Hierarchical Data Structures in Adventure Education and Therapy

Keith Russell and Jim Sibthorp

Hierarchical linear modeling (HLM) is an approach used in data analysis to better understand how program outcomes are affected by the "nested" nature of data collected in many studies. An outcome can be considered variables such as an individual's self-efficacy, social skills, or more targeted outcomes such as demonstrated reading and mathematical skills. Recent research has suggested that individuals within each "nested structure" may exhibit more similar outcomes than another similar research setting. The purpose of this article is to provide examples of nested data structures and illustrate common approaches to dealing with this type of data often found in adventure education and therapy research. Data available from a study on the wilderness treatment outcomes are then analyzed using HLM to illustrate how the process can increase interpretation of findings and inform future research. Results suggest that many of the variables of interest in research on adventure education and therapy, which might explain why outcomes vary for participants, may be missing from research designs due to nested data structures. Future researchers should consider HLM approaches that may be appropriate for nested data structures common in studies on adventure education and therapy.

Keywords: Adventure education, Adventure therapy, Hierarchical linear modeling (HLM)

Keith C. Russell, Ph.D., is an Assistant Professor in the Department of Kinesiology—Outdoor Education, at the University of New Hampshire where he teaches courses in research methods and outdoor education philosophy and methods. He is also the Director of the Outdoor Behavioral Healthcare Research Cooperative.

Jim Sibthorp, Ph.D., is an Assistant Professor in the Department of Parks, Recreation, and Tourism at the University of Utah where he teaches courses in research methods, experiential education, and commercial recreation.

The purpose of this article is to introduce the concept of "nested data structures" to researchers and practitioners in adventure education and therapy, and also to introduce a data analysis technique called Hierarchical Linear Modeling (HLM) that can addresses the specific challenges these nested data structures present. Hierarchical data structures are common in educational (Williams, 2003) and therapeutic settings (Osborne, 2000), and present a unique challenge to researchers. Some researchers go as far as to advocate the treatment of repeated measures within individuals and meta-analytic data as nested hierarchies (Bryk & Raudenbush, 1992). Several recent articles have appeared in the organizational studies and education literature that support the premise that HLM is more appropriate than using the General Linear Model (GLM) when analyzing nested data (Davidson, Kwak, Seo, & Choi, 2002; Osborne, 2000; Schectman, 2003; Weiss, Harris, Catron, & Han, 2003; Williams, 2003). These suggestions served as the primary impetus to explore HLM techniques in the context of adventure education and therapy. To accomplish this, three primary purposes guide this article: (a) to introduce nested data structures and HLM to the fields of adventure education and therapy, (b) to illustrate how HLM might be used to better understand the processes behind adventure education and therapy programs by analyzing data from a study of treatment outcomes conducted at seven different programs, and (c) to discuss implications of this analysis for future HLM research. The goal is to highlight the HLM technique and discuss the implications for future research that may employ HLM, and to better inform practitioners, researchers, and consumers as to why some programs may, or may not, be effective for participants.

Hierarchical Data Structures: Two Examples

In order to understand what a hierarchical data structure might look like in adventure education, two hypothetical studies are presented. The first study seeks to evaluate the effects of an adventure program that is offered over three months, to two 6th grade classes of 22 students each (see Figure 1). The outcomes being evaluated are social skills which are theorized to be enhanced through participation in adventure experiences. The research design includes a treatment group (adventure education programming) and a control group (no adventure education programming), and was implemented at two different schools. The nested data structure in this example begins with the middle school student, who is "nested" in a classroom with a specific teacher. Each of the two classrooms is also

