilderness treatment settings and the methodological shortcomings of those we have, the focus on outcome research is understandable. However, in approaching the present study we wanted to push beyond measuring outcome with adolescents in wilderness therapy; we wanted to assess what other variables might be contributing to the changes occurring with clients in wilderness therapy programs. Therefore, in addition to the Youth Outcome Questionnaire®, we chose questionnaires assessing hope, life effectiveness, and treatment expectancy. Each of the measures was selected to explore the mechanisms of change for youth in wilderness treatment.

Methods

We invited adolescent clients and their parents at a wilderness therapy program in southwest Utah to participate in this pilot study from December 2007 to December 2010. Each client worked with one of three primary therapists. To remain eligible for inclusion in the study, adolescents needed to complete five weeks of the program. Participants were asked to complete measures (Table 1) at intake, week 3, week 5, and discharge from the treatment program. Six months after discharge, participants were asked to complete the Y-OQ® 2.0 again. Due to attrition at the 6-month follow-up, we conducted another follow-up with a random sample of 30 participants one year after the study ended. This sample of clients discharged between one and four years prior to the follow-up; therefore, it is referred to as the “12-month plus follow-up”.

Participants completed several measures for this study, including the Youth Outcome Questionnaire® 2.0 (Y-OQ® 2.0, Burlingame et al., 2001), Youth Outcome Questionnaire®-Self Report (Y-OQ® SR 2.0, Wells, Burlingame, & Rose, 2003), the Life Effectiveness Questionnaire (LEQ, Neill, Marsh, & Richards, 2003), the Hope Scale (HS, Snyder et al., 1997), and the Treatment Expectancy/Credibility Questionnaire (CEQ, Devilly & Borkovec, 2000).

The Y-OQ® 2.0 is a parent report measure of treatment progress over time for children and adolescents, aged 4-17 years old, receiving mental health services (Wells, Burlingame, & Lambert, 1999). It was constructed as a brief measure that was sensitive to change over short periods of time. The items are each rated on a 5-point Likert scale, which range from 0 (never), 1 (rarely), 2 (sometimes), 3 (frequently), to 4 (almost always). Burlingame and colleagues (2001) found the Y-OQ® 2.0 total score internal consistency to be .94 among non-clinical and clinical samples. Test-retest reliability coefficients were .83, indicating high temporal stability (Burlingame et al., 2001).

The Y-OQ® SR 2.0 is a self-report version of the Y-OQ® 2.0 to be completed by adolescents aged 12-18 years old, to provide a parallel to the parent completed version. It has demonstrated reliability including strong internal consistency ($\alpha=.95$; Wells et al., 2003). Similar to the Y-OQ® 2.0 the Y-OQ® SR items are rated on a 5-point Likert scale.

The Life Effectiveness Questionnaire (LEQ) is a 24-question self-report instrument measuring changes in life proficiency and effectiveness as a result of experiential intervention. The LEQ uses an eight-point scale (1 - “false or not like me”, 8 - “true or like me”). It has been shown to have a high internal consistency with alpha levels from eight of the scales ranging from .78 to .93 and test-retest correlations ranging from .60 to .81 (Neill et al., 2003).

PUSHING BEYOND OUTCOME

The Hope Scale (HS) is a 6 item measure assessing goal oriented behavior across two components - agency and pathway thinking. The six-point scale ranges from “*none of the time*” to “*all of the time*”. Snyder et al., (1997) related the HS has establish adequate internal consistency ($\alpha=.77$), is stable over time, and exhibits convergent, discriminant, and incremental validity.

Finally, the Treatment Expectancy/Credibility Questionnaire (CEQ) has 6 items measuring how logical and convincing the treatment is to the participant and how much he or she expects to improve (Devilley & Borkovec, 2000). It demonstrates high internal consistency ($\alpha=.79-.90$) with the factors of expectancy and credibility. It utilizes a nine-point Likert scale assessing how the participant feels about the treatment and how logical it seems to them.

