

Towards an Evidence-Informed Adventure Therapy: Implementing Feedback-Informed Treatment in the Field

Will Dobud

To cite this article: Will Dobud (2017): Towards an Evidence-Informed Adventure Therapy: Implementing Feedback-Informed Treatment in the Field, Journal of Evidence-Informed Social Work, DOI: [10.1080/23761407.2017.1304310](https://doi.org/10.1080/23761407.2017.1304310)

To link to this article: <http://dx.doi.org/10.1080/23761407.2017.1304310>



Published online: 07 Apr 2017.



Submit your article to this journal [↗](#)



View related articles [↗](#)



View Crossmark data [↗](#)

Towards an Evidence-Informed Adventure Therapy: Implementing Feedback-Informed Treatment in the Field

Will Dobud

School of Humanities and Social Sciences, Charles Sturt University, Sydney, Australia

ABSTRACT

As an intervention for adolescents, adventure therapy has evolved considerably over the last three decades with support from multiple meta-analyses and research input from both residential and outpatient services. Tainted by a history of unethical practice and issues of accountability, this article explores the question of how adventure therapy can meet a standard of evidence preferred by policymakers and funding bodies on the international stage. In this case, feedback-informed treatment (FIT) is presented as a means for routine outcome management, creating a framework for adventure therapy which aims to improve the quality of participant engagement while maintaining and operationalizing today's definitions for evidence-based practice. A case vignette illustrates the use of FIT with an adolescent participant engaged on a 14-day adventure therapy program.

KEYWORDS

Adventure therapy;
feedback-informed
treatment; adolescents;
psychotherapy; routine
outcome management

It is hard to come across an adventure therapy (AT) text that does not provide some dialogue on whether there exists an undisputed definition of what AT is, or if there is any consensus on how it works (Bowen & Neill, 2013; Fernee, Gabrielsen, Andersen, & Mesel, 2017; Norton et al., 2014; Russell, 2001; Russell & Hendee, 2000). Addressing this in what is today's most comprehensive AT text, American researchers, Gass, Gillis, and Russell (2012), defined AT as "the prescriptive use of adventure experiences provided by mental health professionals, often conducted in natural settings that kinesthetically engage clients in cognitive, affective, and behavioral levels" (p. 1). The Adventure Therapy International Committee, however, maintain that the "implication of any definition is that there will be clarity about what lies inside and what lies outside the definition—and there is no hard boundary around adventure therapy in a global context" (ATIC, 2016).

The call for understanding how AT works and establishing greater scientific verifiability is nothing new. With a history involving reports of abuse, boot camps, and a government investigation delivered before the U.S. Congress, accountability has been in the forefront of the conversation (GAO, 2007; Gass et al., 2012; Norton et al., 2014). In 1996, the now named Outdoor Behavioral Healthcare Council was established "in an effort to address this scrutiny and promote accountability" (Tucker, Widmer, Faddis, & Randolph, 2016, p. 33) since finding AT safer for adolescents compared to those in the general public (Javorski & Gass, 2013) and to deliver clinically significant outcomes (Russell, 2003; Russell & Hendee, 2000). While noteworthy, Fernee and colleagues (2017) admit AT "has yet to establish itself as a viable treatment

CONTACT Will Dobud, MSW  info@willdobud.com  2 Kintore Avenue, Hazelwood Park, SA 5066, Australia.

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/webs.

© 2017 Taylor & Francis

option within the continuum of care for adolescents in need of mental health treatment” (p. 114). Additionally, Australian psychologists Bowen and Neill (2013) maintain that less than 1% of worldwide AT practice undergo program evaluation.

This article will explore routine outcome management (ROM) using feedback-informed treatment (FIT), a “pantheoretical approach for evaluating and improving the quality of mental health services” (Bertolini & Miller, 2012, p. 2). Capable of being integrated into any therapeutic service, FIT has been shown to improve outcomes, shorten inpatient stays, and reduce client deterioration (Miller, 2011; Miller, Hubble, Chow, & Seidel, 2015). For example, the Norway Feedback Trial, the largest ever randomized clinical trial in couples therapy, found couples in the feedback group achieved “nearly 4 times the rate of clinically significant change, and maintained a significant advantage on the primary measure at 6-month follow-up while attaining a significantly lower rate of separation or divorce” (Anker, Duncan, & Sparks, 2009, p. 693). For this article, a case vignette will illustrate how FIT can be implemented on an AT program.