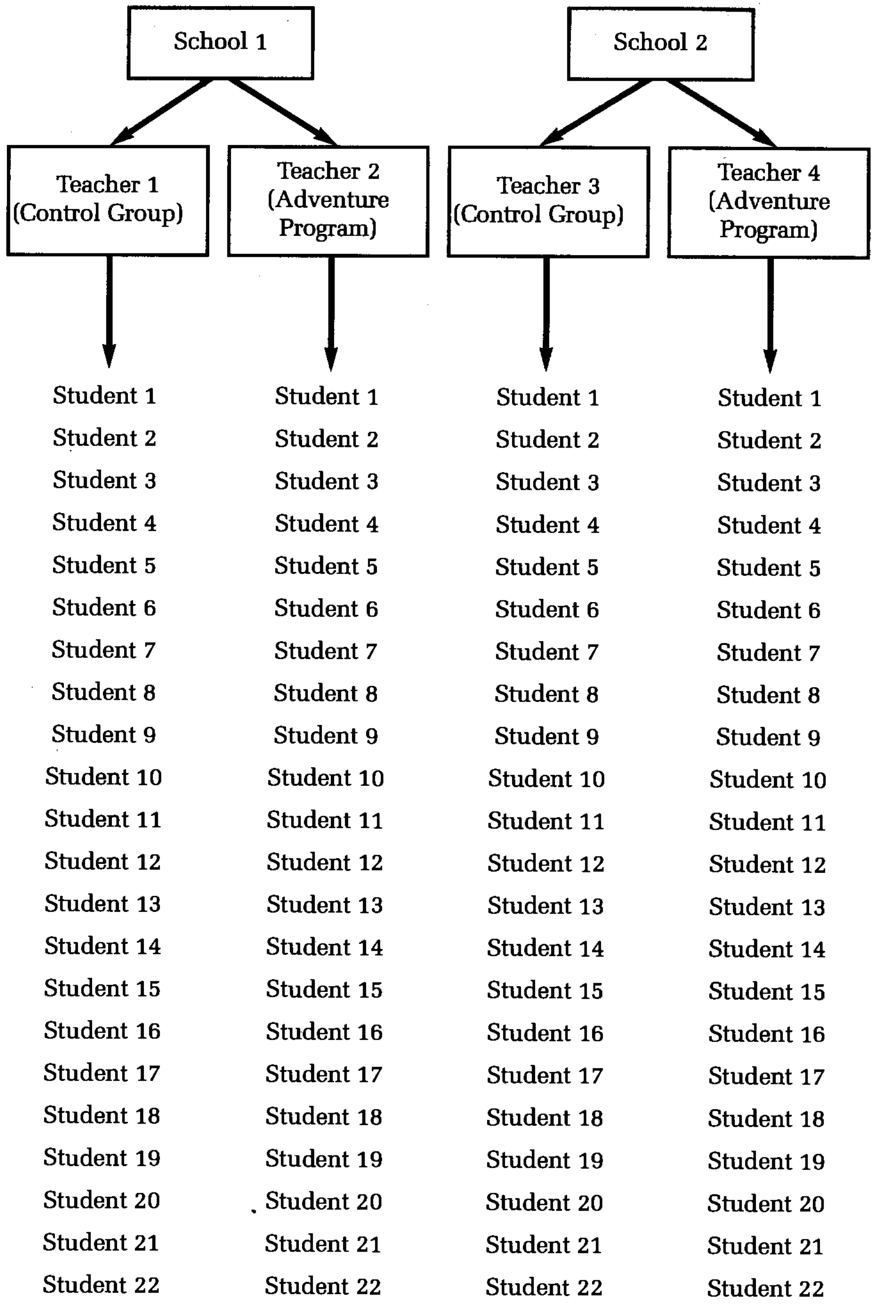

Figure 1. **Hypothetical study of adventure education programming outcomes.**

"nested" in the school within which they reside. It is reasoned that the nested nature of the data collected could impact analyses of association or differences found across treatment and control groups. For example, if the researcher identified an increase in social skills for the individuals that received the adventure programming (treatment), when compared to those individuals in the classrooms that did not receive adventure programming (control), one may conclude that adventure education can be effective in developing these skills.[1]