Between December 2007 and December 2010, 332 adolescent clients entered the wilderness program and completed at least 5 weeks. Of these, a total of 118 adolescent clients and their parents participated in the study (36% participation rate); 36 were female (30.5%) and 82 were male (69.5%). Clients ranged in age from 13 to 17 years, with the average age being 15.9 years. Sixty-eight percent of students reported having treatment prior to the wilderness program. The median length of stay was 10.6 weeks. Diagnostic data on this sample was not collected; however, we conducted a secondary analysis of records of clients who enrolled between October 2010 and November 2011 (Hoag, Massey, & Roberts, in press). Of that sample, 74% had four or more diagnoses. The most common primary diagnostic classifications were Mood (39%), Behavior (19%), Substance-Related (17%), and Anxiety (15%). Forty-eight percent of the participants completed the Y-OQ[®] SR 2.0 at each of the four in-program data points, and 68% completed the Y-OQ[®] SR 2.0 at two of the four in-program data points (i.e., intake and discharge). We conducted a *t*-test to examine differences in Y-OQ[®] SR 2.0 scores at intake and discharge between those with complete in-program datasets (all four questionnaires) and those without. Scores were similar at intake, $t(109) = -0.475, p = 0.636$, and discharge, $t(82) = 0.742, p = .959$, suggesting that those with completed in-program datasets were representative of all participants in the study.

Results

Outcomes during the Program

We conducted paired *t*-tests to examine change from intake to discharge on each measure. We found statistically significant change, with large effect sizes on each measure (Table 2). The Y-OQ[®] 2.01 defines scores below 46 to be in the community or normal range of functioning, and a change of 13 points to be reliable change (Burlingame et al., 2005). The reliable change index (Jacobson & Truax, 1991) identifies whether the magnitude of change is clinically significant, as statistical significance does not always equate to clinical significance. The 66 point decrease in parent scores on the Y-OQ[®]-2.01 from intake to discharge is both clinically and statistically significant. The Y-OQ[®] SR defines scores below 47 to be in the community range of functioning, and a change of 18 points to be reliable change (Wells et al., 2003). Therefore, the 38-point decrease in adolescent scores from intake to discharge is clinically and statistically significant (Table 2). In addition to Y-OQ[®] scores, adolescents' assessments of their level of hope, life effectiveness, and treatment expectancies also saw statistically significant improvements with large effect sizes over the course of the program (Table 2).

Table 3 displays the mean scores for each measure over the course of the program. On the Y-OQ[®] SR, symptom change was consistent over time and it took five weeks to reach the reliable change threshold. Those who completed the Y-OQ[®] SR at all four in-program data points changed an average of 14 points from intake to week 3, 26 points from intake to week 5, and a total of 37 points from intake to discharge. A similar rate of change is seen in those who had incomplete datasets for the Y-OQ[®] SR.

Adolescents also reported consistent improvements in life effectiveness, hope, therapeutic alliance, and treatment expectancy. However, these factors accelerated at different points in treatment. Life effectiveness and hope increased only slightly in the first five weeks of treatment and made more dramatic changes during the second half of treatment. Conversely, treatment expectancy/credibility increased nearly as much during the first three weeks as it did over the remaining weeks of treatment (Table 3).

We explored the relationship between student and parent assessments on the Y-OQ[®] as well as the relationship between the Y-OQ[®] SR and the other client self-assessments. The Pearson's *r* showed that

PUSHING BEYOND OUTCOME

Y-OQ[®] parent and student assessments did not correlate at intake ($r = .220$, $n = 39$, $p = .174$), though they did at discharge ($r = .540$, $n = 17$, $p = .025$). Adolescent self-assessments of outcome correlated to self-assessments of hope and life effectiveness. The HS and Y-OQ[®] SR had a moderate correlation with a Pearson's r of $-.453$ ($n = 111$, $p = .000$) at intake and $-.473$ ($n = 83$, $p = .000$) at discharge (indicating that as symptoms on the Y-OQ[®] SR decreased, hope increased). The LEQ and the Y-OQ[®] SR showed moderate correlation as well with a coefficient of $-.490$ at intake ($n = 111$, $p = .000$), and $-.444$ at discharge ($n = 83$, $p = .000$) (suggesting that as the young people felt more effective in their life, symptoms decreased). The CEQ did not have significant correlations with the Y-OQ[®] SR.