Evidence and adventure therapy

With recent meta-analyses demonstrating moderate to large effect-sizes for AT participants (Bettman, 2012; Bowen & Neill, 2013; Gillis et al., 2016b), Gillis and colleagues (2016b) see that “what is missing is the routine monitoring of outcomes . . . while youth are in treatment” leaving the field stuck with a “proverbial ‘black box’ of pre- and post robust outcomes” (p. 860). This black box problem for researchers, addressed in a qualitative synthesis by Fernee and colleagues (2017), “commonly refers to outcomes studies that are concerned only with effects, paying little attention to how these effects are produced” (p. 115). Norton and colleagues (2014) also cite “rising concern about documenting what specific factors . . . are involved in the change process” (p. 51).

Specifying time spent in nature, therapist contact time, novel group experiences, mindfulness, and reflection, or the development of success and mastery to name a few, many have proposed what specific factors are at play within AT (Gass et al., 2012; Harper, 2007; Hoag, Massey, & Roberts, 2014; Norton, 2010; Russell, 2003; Russell, Gillis, & Heppner, 2016; Russell & Hendee, 2000). Nearly half a century of psychotherapy research, however, suggests that general therapeutic factors, like the therapeutic alliance, the idiosyncratic characteristics of both the therapist and client, and the quality of participant engagement contribute more to therapeutic outcomes (Asay & Lambert, 1999; Lambert, 2013; Orlinsky, Grawe, & Parks, 1994; Wampold, 2010). In the only routine outcome measurement AT study to date, Gillis, Kivilighan, and Russell (2016) studied the engagement/outcome relationship of 68 young adults in a residential wilderness treatment facility finding that when clients were less engaged, outcomes deteriorated. In contrast, higher levels of reported engagement predicted improved outcomes.

To achieve the desired standard of scientific evidence, researchers use the supposed gold standard randomized control trial (RCT). RCTs have found psychotherapy to be largely effective with reported effect sizes ranging from .8 and 1.2 (Asay & Lambert, 1999; Bertolini & Miller, 2012; Wampold, 2001) despite no significant improvement since psychotherapy’s first meta-analysis conducted by Smith and Glass (1976). However, no specific ingredient required for therapeutic change has been identified (Asay & Lambert, 1999; Lambert, 2013; Lambert & Bergin, 1994; Wampold, 2010).

In AT, RCTs have been a challenge. For Gabrielsen, Fernee, Aasen, and Eskedal (2015), the struggle came while implementing an AT program within the Department for Child and Adolescent Mental Health in a hospital in southern Norway. Attempting to answer the call for comparison groups and scientific verifiability (Behrens, Santa, & Gass, 2010; Norton et al., 2014), the researchers found greater ethical concerns when randomly assigning young participants to treatment-as-usual control groups (Gabrielsen et al., 2015). Adolescents referred to the hospital were provided the opportunity to meet with the researchers to discuss the AT program versus engaging in the hospital's other therapeutic services. Privileging the adolescents' treatment preference and taking previous treatment failures into account, Gabrielsen and colleagues (2015) asked "if we assume that the members of a control group are pre-conditioned to respond poorer to the health care that they are offered, is it right to have them in a control group in the first place?" Although evidence from RCTs is required to establish an accepted evidence base, an RCT was not a viable option for these researchers.

Another concern, as pointed out by Diaz and Drewery (2016), is that in contrast to medicine, there is likelihood for greater ambiguity of evidence within the helping professions. In the medical model's view of therapeutic change, diagnostic criteria and empirically supported treatments are perceived causal agents. For psychotherapy, however, research suggests that client and therapist factors have more influence over successful treatment than the model employed. Miller, Hubble, Chow, and Seidel (2013) also report "consistently greater variance in outcomes between psychotherapists . . . than between the types of therapy they are practicing" (p. 89).

With the contribution these general factors have to therapeutic outcomes, a presidential task force formed by the American Psychological Association (2006, pp. 276–277) granted that "clinical expertise also entails the monitoring of patient progress . . . [and] if progress is not proceeding adequately, the psychologist alters or addresses problematic aspects of the treatment (e.g., problems in the therapeutic relationship or in the implementation of the goals of the treatment) as appropriate." Likewise, in their 2016 practice standards, the National Association for Social Workers specified that evidence-informed social work involved the "solicitation and incorporation of feedback from clients regarding the extent to which social work services have helped them identify and achieve their goals" accomplished through "measurement of both process and objectives" (NASW, 2016, p. 34).