However, several questions still remain that are related to the nested structure of the data that were analyzed. These include: (a) How did the classroom environment and the style of teaching affect the variable of interest, in this case social skill development of students? (b) How did the school philosophy and curriculum affect the variable of interest? (c) Were the schools located in different regions and did that have an affect on the outcomes of the study? The underlying process of HLM addresses these questions by analyzing outcomes at each nested level of the data. This is based on two important assumptions: (a) that students in each classroom are reasoned to have an experience that is more similar than those in any other classroom (e.g., this can be due to classroom environment and teaching style); and (b) that students in each school are reasoned to be more similar than the students in the other school (e.g., this can be due to adolescent socio-economic backgrounds and educational philosophies at each school). In other words, HLM recognizes that individuals nested within hierarchies may be more similar to one another than if they were placed into various settings using random sampling techniques.[2]

Hierarchical data structures can also be found in therapeutic settings, though the nested nature of the data structures may be more subtle. This can be illustrated using another hypothetical study, this time shifting the focus to adventure therapy outcomes from two outpatient treatment programs for adolescents (see Figure 2). In this example, researchers are interested in examining a hypothesized reduction of negative behavioral incidences in school. Data analysis would test whether or not differences exist in behavioral incidents for adolescents who participated in an adventure therapy program (treatment), or a traditional counseling program (control) delivered by two mental health agencies. The nested data structure in this example, once again, begins with the individual who is nested within a program, who is participating in several on-going treatment groups, being facilitated by different therapists. It is assumed that each therapist may create different group dynamics given his/her therapeutic orientation or style (e.g., cognitive behavioral or family therapy), thus creating similar group dynamics for those participants. Similar to the education example, any differences in outcomes found between the adventure therapy program participants and the traditional program participants