We computed independent t -tests to examine differences in Y-OQ[®]-SR scores by several demographic factors: gender, parent's marital status, previous treatment, and ethnicity. The only statistically significant difference found was at discharge for parent's marital status. Students with parents who were married scored 13.6 points higher at discharge, $t(79) = -2.23$, $p = .029$, $d = -0.495$, than those whose parents were not married. Gender showed differences that approached statistical significance. The overall change in Y-OQ[®] SR scores from intake to discharge was 12 points greater for females than it was for males, $t(78) = 1.76$, $p = .082$. Girls assessed themselves to be worse at intake ($M_{\text{girls}} = 64$, $M_{\text{boys}} = 59$) and better at discharge compared to boys ($M_{\text{girls}} = 15$, $M_{\text{boys}} = 25$). The small sample of females may explain the lack of statistical significance.

Outcomes at Follow-up

Response rates at the 6-month follow-up were low ($N_{\text{adolescent}} = 8$, $N_{\text{parent}} = 10$) with a mean score of 36.1 on the Y-OQ[®] SR and 56.4 on the Y-OQ[®]-2.01. In order to augment the 6-month follow-up, we randomly selected a sample of 30 students in February 2012 for another follow-up. This sample of students discharged between one and four years prior to their discharge, therefore it is referred to as the "12-month plus follow-up". Of this subset of 30 students and parents, 20 parents and six students completed questionnaires for a mean score of 48.1 on the Y-OQ[®] 2.01 and 19.5 on the Y-OQ[®]-SR

Due to low sample sizes at the follow-ups, statistical testing was not appropriate. Although the follow-up responses are to be interpreted with great caution, they suggest the possibility that, at 12 months or more after discharge, parents on average assessed their children to be close to the community functioning cutoff score, and students assessed themselves to be well within community functioning.

Discussion

We found that adolescent clients made clinically and statically significant change on the parent and self-report of the Y-OQ[®] during treatment. Adolescents reported statistically significant change on measures of hope, life effectiveness, and treatment expectancy, though these factors changed at different rates at different points during the program. There were no statistically significant differences between demographic variables examined (gender, previous treatment, and ethnicity), except for parental marital status. Outcome differences between males and females were nearly significant and warrant further exploration in future research. This study corroborates several major themes in the literature for adolescents in wilderness therapy: adolescents in wilderness therapy undergo positive change, females appear to respond more to wilderness than males, and attrition is challenging.

The growing body of evidence suggests that wilderness therapy has a positive effect on adolescents. In a recent analysis of the NATSAP database, clients in OBH programs showed clinically and statistically significant change from intake to discharge on the Y-OQ[®] 30 SR (Magle-Haberek, Tucker, & Gass, 2012) and on the Y-OQ[®] 30 parent assessment (Tucker et al., 2011). The Y-OQ[®] 30 is a shorter version of the Y-OQ[®] 2.0 that provides a global index score of an adolescent's behavioral and emotional distress. In 2003, Russell conducted an outcome study with 858 adolescents from seven OBH programs using the Y-OQ[®] 2.0. This study also found clinically and statistically significant change on the Y-OQ[®] with adolescents self-reporting a decrease of 22 points and parents reporting a decrease of 52 points from intake to discharge. As in our sample, the improvement in parent scores was nearly two times greater than that reported by clients (Russell, 2003). Other studies using measures other than the Y-OQ[®] have reported significant change over the course of OBH treatment (Behrens et al., 2010).

Our sample also reflects a gender pattern that has appeared in the OBH literature: females entering with higher levels of dysfunction and a greater response to treatment than males (Magle-Haberek et al., 2012; Russell, 2003; Tucker, Javorski, Tracy, & Beale, 2013; Tucker et al., 2011). In the present