Because of the unique context of each therapeutic encounter, routine measurement of client outcomes and the participant's perception of the therapeutic alliance directs attention towards the most robust variables in the therapeutic experience. Gillis and colleagues (2016a) encourage this monitoring for adventure therapists, recommending "tracking week-to-week changes in . . . engagement to identify group members who are not getting optimal program benefits" (p. 413). In the following section, FIT will be presented as a reliable and feasible option for collecting and tracking this feedback.

Feedback-informed treatment (FIT)

The concept of routinely monitoring outcomes was first put forth by Howard, Moras, Brill, Martinovich, and Lutz (1996) speculating that session to session monitoring could help practitioners determine if their services were effective for a specific client. Exploring the benefits and obstacles of implementing ROM, Boswell, Kraus, Miller, and Lambert (2013) found that when "implemented, the risk of patient deterioration is significantly decreased, . . . effect sizes are

enhanced, and in some extreme cases tripled” (p. 7). Meta-analyzing the results of over a dozen RCTs, Miller (2011) used a diverse sample of 12,374 cases noting that providing feedback to therapists about the progress of treatment and the client’s perspective of the therapeutic alliance led to as much as a doubling in reliable and clinically significant change and decreases in dropouts, deterioration, and hospital stays. Alerting therapists to deterioration is critical as Hannan and colleagues (2005) studied if therapists could accurately predict deterioration. Of the 550 cases used in the study, the therapist group was unable to identify 39 of the 40 deteriorating clients.

In 2013, FIT was added to the Substance Abuse and Mental Health Services’ National Registry of Evidence-based Programs and Practices receiving perfect marks for readiness of dissemination identifying “no adverse effects, concerns, or unintended consequences” (SAMHSA, 2012). FIT provides a method by which therapists gather real-time feedback on the outcome of treatment and therapeutic alliance at each therapeutic encounter. Though FIT can be implemented using other measures, this article discusses the possibility of incorporating two popular and feasible ultra-brief scales, the outcome rating scale (ORS), the session rating scale (SRS), and the group session rating scale (GSRS) for AT (Miller et al., 2015).

The ORS was developed as a briefer alternative to the popular Outcome Questionnaire 45.2 (OQ 45.2; Lambert et al., 1996) hoping to create a reliable and valid outcome measurement tool feasible in naturalistic settings. For programs operating in wilderness environments, the feasibility of collecting data is an important consideration. The scale has high internal consistency ($\alpha = .93$), test-retest reliability, and moderate concurrent validity (.59) with the OQ 45.2, used commonly in AT literature (Gass et al., 2012; Miller, Duncan, Brown, Sparks, & Claud, 2000). While longer scales do provide greater detail, the ORS allows for a psychometric description of overall distress (Miller et al., 2000). The scale is broken into four areas: (a) individual, (b) interpersonal, (c) social, and (d) overall well-being. Each item has a corresponding line, 10 cm long, against which clients place a check mark to describe how they are experiencing these areas. For adults, the ORS has a clinical cutoff of 25 and for adolescents of 28. In presenting the psychometrics of the ORS, Bertolini and Miller (2012) note that a clinical cutoff is helpful in defining the “boundary between a normal and clinical range of distress” and providing “a reference point for evaluating the severity of distress for a particular client or client sample” (p. 4). This cutoff is based on a robust sample of 34,790 ORS administrations.

The SRS, a similar four-item tool, provides a measure for assessing the quality of the therapeutic alliance. Administered at the end of an encounter, the SRS measures areas of the working alliance conceptualized by Bordin (1979): (a) the relational bond, (b) the importance of the goals or topics discussed in the session, (c) the therapist’s approach or method, and (d) an overall rating for the session. After clients place their check marks, the therapist can measure where the check mark falls on each 10-cm line, providing a score out of 10 and adding to a total possible score of 40 in less than a minute.

The SRS was tested against the revised Helping Alliance Questionnaire (HAQ-II; Luborsky et al., 1996) to establish reliability and validity. Although it is expected that a four-item measure have lower reliability, Duncan and colleagues (2003) found the SRS to have high internal consistency ($\alpha = .88$), which compared positively to the HAQ-II ($\alpha = .90$). Using Pearson’s *r*, test-retest reliability for the SRS was .64, which again compared favorably to the HAQ-II (.63). For concurrent validity, Duncan and colleagues (2003) found “all correlations between SRS items and total HAQ-II scores were within a range of .39 to .44” (p. 9). The

authors also found a 98% compliance rate with the SRS making it very feasible for clinical practice. Confirming the reliability and validity of the ORS and SRS, Campbell and Hemsley (2009) also point out that brief scales are beneficial for use in real world practice settings as they are easy to interpret, require minimal training, and are low-cost.