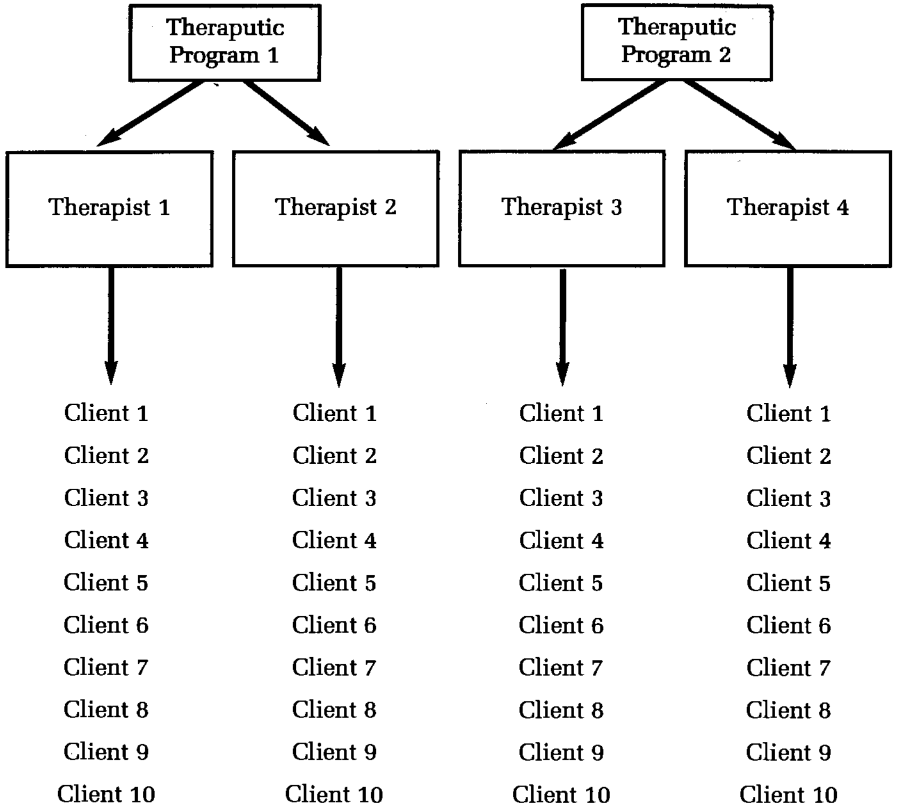

Figure 2. **Hierarchical example from an adventure therapy program.**

may not be due to treatment effects, but rather due to the effects of the "nest" within which they reside. These dynamics can have significant effects on outcomes for each participant. Researchers have addressed these issues prior to the advent of HLM by using the GLM, but are limited in their ability to understand how each level affects the variance in outcomes for individuals, and may be violating key assumptions required with these approaches.

Previous Strategies to Analyze Nested Data

The previous examples and subsequent implications highlight a common error which is made when analyzing data that is nested. This is the assumption of independent observations that is required for most statistical analysis techniques; each participant's process and outcome is assumed to be independent from other participants in the study. To make comparisons and examine associations in nested data structures two approaches are commonly applied. These two approaches are graphically depicted in Figure 2 in order to illustrate how they may violate the independence assumption required for their use and may lead to misleading conclusions based on a loss of statistical power. The first technique occurs when researchers are interested in how program level variables (e.g., philosophy, approach, program design) and group level variables (e.g., group dynamics, therapist) may influence individual outcomes (e.g., relapse; increase in self-concept). The strategy requires researchers to "disaggregate" the data, where the "group" or "therapist" level is ignored and the individual outcomes (e.g., change in self-concept) are compared. In this example, a *t* test could then be conducted to examine differences in self-concept change between program one and program two. The issue with this approach relates to assumptions required for analyzing data using techniques based on the GLM, namely, that all observations are *independent* of one another. If a difference is detected between program one and program two, it could be because the programs are different (program design, philosophy, or approach), *or* it could be because therapists one and two were different than therapist three and four. Ignoring this "group" level variance can lead to conclusions that suggest differences exist when in fact they do not. This error is referred to as a Type I error, a rejection of the null hypothesis (that there are no differences) and concluding that differences or associations *do* exist across programs, when in fact they may be due to "group" level effects (Osborne, 2000).

The other common approach is to "aggregate" the individual level data "up to" the group level. When using this approach, researchers are interested in exploring the effect of program level variables on group outcomes. Using the hypothetical example in the therapeutic settings previously described, we might aggregate individual rates of self-concept up to the group level by using a group average. The main issue with this approach is that most of the variance at the individual level is lost (up to 80-90%), and statistical power is lowered as the degrees of freedom are reduced from the number of individuals to the number of groups (Raudenbush & Bryk, 2002).[3]

One alternative technique is to use HLM in place of GLM applications. Traditional education and therapeutic program research is increasingly using the HLM technique to analyze data structures, like the ones illustrated in the previous hypothetical examples, in order to better understand how process elements affect outcomes (Harachi, Abbott, Catalano, Haggerty, & Fleming, 1999; Latimer, Winters, Stinchfield, & Traver, 2000). HLM is also well suited for research in adventure education and therapy for similar reasons, allowing researchers the opportunity to explore how various process elements found in program level nests may explain the variance in outcomes gathered at individual levels.

How Does HLM Work?

The basic process underlying HLM is similar to Ordinary Least Squares (OLS) regression techniques in the GLM. The easiest way to explain HLM is to discuss a simple two-level model in which a researcher may be interested in individual outcomes nested within a program, and the researcher is studying more than one program. The process begins at the individual level, where an outcome variable is predicted based on a linear combination of variables. Using this process, a separate equation is predicted for each program. Next, HLM uses the parameter estimates generated from the Level-1 analysis (individual) as outcome variables in the Level-2 analysis (program). For example, the equation below represents the Level-1 analysis of a study examining the relationship between gender and treatment outcome of the individual leaving a treatment program across several programs. Subscripts reference individuals (i) and programs (j).