PUSHING BEYOND OUTCOME

sample, girls improved an average of 12 points more on the Y-OQ[®]-SR than males, $t(78)=1.76, p=.082$. Though this was not quite statistically significant, probably due to the small sample of females, it approaches significance and is worth exploring further. Using the Y-OQ[®], Russell (2003) and the NATSAP Research and Evaluation Network (Magle-Haberek et al., 2012; Tucker et al., 2011) found that females scored higher at admission and made greater overall change during treatment in wilderness therapy. In Russell's study both the adolescent self-report and the parent report saw greater intake scores (than males) and greater change over the program. At Russell's 12 month follow-up, Y-OQ[®] scores between males and females were similar. Russell suggests that this gender pattern could be related to differences in subscale changes (Russell, 2003) given that Burlingame et al. (1996) found that males had higher scores on behavioral dysfunction subscale and females had higher scores with the somatic subscale. In some research published using data from the NATSAP database, females self-reported higher scores at admit and made greater overall change than males; by discharge however, male and female self assessments and parent assessments of males and females on the Y-OQ[®] were comparable (Magle-Haberek et al., 2012; Tucker et al., 2011). Similarly, a study of young adults in wilderness therapy using the Outcome Questionnaire[®] (Lambert et al., 2004) found that young adult females assessed themselves to be functioning worse at intake and made greater change over the course of treatment than their male counterparts (Hoag, Massey, Roberts, & Logan, 2013).

Why these gender-based outcome differences seem to exist in OBH programs is unclear. Males largely outnumber females in wilderness treatment, though females appear to be responding more to this approach. The trend of girls entering treatment with greater dysfunction is not specific to wilderness therapy though. Several studies among various residential treatment centers reported that females also have more psychopathology than males (Baker, Archer, & Curtis, 2005; Connor, Doerfler, Toscano, Volungis, & Steingard, 2004; Handwerk et al., 2006; Hussey & Guo, 2002; Wells et al., 2003). However, unlike the wilderness therapy research, results are mixed in terms of outcome differences by gender (Cohen, 1989; Connor et al., 2004; Handwerk et al., 2006; Weis, Whitmarsh, & Wilson, 2005). Tucker, Javorski, Tracy, and Beale (2013) propose that girls may respond particularly well to OBH due to its focus on empowerment and self-efficacy, as well as using a social group format. Another possibility is the phenomenon of regression to the mean (Barnett, van der Pols, & Dobson, 2005). We agree with Tucker and colleagues' (2013) conclusion that "future research is needed to explore in more depth why or in fact if this modality truly impacts youth differently based on gender" (p. 174).

Our results show a correlation between outcome scores on the Y-OQ[®] and both the Life Effectiveness Questionnaire and Hope Scale. This suggests that participants were experiencing more hope in their lives and feeling more effective with general life skills. However, participants do not report significant change in their hopefulness or beliefs about the effectiveness of their coping skills until later in the treatment process. This may be attributed to the need for the client to stabilize, take inventory of their situation, and begin to build confidence in their ability to heal and move forward in a healthier manner. This finding points to the value in using a variable length of stay for clients to allow for internalization of changes and to ensure they have had sufficient time to experience increased hopefulness and belief in the effectiveness of their coping skills.

We found that the Y-OQ[®] scores for parents and adolescents, while not correlated at intake, were correlated at discharge. Parents are often highly emotional and in a state of crisis upon their son or daughter's admission to wilderness. Conversely, the adolescent is often in denial and externalizing fault for struggles or discord. This tension may account for the divergent scores of parents and clients on the Y-OQ[®] at intake. Similarly, it is possible that, as parents move out of their heightened emotional state and clients begin to see their process more clearly, scores on these measures more closely align at discharge. This reflects the movement to a more objective and unified family system, something that is anticipated and suggested as a goal in a wilderness therapy program.

Limitations

The chief limitations of this study were low parent participation and follow-up response rates. Only 40% of parents who agreed to participate actually completed a questionnaire at intake, and parent participation decreased further over the course of the study. Participation from both parent and adolescent clients post-discharge was particularly challenging. While in-program participation for adolescent clients was strong with 68% completing intake and discharge questionnaires, only 14% of parents responded at both intake and discharge, and post-discharge response rates were too low to conduct statistical analyses. It is also worth noting that all participants were from one wilderness

PUSHING BEYOND OUTCOME

program site; and therefore the results may not be generalizable to wilderness therapy as a whole.

These challenges are not specific to this study or to OBH research. Difficulty with parent participation has been experienced in a number of efforts to evaluate outcomes in residential treatment (Behrens & Satterfield, 2007; Russell, 2007; Tucker et al., 2011). For example, in an analysis of the data in the NATSAP database, Tucker et al. (2012) report that there were 879 adolescent matched pairs for intake and discharge in OBH programs, however there were only 171 matched pairs for parents.