The SRS “has the limitation of focusing only on the individual’s experience of the alliance” (Quirk, Miller, Duncan, & Owen, 2012, p. 1). For adventure therapists working with groups, the GSRS (Duncan & Miller, 2007) may be used by “leaders to better identify group members who do not feel the group experience is assisting them to reach their goals and consequently could prevent therapeutic failures” (Quirk et al., 2012, p. 6). After administering the GSRS to a group, the leader may discuss participant scores with the group or concerned individuals.

While these tools have been used for evaluation, FIT is a clinical tool that relies on therapist knowledge of warning signs while working with clients to interpret outcomes (Duncan et al., 2003; Tilsen & McNamee, 2015). Similar to the findings of Gillis and colleagues (2016a), higher ratings of the therapeutic alliance early on in therapy are predictive of success, while poorer ratings are likely to produce null outcomes and dropouts (Baldwin, Wampold, & Imel, 2007). When using these alliance measures, therapists should look out for scores falling under the clinical cutoff of 36, and engage clients in conversations about changing course. Second, significant outcomes are more likely to occur early on in treatment. If no early change is present, Lambert (2010) recommends therapists “enter frank and open discussions with clients about their progress” (p. 241) as the majority of clients, 60–65%, exhibit some symptom relief between the first and seventh visit (Bachelor & Horvath, 1999).

The following section will present a case vignette of how FIT was implemented in an Australian AT program illustrating how adventure therapists may incorporate ROM into their practice.

Evidence-informed adventure therapy: A case vignette

Referred to the AT program by his clinical psychologist, Sammy (16) was a therapy veteran. With previous diagnoses of conduct disorder, attention deficit hyperactivity disorder, and depression, Sammy had many treatment failures to his name. Sammy agreed to attend this AT program as a friend of his completed the program previously a year earlier and found it to be beneficial.

Upon arriving, the social worker (SW) leading the program met individually with each of the six adolescents and introduced the ORS and GSRS, their purpose, and the process of FIT. Sammy willingly completed the ORS with a total score of 13.4. Using a web-based computer system, SW printed a graph (See [Figure 1](#)) for Sammy to discuss predicted success trajectories and how the two, together, would monitor Sammy’s progress. SW asked if this score provided an accurate snapshot of how his life had been going. Sammy identified that his biggest concerns were “lack of direction,” such as not going to school, and his marijuana use. These became initial goals for Sammy.

On Day 4, SW again met with each participant and administered the GSRS. Sammy scored a 28.1, well below the clinical cutoff. Because the GSRS considers the relationship between participants, as well as the group leaders, SW held a group discussion to address potential issues. Sammy said that although he trusted SW, he did not feel the program was appropriately

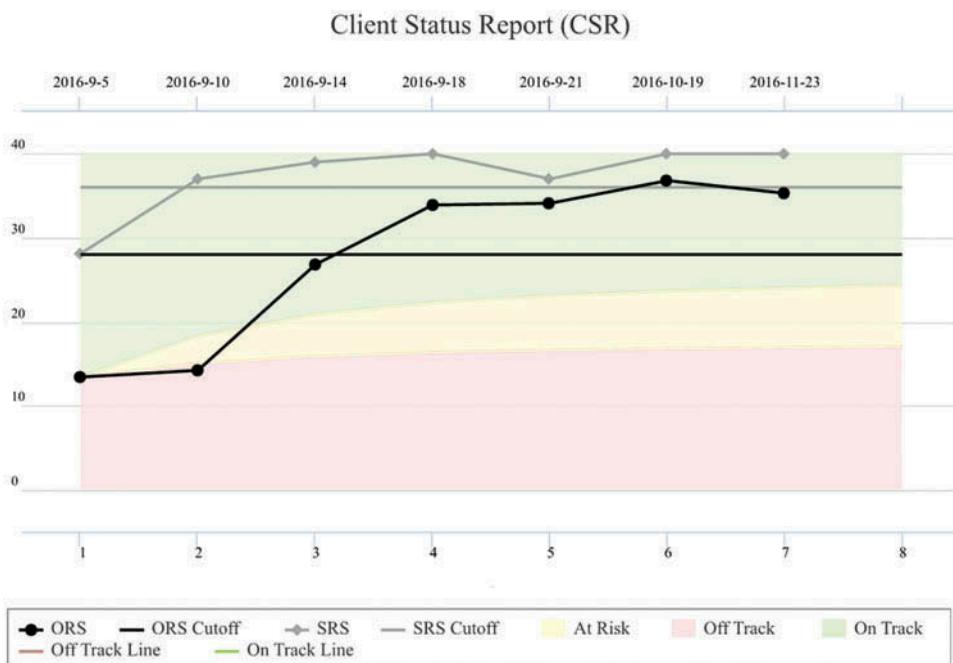


Figure 1. Sammy's ORS and GRS scores.