$$\text{Level 1: Outcome}_{ij} = \beta_{0j} + \gamma_{10}\text{Gender} + r_{ij}$$

A separate regression equation for each program is calculated which includes the following parameters: (a) an intercept (β_{0j}), (b) the influence of gender on relapse (γ_{10}), and (c) the error plus participant level variance (r_{ij} is the parameter, σ_2 is the variance estimate for this parameter).[4] This is referred to as an "intercept as outcome" HLM model (Raudenbush et al., 2001). The intercept for each organization becomes the dependent variable in the Level-2 equation. Variables from the Level-1 equation can then be centered around the mean for the organization, creating a program level intercept which represents the mean level of treatment outcome for each program (γ_{00}). This provides an estimation of the amount of variance explained at the program level (μ_{0j} is the parameter; τ_{00} is the variance estimate for this parameter).

$$\beta_{0j} = \gamma_{00} + U_{0j}$$

Finally, a covariate can be added at the program level to determine the effect of, for example, length of program, on the dependent variable:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{Length}_j + U_{0j}$$

The combined two-level model can then determine how much of the variation in treatment outcome is explained by the program (U_{0j}), the length of the program (γ_{01}), and the gender (γ_{10}) of the participants.

$$\text{Level 2: Outcome}_{ij} = \gamma_{00} + \gamma_{01}\text{Length}_j + \gamma_{10}\text{ Genderi} + U_{0j} + r_{ij}$$

There are several advantages to the HLM approach. First, researchers can better understand if a relationship exists between both individual level variables (e.g., age) and program level variables (e.g., length) and treatment outcome. Second, parameter estimates are more stable using HLM with maximum likelihood estimation and precision weighting than in standard OLS regression techniques (Raudenbush & Bryk, 2002). Finally, as each level of analysis is conducted, it is possible to include additional predictor variables and variable interactions across levels that are theorized to explain the dependent variable in the equation, lending increased understanding of what may be driving variance across outcomes.

An Example of HLM

A recent study of treatment outcomes in seven outdoor behavioral healthcare (OBH) programs for resistant adolescents suggested that the treatment was effective at reducing the behavioral and emotional symptoms of adolescents (Russell, 2003). However, the approach used to analyze the data may not have fully captured the multivariate structure of the data. It may have allowed limited insight into interactions, and may have violated the assumption of independence because of the nested structure of the data. Though this research addressed the question of adolescent improvement as a result of treatment, several unanswered questions remain: Were there any differences in treatment outcome attributable to participant characteristics? What role did the group, leaders, or therapists play? Does treatment affect treatment outcome? How do specific program elements associated with wilderness treatment enhance outcomes when compared to more traditional program approaches? Though many of these questions remain out of reach due to the limitations of the study, some can be addressed by employing alternative data analysis techniques like HLM. This inquiry is especially critical as there is a growing need for researchers to begin investigating process factors, like those previously, listed that are most likely to foster personal development through adventure (Ewert, 1989; Hattie, Marsh, Neill, & Richards, 1997; Henderson & Fox, 1994; Warner, 1999).

HLM was used to analyze the data set reported by Russell (2003) to illustrate the additional explanation afforded by HLM analysis by comparing these results to conclusions developed from the original study. This process also offered insight into potential variables that could be examined in similar therapeutic studies that may better explain the variance in outcome. The data included 523 adolescent self-reports, and 372 parent assessments collected at admission and discharge at seven OBH programs. Treatment averaged 45-days in length. The Youth-Outcome Questionnaire (Y-OQ) and Self Report Y-OQ (herein referred to simply as the Y-OQ except where distinction is important) was used to assess outcomes. The Y-OQ offers parent assessment and adolescent self-reports designed for repeated measurement of adolescent emotional and behavioral symptoms (Burlingame et al., 1996; Lambert & Cattani-

Thompson, 1996; Lambert, Ogles, & Masters, 1992; Russell, 2000; Wells, Burlingame, Lambert, Hoag, & Hope, 1996; Wells, 1990). The 64 items contained in the Y-OQ are summed across six content areas to produce a total score.[5] The higher the Y-OQ score, the more serious are the adolescent's symptoms. A significant outcome in this study would be a significant change in scores from admission to discharge indicating the participant had improved.