Though gains are being made, longitudinal data is a major gap in the literature for OBH and the broader field of private residential treatment. OBHC and NATSAP are making new commitments to collecting in-program and post-discharge data (OBHC, 2013a), and using technology (Outcome Tools) that allows parents and clients to complete questionnaires online. In 2011 the NATSAP Research and Evaluation Network began using Outcome Tools (Outcome Tools, 2012). Looking at the different methods and response rates from our 6-month and 12-month plus follow-up, one can see hope for future post-discharge data with this technology. With this online system our response rates increased from less than 10% for parents and clients to 67% for parents and 20% for adolescents.

Lessons Learned

This exploratory effort provided critical lessons on how to conduct research in a clinical setting. The essential lessons we took from this pilot study were: having an appointed staff with the skills, interest, and time to coordinate the study; constant communication between all levels of staff including therapists, researchers, and field and office staff; close monitoring of data collection; utilization of technology for data collection; and greater investment in post-discharge follow-up. Quality research demands consistent attention. That attention is hard to give if researchers are simultaneously serving as full-time clinicians or administrators. The level of consistent monitoring and communication needed to carry out research is not sustainable without appointing or hiring a staff member who has the appropriate skills, interest, and time.

As discussed above, post-discharge follow-up was a major challenge and limitation in this study and in the literature of OBH. Utilizing technology that allows participants to respond more easily will likely improve this. However, we also believe that increased efforts and investment in reaching clients after they leave a program is necessary to attain representative samples. We recognized that there is no easy way to reduce attrition, and that there is no replacement for the hours spent attempting to solicit responses.

Conclusions

This study supports the consistent finding in OBH literature that adolescents change dramatically over the course of wilderness therapy. Our sample of adolescents showed a marked decrease in symptoms over the course of this wilderness program according to parent report and adolescent self-assessments. Clients also reported significant improvements in hope, life effectiveness, and treatment expectancy and credibility. Hope and life effectiveness measures were significantly correlated to outcome and accelerated in the second half of the program.

Weaknesses that have riddled the OBH literature were also present in this study. However, this pilot study taught us valuable lessons in overcoming issues of post-discharge attrition and low parent participation. We believe that it is essential to invest more resources in post-discharge follow-up, to appoint staff with the appropriate skills and time to coordinate a research project, and to utilize technology that makes it easier for participants to complete questionnaires and for staff to manage and monitor the research. These lessons helped us further develop our research program at a multi-site wilderness program, and will hopefully aid in the collective responsibility of the industry to build a thorough and diverse body of literature on the change process in wilderness therapy.