addressing his concerns. He did not feel comfortable talking in front of the group. SW thanked Sammy for the feedback and the two agreed to continue this discussion away from the group.

The second administration of the ORS occurred the following day. Sammy's score of 14.2 showed no significant change alerting SW that Sammy's score was falling off-track. Sammy's score was plotted on the graph and the two discussed what changes needed to be made to illicit improvement. Sammy said that while he enjoyed the program thus far, he did not feel as though it was "going to help with issues at home." Sammy said he needed to focus on "confidence and relationship with parents" to improve. The two agreed that more journal assignments around these topics would be helpful for Sammy and arranged to spend more time together away from the group.

Day 9, Sammy completed his second GRS rating scoring a 37, above the clinical cutoff. Addressing the improvement, Sammy said he felt "the group was functioning well together" and believed they were self-sufficient. The following morning, Sammy's ORS score improved to 26.8. Sammy said that the increased time spent journaling and reflecting were beneficial to his improvement. He reported feeling remorse about hurting his parents and causing them to "lose trust in him" but "felt that he could earn it back." He said his score would improve should he earn back their trust, which required reducing his drug use and attending school.

The final initiative of the program was a solo experience inviting participants to spend a night on their own away from the group. Sammy told the group he wanted to "prove" he could "survive in the Australian bush" expressing that a two-night solo would give him the time to reflect on his situation, write in his journal, and think about what he needed to do before returning home. Not without apprehension, the group decided that a two-night solo was

preferred to “put their skills to the test.” Each participant was given a spot for the two nights within sight of the program leaders but far enough away as to not interrupt their experience.

As the sun set on the second night, Sammy visited SW declaring he had “discovered the meaning of life” stating “the only thing that matters was the present moment.” He said he was bored during the second day of solo but realized that this opportunity was about “becoming comfortable with himself.” As the program leaders were thrilled with Sammy’s positivity, SW shared a letter sent from Sammy’s father. Sammy was receptive to the new boundaries his father intended to set and returned to his campsite to write a reply. His lengthy reply addressed his substance abuse, and though he felt unsure about returning to school, he wanted to join the army and knew this would require him to be clean and sober.

After the solo, Sammy’s ORS score improved to a 33.9. His final GSRS was a 40. He acknowledged that mindfulness strategies, remaining drug-free, and remembering the importance of his solo experience were necessary for maintaining such significant improvement. After the program, Sammy’s case was returned to his previous psychologist. Although this AT program offers aftercare support, Sammy had a “close relationship” with his psychologist, and this was respected. Sammy and his father did, however, schedule monthly Skype sessions with SW to discuss how things were going. Sammy had a relapse 2 weeks post-program and addressed this with his psychologist. He had joined a volunteer fire department and returned to school to complete an apprenticeship program. Because Sammy’s scores had sustained at the third meeting, it was agreed that this service was completed.

Six months post-program, a survey was sent to Sammy and his father to gather qualitative feedback about Sammy’s AT experience. Sammy reported that he had remained drug-free and had returned to school full-time. His father said Sammy’s relationship with his mother was “still rocky” but Sammy was coping much better. Sammy asked if he could return to the AT program as a peer group mentor in the future.

Implications for social workers and adventure therapists

Evidence-informed AT places participants and their families as key stakeholders. By maintaining a focus on the engagement and quality of participation (Orlinsky et al., 1994), relational responsibility takes center stage. As Gillis and colleagues (2016a) showed engagement to be a predictor of outcome, FIT may be a feasible and reliable method, which is low cost and requires minimal training, for measuring both outcome and engagement. In this framework, theories of AT rely first and foremost on a relational allegiance to improve engagement before the specific ingredients of the model itself.