The data is reasoned to be nested because Y-OQ scores were collected across multiple programs that are similar because they utilize a wilderness therapy approach, but also contain differences in treatment length and approach. The goal is to better understand what characteristics of the programs or adolescents were related to larger or smaller changes in scores. The primary focus of this analysis was to determine the portion of variance in outcome indicated by a change in Y-OQ scores that could be explained by program structure (specifically program length and amount of time in the wilderness), and to investigate the amount of variance accounted for by adolescent characteristics such as sex and age. Hierarchical models for both parent and adolescent (participant) reports on the Y-OQ were examined.

Interpreting the Results

The initial results for a two-level model examining changes in adolescent-reported Y-OQ scores from admission to discharge (Level-1 contains participant variables and Level-2 contains program level variables) indicated that a significant amount of the variance in the sample was attributable to the program level (τ_{00} = 106.1, σ^2 = 681.9, where τ_{00} is the program level variance and σ^2 is the participant level and error variance). This indicates that about 13.5% of the variance in the sample could be accounted for by the program level variables (see Table 1). The percent of

Table 1
Final Estimation of Variance Components

Random Effect	Standard Deviation	Variance Component	*df*	Chi-square	*p*-value
U_{0j}	10.3	106.1 (τ_{00})	6	122.8	< .001
r_{ij}	26.1	681.9 (σ^2)			

Note. U_{0j} is the parameter estimate representing the program level effects. r_{ij} is the parameter estimate representing the individual level effects plus the unexplained variance. Dependent variable is the change in parent Y-OQ score.

variance accounted for by the program level is computed by dividing the variance attributable to the program level (106.1) by the variance attributable to the entire model, which includes program, individual, and error variance (106.1/[106.1+681.9]). Subsequent analyses looked at the importance of sex and age (level-one or participant level predicators) and the number of program days, and number of program days in the wilderness, as potential Level-2 predictors (program level). The model with the most explainable variance included only the number of program days as a Level-2 predictor (t = 5.24, p < .001). This reduced τ_{00} to 34.7 from 106.1, thus indicating that program length accounts for about 67% ([106.1-34.7]/106.1) of the variance attributable to the program level, or about 9% (67% of the 13.5% explained at the program level) of the variance in the sample data. What this means is that program characteristics, such as treatment length and percentage of time in wilderness, explain 13% of variance in outcomes and suggests that the longer adolescents remain in wilderness, the larger the reduction in Y-OQ scores. While the remaining variance at the program level (about 4.5%) remains statistically significant and warrants additional work, perhaps more importantly, 87% of the variance is *not* explained by program level variables, but is likely attributable to other factors including random error.

The initial results for a two-level model of the parent scores (Level-1 contains participant variables and Level-2 contains program level variables) indicated that a significant amount of the variance in the sample was attributable to the program level (τ_{00} = 148.6, σ^2 = 584.6). This indicates that about 20.3 % (148.6/[148.6+584.6]) of the variance in the sample could be accounted for by the program level variables (see Table 2).

Similar to the analysis of adolescent self-reports, the next step was to examine the importance of sex and age (Level-1 or participant level predicators) and the number of program days in the wilderness as potential

Table 2
Final Estimation of Variance Components

Random Effect	Standard Deviation	Variance Component	*df*	Chi-square	*p*-value
U_{0j}	12.2	148.6 (τ_{00})	6	195.43	< .001
r_{ij}	24.2	584.6 (σ^2)			

Note. U_{0j} is the parameter estimate representing the program level effects. r_{ij} is the parameter estimate representing the individual level effects plus the unexplained variance. Dependent variable is the change in parent Y-OQ score.