PUSHING BEYOND OUTCOME

References

- Baker, A. J., Archer, M., & Curtis, P. A. (2005). Age and gender differences in emotional and behavioral problems during the transition to residential treatment: The Odyssey Project. *International Journal of Social Welfare, 14*(3), 184-194.
- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: what it is and how to deal with it. *International Journal of Epidemiology, 34*(1), 215-220.
- Behrens, E., Santa, J., & Gass, M. (2010). The evidence base for private therapeutic schools, residential programs, and wilderness therapy programs. *Journal of Therapeutic Schools and Programs, 4*(1), 106-117.
- Behrens, E., & Satterfield, K. (2007). Longitudinal family and academic outcomes in residential programs: How students function in two important areas of their lives. *Journal of Therapeutic Schools and Programs, 2*(1), 81-94.
- Burlingame, G. M., Mosier, J. I., Wells, G., Atkin, Q. G., Lambert, M. J., Whoolery, M., & Latkowski, M. (2001). Tracking the influence of mental health treatment: The development of the Youth Outcome Questionnaire. *Clinical Psychology and Psychotherapy, 8*(5), 361-379.
- Burlingame, G. M., Wells, M. G., Cox, J. C., Lambert, M. J., Latkowski, M., & Justice, D. (2005). Administration and scoring manual Y-OQ 2.01 (Youth Outcome Questionnaire). Wharton, NJ: American Professional Credentialing Services.
- Cohen, N. J. (1989). Sex differences in child psychiatric outpatients: Cognitive, personality, and behavioral characteristics. *Child Psychiatry & Human Development, 20*(2), 113-121.
- Connor, D. F., Doerfler, L. A., Toscano, P. F., Volungis, A. M., & Steingard, R. J. (2004). Characteristics of children and adolescents admitted to a residential treatment center. *Journal of Child and Family Studies, 13*(4), 497-510.
- Devilley, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectance questionnaire. *Journal of Behavior Therapy and Experimental Psychiatry, 31*(2), 73-86.
- Gass, M. A., Gillis, H. L., & Russell, K. C. (2012). Adventure therapy: Theory, practice, & research. NY: Routledge Press.
- Handwerk, M. L., Huefner, J. C., Smith, G. L., Clopton, K., Hoff, K. E., & Lucas, C. P. (2006). Gender differences in adolescents in residential treatment. *The American Journal of Orthopsychiatry, 76*(3), 312- 324.
- Hoag, M. J., Burlingame, G. M., Reedy, B., Parsons, P., & Hallows, G. (1999, May). The efficacy of wilderness therapy: Analysis of change using the Y-OQ. Poster presentation at the annual meeting of the Western Psychological Association, Irvine, CA.
- Hoag, M. J., Massey, K. E., Roberts, S. D., & Logan, P. (2013). Efficacy of wilderness therapy for young adults: A first look. *Journal of Residential Treatment of Children and Youth, 30*(4), 294-305. Advance online publication. doi:10.1080/0886571X.2013.852452.
- Hoag, M. J., Massey, K. E., & Roberts, S. D. (in press). Dissecting the wilderness therapy client: Examining clinical trends, findings, and industry patterns. *The Journal of Experiential Education*.
- Hoag, M. J., Savicki, K., & Burlingame, G. M. (2001). The efficacy of wilderness therapy: Individual and familial aspects of change. In Michael C. Roberts & Keri J. Brown (Chairs), Evaluation of innovative interventions - Therapeutic camping for children. *In a paper presented as a part of a symposium conducted at the meeting of the American Psychological Association Convention, San Francisco, CA.*

PUSHING BEYOND OUTCOME

- Hussey, D. L., & Guo, S. (2002). Profile characteristics and behavioral change trajectories of young residential children. *Journal of Child and Family Studies, 11*(4), 401–410.
- Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12–19.
- Massey, K. E., Hoag, M. J., & Roberts, S. D. (2013, September). How to do research in a clinical setting without ‘failing’: Ideas and strategies. Presentation at the annual Wilderness Therapy Symposium, Boulder, CO.
- Lambert, M. J., Morton, J. J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R. C., Shimokawa, K., Christopherson, C., & Burlingame, G. M. (2004). *Administration and scoring manual for the OQ-45*. Orem, UT: American Professional Credentialing Services.
- Lewis, S. (2007). The outdoor division of Aspen Education Group: Examining treatment effectiveness. Retrieved March 18, 2013, from <http://natsap.org/research/research-at-natsap-programs>.
- Magle-Haberek, N. A., Tucker, A. R., & Gass, M. A. (2012). Effects of program differences with wilderness therapy and residential treatment center (RTC) programs. *Residential Treatment or Children & Youth, 29*(3), 202–218.
- Neill, J. T., Marsh, H. W., & Richards, G. E. (2003). The Life Effectiveness Questionnaire: Development and psychometrics. Unpublished manuscript, *University of Western Sydney*, Sydney, NSW, Australia.
- OBHC (Outdoor Behavioral Healthcare Council). (2013a). Research. Retrieved October 30, 2013, from <https://www.obhic.com/research/>.
- OBHC (Outdoor Behavioral Healthcare Council). (2013b). Wilderness Therapy Symposium. Retrieved November 4, 2013, from <https://www.obhic.com/symposium/>
- Outcome Tools. (2012). Outcome Tools. Retrieved March 18, 2013, from <https://outcometools.com/site/>
- Russell, K. C. (2003). An assessment of outcomes in outdoor behavioral healthcare treatment. *Child and Youth Care Forum, 32*(6), 355–381.
- Russell, K. C. (2005). Two years later: A qualitative assessment of youth well-being and the role of aftercare in outdoor behavioral healthcare treatment. *Child and Youth Care Forum, 34*(3), 209–239.
- Russell, K. C. (2007). Summary of research in the Outdoor Behavioral Healthcare Research Cooperative from 1999-2006. Minneapolis, MN.
- Russell, K. C., Gillis, H. L., & Lewis, G. T. (2008). A five-year follow-up of a survey of North American outdoor behavioral healthcare programs. *Journal of Experiential Education, 31*(1), 55-77. doi:10.5193/JEE.31.1.55
- Scott, D. A., & Duerson, L. M. (2009). Continuing the discussion: A commentary on “Wilderness Therapy: Ethical considerations for mental health professionals”. *Child & Youth Care Forum, 39*(1), 63-68.
- Snyder, C. R., Hoza, B., Pelham, W. E., Rapoff, M., Ware, L., Danovsky, M., . . ., Stahl, K. J. (1997). The development and validation of the children’s hope scale. *Journal of Pediatric Psychology, 22*(3), 399-421.
- Tucker, A. R., Javorski, S., Tracy, J., & Beale, B. (2013). The use of adventure therapy in community-based mental health: Decreases in problem severity among youth clients. *Child & Youth Care Forum, 42*(2), 155-179. Springer US.