In practice, using FIT measures may also reduce the potential for a power differential in treatment. As some adolescents are involuntarily mandated to attend AT programs, it is important that AT or any residential treatment uphold the values of social justice and advocacy, ensuring that treatment does not become another branch of social power privileging therapist theories of pathology over a participant’s theory of change. Sparks and Muro (2009) offer two philosophies worth considering, “one that values client direction . . . and one that mistrusts it” (p. 72). FIT privileges systematic client feedback to alarm therapists to cases at-risk of deterioration placing clients at center stage in the therapeutic process.

Lundervold and Belwood (2000) perceive these single-case designs for evaluation as the “best kept secret in counseling” as they are sufficient for scientific evaluation in real work practice settings. Addressing specific internal or external threats of validity, Nielson (2015)

notes that FIT may meet the needs for program evaluation and single-case design studies though rigor is required to strengthen this evaluation process. For evaluation, research-practitioners may obtain three to five ORS administrations before and after the intervention phase until consistent patterns are found. This baseline “allows the analysis to demonstrate that change in a particular client may be primarily attributed to the interventions and is not due to other factors” (Nielson, 2015, pp. 370–371). Using this research design, research-practitioners can control for the effects of AT, therapist/interaction and many specific ingredients. Because RCTs have been an obstacle in AT, experimental designs such as these may help to improve the scientific verifiability called for in AT literature (Gabrielsen et al., 2015; Gillis et al., 2016b; Norton et al., 2014).

Although FIT can be incorporated into any therapeutic approach, FIT does not argue for therapists operating without a structured therapeutic model. What it does suggest is that when adding FIT to existing research-supported treatments, therapist outcomes improve. As clinical expertise has become a trend in psychotherapy (see Chow et al., 2015), researchers may benefit from identifying practical skills adventure therapists can practice to improve their performance. Additionally, therapists can use FIT to establish a baseline for their effectiveness, which they can use to monitor their development. RCTs could explore the effects of FIT with AT participants, as well as help to locate highest achieving adventure therapists. With the quality of research in AT improving exponentially over the last 20 years (Gass et al., 2012), it is promising that ROM can be studied within AT in more depth re-humanizing client experiences and potentially allowing for a more collaborative AT process.

Conclusion

In AT literature, one will find arguments benefiting time spent in nature, the metaphoric framework of the experience, and the length of time adventure therapists interact with their clients, among others, as being the key contributors to change (Fernee et al., 2017; Gass et al., 2012; Norton et al., 2014; Russell, 2001; Russell & Hendee, 2000). Borrowing from over 40 years of outcome research, general factors, such as therapist and client’s characteristics, the therapeutic alliance, and hope, placebo, and expectancy factors seem to play a larger role than the precise differences between each therapeutic approach (Asay & Lambert, 1999; Miller et al., 2013; Orlinsky et al., 1994; Wampold, 2001). Research is showing continually that AT can deliver significant outcomes for adolescents, and those effects can be maintained long-term (Bowen & Neill, 2013; Gass et al., 2012; Gillis et al., 2016b; Norton et al., 2014; Tucker et al., 2016).

This article hopes to present that, instead of finding consensus on what AT is or how it works, accountability in outcomes and partnering with clients can ensure the most efficacious delivery of this experiential approach. One option would be to look at FIT as a method for tailoring each AT experience to the preference, feedback, and unique needs of each AT client and family. Within this framework, improving outcomes can only be accomplished with one adventure therapist partnering with one client at a time.

References

- Adventure Therapy International Committee (ATIC). (2016). *What is international adventure therapy?—International Adventure Therapy*. Retrieved from <https://internationaladventuretherapy.org/what-is-iat/>