Level-2 predictors (program level). The most significant Level-2 predictor was the number program days principally in the wilderness ($t = 3.39$, $p < .05$). This reduced τ_{00} to 80.2, indicating that program days in the wilderness account for about 46% ([148.6-80.2]/148.6) of the variance attributable to the program level, or about 9.3% (46% of 20.3% of the variance at the program level) of the variance in the sample data. The 11% of the program level variance remaining unexplained is statistically significant, and could be explained by additional program level variables. Subsequently, sex also explained a statistically significant, albeit small, amount of the variance at the participant level ($t = 2.79$, $p < .05$). (It is impossible to accurately calculate the percent variance for Level-1 predictors using the equations indicated because the denominator includes both Level-1 and all remaining error variance, thus the percent is somewhat meaningless). The positive coefficient suggests that females may experience greater change than males.

Discussion and Implications

The applied example suggests that longer programs, and those spending more time in the wilderness, may have the potential for greater participant impacts in wilderness therapy programs. It also suggests that females may have more positive outcomes than males. Although similar conclusions were reported by Russell (2003), HLM analysis offered slightly different and more detailed explanations of the variance in Y-OQ scores, and did not require an assumption of independence. The original study reported an ANOVA that showed significant differences in score reductions across program models which suggested that longer programs had greater reductions in scores (Russell, 2003, p. 367). HLM analyses also identified these differences but revealed that this program level characteristic accounted for only 9% of all the variance in outcomes. This shows researchers that upwards of 80% of the variance remain unexplained. While, undoubtedly, much of this variance is simply error variance and will never be identified, it is also likely that some of the unexplained variance in this data set is attributable to nested factors or predictors that were not collected in the study. Two of the most promising nested factors seem to be group level data and instructor/therapist level data. While an abundance of literature supports both the role and importance of the specific group (Ewert & Heywood, 1991) and the leader (Propst & Koesler, 1998; Russell & Phillips-Miller, 2002), little research has specifically identified the role each of these plays in the growth process through adventure education or therapy.

Additionally, there may be important participant characteristics or predictors at the individual level. Some of these have traditionally been examined (e.g., age, sex), and others are currently being explored, such as motivations, expectations, and on-program perceptions (McKenzie, 2003; Sibthorp, 2003). While it is possible that individuals are too idiosyncratic and that a meaningful amount of individual level variance will be difficult to track, the amount of unexplained variance in a sample makes this issue more prominent. Thus, to adequately explore the role that participant characteristics play in educational and therapeutic processes, the variance attributable to a sample's nested structure must be identified. If the sources of variance remain unidentified, they are simply relegated to the error term, which reduces the statistical power of the analysis and limits conclusions that can be drawn from studies.

HLM offers a viable alternative for analyzing new data and reanalyzing existing adventure education and therapy data involving nested structures. Other approaches, such as hierarchical regression, hierarchal ANOVA, and two-stage hypothesis testing, also offer alternatives to dealing with nested effects. However, specific techniques embedded in the HLM process, like maximum likelihood estimation and precision weighting, makes HLM more robust when the treatment levels are random and the cell sizes are unequal, common occurrences in field-based adventure education and therapy research.

Regardless of the specific approach, future researchers should consider alternative methods of analysis appropriate for nested data structures common in adventure education and therapy program studies. In addition, data collection across more heterogeneous programs, inclusion of additional nested level variables, and more sophisticated measures at all levels will allow researchers to better understand the "hows" and "whys" of outcomes identified through adventure programming. Ultimately, this will allow research to more effectively dissect the adventure experience and to provide better-informed guidance to adventure programs and practice.

References

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

Burlingame, G. M., Wells, M. G., Hoag, M. J., Hope, C. A., Nebeker, R. S., Konkel, K., et al. (1996). *Manual for youth outcome questionnaire (Y-OQ).* Stevenson, MD: American Professional Credentialling Services.