PUSHING BEYOND OUTCOME

- Tucker, A. R., Zelov, R., & Young, M. (2011). Four years along: Emerging traits of programs in the NATSAP Practice Research Network (PRN). *Journal of Therapeutic Schools and Programs*, 5(1), 10-28.
- Weis, R., Whitmarsh, S. M., & Wilson, N. L. (2005). Military-style residential treatment for disruptive adolescents: Effective for some girls, all girls, when, and why? *Psychological Services*, 2(2), 105-122.
- Wells, M. G., Burlingame, G. M., & Rose, P. M. (2003). Administration and scoring manual for the Y-OQ SR-2.0 (Youth Outcome Questionnaire-Self Report). Wilmington, DE: American Professional Credentialing Services.
- Young, M. C., & Gass, M. A. (2010). Preliminary data from the NATSAP research and evaluation network: Client characteristics at admission. *Journal of Therapeutic Schools and Programs*, 4(1), 80-105.

PUSHING BEYOND OUTCOME

Table 1
Description of Measures

Measure	Description
Youth Outcome Questionnaire® 2.01 (Y-OQ® 2.01)	The Y-OQ® 2.01 is a 64 item report completed by the parent or guardian. It measures treatment progress for children and adolescents (ages 4-17) receiving a mental health intervention, and tracks actual change in functioning (Burlingame et al., 2005).
Youth Outcome Questionnaire®-Self Report (Y-OQ® SR 2.0)	A self-report version of the Y-OQ® 2.01. It serves as an additional source of data in tracking treatment progress for adolescents ages 12-18 receiving mental health treatment (Wells, Burlingame, & Rose, 2003).
Life Effectiveness Questionnaire (LEQ)	A self-report instrument measuring the effects of experiential education and the extent to which a person's actions, behaviors, and feelings are effective in succeeding in life and general life skills (Neill, Marsh, & Richards, 2003).
Hope Scale (HS)	A 6 item self-report index measuring agency and pathway thinking toward goals. Agency thoughts reflect the perception that children can initiate and sustain action toward a desired goal; pathways thoughts reflect one's perceived capability to produce routes to those goals (Snyder et al., 1997).
Treatment Expectancy/Credibility Questionnaire (CEQ)	A 6 item instrument measuring treatment expectancy and credibility. It assesses how logical and convincing the treatment is to a client and how much a client expects to improve (Deville & Borkovec, 2000).

Table 2
Paired t-Tests on Change in Score from Intake To Discharge

Measure	M_{intake} (SD)	M_{exit} (SD)	t	df	d
Y-OQ® 2.01	93.3 (25.9)	27.8 (22.1)	6.97**	15	2.7
Y-OQ® SR	58.9 (32.8)	20.7 (27.8)	11.75**	79	1.3
HS	23.1 (5.7)	29.6 (4.1)	-9.26**	79	-1.3
LEQ	6.1 (1.0)	7.0 (0.8)	-7.48**	78	-1.0
CEQ	38.1 (11.3)	48.4 (8.7)	-6.94**	80	-1.0

Note: ** indicates a $p < .001$