- American Psychological Association. (2006). Evidence-based practice in psychology. *American Psychologist*, 61, 271–285.
- Anker, M. G., Duncan, B. L., & Sparks, J. A. (2009). Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, 77, 693–704.
- Asay, T. P., & Lambert, M. J. (1999). The empirical case for the common factors of therapy: Quantitative findings. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 33–56). Washington, DC: American Psychological Association.
- Bachelor, A., & Horvath, A. (1999). The therapeutic relationship. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds.), *The heart and soul of change: What works in therapy* (pp. 133–178). Washington, DC: APA Press.
- Baldwin, S. A., Wampold, B. E., & Imel, Z. E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology*, 75(6), 842–852.
- 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, 106–117.
- Bertolini, B., & Miller, S. D. (2012). *ICCE manuals of feedback-informed treatment* (Vol. 1). Chicago, IL: ICCE Press.
- Bettman, J. E. (2012, May). *Therapeutic outcomes of wilderness therapy for adolescent and young adult populations*. Paper presented at the American Psychological Association Convention, Orlando, FL.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice*, 16, 252–260.
- Boswell, J. F., Kraus, D. R., Miller, S. D., & Lambert, M. J. (2013). Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, 25, 6–19.
- Bowen, D. J., & Neill, J. T. (2013). A meta-analysis of adventure therapy outcomes and moderators. *The Open Psychology Journal*, 6, 28–53.
- Campbell, A., & Hemsley, S. (2009). Outcome rating scale and session rating scale in psychological practice: Clinical utility of ultra-brief measures. *Clinical Psychologist*, 13, 1–9.
- Chow, D. L., Miller, S. D., Seidel, J. A., Kane, R. T., Thurnton, J. A., & Andrews, W. P. (2015). The role of deliberate practice in the development of highly effective psychotherapists. *Psychotherapy*, 52, 337–345.
- Diaz, C., & Drewery, S. (2016). A critical assessment of evidence-based policy and practice in social work. *Journal of Evidence-Informed Social Work*, 13, 425–431.
- Duncan, B. L., & Miller, S. D. (2007). *The Group Session Rating Scale*. Jensen Beach, FL: Author.
- Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R., Brown, J., & Johnson, L. D. (2003). The session rating scale: Preliminary psychometric properties of a “working” alliance measure. *Journal of Brief Therapy*, 3(1), 3–12.
- Ferneer, C. R., Gabrielsen, L. E., Andersen, A. J., & Mesel, T. (2017). Unpacking the black box of wilderness therapy: A realist synthesis. *Qualitative Health Research*, 27, 114–129.
- Gabrielsen, L. E., Ferneer, C. R., Aasen, G. O., & Eskedal, L. T. (2015). Why randomized trials are challenging within adventure therapy research: Lessons learned in Norway. *Journal of Experiential Education*, 39(1), 5–14.
- Gass, M. A., Gillis, H. L., & Russell, K. C. (2012). *Adventure therapy: Theory, research, and practice*. New York, NY: Routledge.
- Gillis, H. L., Kivlighan, D. M., & Russell, K. C. (2016). Between-client and within-client engagement and outcome in a residential wilderness treatment group: An actor partner interdependence analysis. *Psychotherapy*, 53, 413–423.
- Gillis, H. L., Speelman, E., Linville, N., Bailey, E., Kalle, A., Oglesbee, N., ... Jensen, J. (2016). Meta-analysis of treatment outcomes measured by the YO-Q and Y-OQ-SR comparing wilderness and non-wilderness treatment programs. *Child and Youth Care Forum*, 45, 851–863.

- Government Accountability Office (GAO). (2007). *Residential treatment programs: Concerns regarding abuse and death in certain programs troubled youth*. Retrieved from <http://www.gao.gov/new.items/d08146t.pdf>
- Hannan, C., Lambert, M. J., Harmon, C., Nielsen, S. L., Smart, D. W., Shimokawa, K., & Sutton, S. W. (2005). A lab test and algorithms for identifying clients at risk for treatment failure. *Journal of Clinical Psychology, 61*, 155–163.
- Harper, N. J. (2007). *A mixed methods examination of family involvement in adolescent wilderness therapy*. Ann Arbor, MI: ProQuest.
- Hoag, M. J., Massey, K. E., & Roberts, S. D. (2014). Dissecting the wilderness therapy client: Examining clinical trends, findings, and patterns. *Journal of Experiential Education, 37*, 382–396.
- Howard, K. I., Moras, K., Brill, P. L., Martinovich, Z., & Lutz, W. (1996). Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist, 51*, 1059–1964.
- Javorski, S., & Gass, M. A. (2013). 10-year incidence monitoring in outdoor behavioral healthcare: Lessons learned and future directions. *Journal of Therapeutic Schools and Programs, 6*, 80–105.
- Lambert, M. J. (2010). Yes, it is time for clinicians to routinely monitor treatment outcome. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart & soul of change* (2nd ed., pp. 239–266). Washington DC: American Psychological Association.
- Lambert, M. J. (2013). The efficacy and effectiveness of psychotherapy. In M. J. Lambert (Ed.), *Bergin and Garfield's handbook of psychotherapy and behavior change* (6th ed., pp. 169–218). Hoboken, NJ: Wiley.
- Lambert, M. J., & Bergin, A. E. (1994). The effectiveness of psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 143–189). New York, NY: Wiley.
- Lambert, M. J., Burlingame, G. M., Umphress, V., Hansen, N. B., Vermeersch, D. A., & Clouse, G. C. (1996). The reliability and validity of the outcome questionnaire. *Clinical Psychology and Psychotherapy, 3*, 249–258.
- Luborsky, L., Barber, J., Siqueland, L., Johnson, S., Najavits, L., & Frankl, A. (1996). The revised helping alliance questionnaire (HAQ-II). *Clinical Psychology and Psychotherapy, 3*, 260–271.
- Lundervold, D. A., & Belwood, M. F. (2000). The best kept secret in counseling: Single-case ($N = 1$) experimental designs. *Journal of Counseling and Development, 78*, 92–102.
- Miller, S. D. (2011). *Psychometrics of the ORS and SRS. Results from RCTs and meta-analyses of routine outcome monitoring & feedback. The available evidence*. Retrieved from <http://www.slideshare.net/scottdmiller/measures-and-feedback-january-2011>
- Miller, S. D., Duncan, B. L., Brown, J., Sparks, J. A., & Claud, D. A. (2000). The outcome rating scale: A preliminary study of reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy, 2*, 91–100.
- Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2013). The outcome of psychotherapy: Yesterday, today, and tomorrow. *Psychotherapy, 50*, 88–97.
- Miller, S. D., Hubble, M. A., Chow, D. L., & Seidel, J. A. (2015). Beyond measures and monitoring: Realizing the potential of Feedback-Informed Treatment. *Psychotherapy, 52*, 449–457.
- NASW. (2016). *NASW standards for social work practice in health care settings*. Retrieved from <https://www.socialworkers.org/practice/standards/NASWHealthCareStandards.pdf>
- Nielson, T. (2015). Practice-based research: Meeting the demands of program evaluation through the single-case design. *Journal of Mental Health Counseling, 37*, 364–376.
- Norton, C. L. (2010). Into the wilderness—A case study: The psychodynamics of adolescent depression and the need for a holistic intervention. *Clinical Social Work Journal, 38*, 226–235.
- Norton, C. L., Tucker, A., Russell, K. C., Bettmann, J. E., Gass, M., Gillis, H. L., & Behrens, E. (2014). Adventure therapy with youth. *Journal of Experiential Education, 37*, 46–59.
- Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy—Noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270–378). New York, NY: Wiley.
- Quirk, K., Miller, S. D., Duncan, B. L., & Owen, J. (2012). Group session rating scale: Preliminary psychometrics in substance abuse group interventions. *Counselling and Psychotherapy Research, 13*, 1–7.

- Russell, K. C. (2001). What is wilderness therapy? *Journal of Experiential Education, 24*, 70–79.
- Russell, K. C. (2003). An assessment of outcomes in outdoor behavioral healthcare treatment. *Child and Youth Care Forum, 32*, 335–381.
- Russell, K. C., Gillis, H. L., & Heppner, W. (2016). An examination of mindfulness-based experiences through adventure in substance use disorder treatment for young adult males: A pilot study. *Mindfulness, 7*, 320–328.
- Russell, K. C., & Hendee, J. C. (2000). *Outdoor behavioral healthcare: Definitions, common practice, expected outcomes, and a nationwide survey of programs*. Retrieved from <http://www.webpages.uidaho.edu/wrc/Pdf/OBHPublication.pdf>
- SAMHSA. (2012). SAMSHA's national registry of evidence-based programs and practices. Retrieved from <http://legacy.nreppadmin.net/ViewIntervention.aspx?id=249>
- Smith, M. L., & Glass, G. V. (1976). Meta-analysis of psychotherapy outcome studies. *American Psychologist, 32*, 752–760.
- Sparks, J. A., & Muro, M. L. (2009). Client-directed wraparound: The client as connector in community collaboration. *Journal of Systemic Therapies, 28*, 63–76.
- Tilsen, J., & McNamee, S. (2015). Feedback informed treatment: Evidence-based practice meets social construction. *Family Process, 54*, 124–137.
- Tucker, A. R., Widmer, M. A., Faddis, T., & Randolph, B. (2016). Family therapy in outdoor behavioral healthcare: Current practices and future possibilities. *Contemporary Family Therapy, 38*, 32–42.
- Wampold, B. E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Hillsdale, NJ: Lawrence Erlbaum.
- Wampold, B. E. (2010). The research evidence for common factors models: A historically situated perspective. In B. L. Duncan, S. D. Miller, B. E. Wampold, & M. A. Hubble (Eds.), *The heart and soul of change: Delivering what works in therapy* (2nd ed.). Washington, DC: American Psychological Association.