Davidson, M., Kwak, N., Seo, Y. S., & Choi, J. (2002). Using hierarchical linear models to examine moderator effects: Person-by-organization interactions. *Organizational Research Methods, 5*(3), 231-254.

Ewert, A. (1989). *Outdoor adventure pursuits: Foundations, models, and theories.* Worthington, OH: Publishing Horizons.

Ewert, A., & Heywood, J. (1991). Group development in the natural environment: Expectations, outcomes and techniques. *Journal of Environment and Behavior, 23*(5), 592-615.

Harachi, T. W., Abbott, R. D., Catalano, R. F., Haggerty, K. P., & Fleming, C. B. (1999). Opening the black box: Using process evaluation measures to assess implementation and theory building. *American Journal of Community Psychology, 27*(5), 711-731.

Hattie, J., 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.

Henderson, K., & Fox, K. (1994). *Methods, measures, and madness: Possibilities for outdoor education research.* Paper presented at the Coalition for Education in the Outdoors Second Research Symposium Proceedings, Cortland, NY.

Lambert, M., & Cattani-Thompson, K. (1996, July/August). Current findings the effectiveness of counseling: Implications for practice. *Journal of Counseling and Development, 74*, 601-607.

Lambert, M., Ogles, B., & Masters, K. (1992, March/April). Choosing assessment devices: An organizational and conceptual scheme. *Journal of Counseling and Development, 70*, 527-532.

Latimer, W. W., Winters, K. C., Stinchfield, R., & Traver, R. E. (2000). Demographic, individual, and interpersonal predictors of adolescent alcohol and marijuana use following treatment. *Psychology of Addictive Behaviors, 14*(2), 162-173.

Mckenzie, M. (2003). Beyond the Outward Bound process: Rethinking student learning. *Journal of Experiential Education, 20*(1), 8-23.

Osborne, J. W. (2000). Advantages of hierarchical linear modeling. *Practical Assessment, Research and Evaluation, 7*(1), 1-7.

Propst, D. B., & Koesler, R. A. (1998). Bandura goes outdoors: Role of self-efficacy in the outdoor leadership development process. *Leisure Sciences, 20*, 319-344.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Newbury Park, CA: Sage.

Raudenbush, S. W., Bryk, A. S., Cheong, Y. K., & Congdon, R. (2001). *HLM 5: Hierarchical linear modeling.* Lincolnwood, IL: Scientific Software International.

Russell, K. C. (2000). Exploring how the wilderness therapy process relates to outcomes. *Journal of Experiential Education, 23*(3), 170-176.

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., & Phillips-Miller, D. (2002). Perspectives on the wilderness therapy process and its relation to outcome. *Child and Youth Care Forum, 31*(6), 415-437.

Schectman, Z. (2003). Therapeutic factors and outcomes in group and individual therapy of aggressive boys. *Group Dynamics: Theory, Research, and Practice, 7*(3), 225-237.

Sibthorp, J. (2003). An empirical look at Walsh and Golins' adventure education process model: Relationships between antecedent factors, perceptions of characteristics of an adventure education experience, and changes in self-efficacy. *Journal of Leisure Research, 35*(1), 80-106.

Warner, A. (1999). Improving program quality through evaluation. In J. C. Miles & S. Priest (Eds.), *Adventure programming* (pp. 299-308). State College, PA: Venture.

Weiss, B., Harris, V., Catron, T., & Han, S. (2003). Efficacy of the RECAP Intervention Program for children with concurrent internalizing and externalizing problems. *Journal of Consulting and Clinical Psychology, 71*(2), 364-374.

Wells, M. G., Burlingame, G. M., Lambert, M. J., Hoag, M. J., & Hope, C. A. (1996, Summer). Conceptualization and measurement of patient change during psychotherapy: Development of the Outcome Questionnaire and Youth Outcome Questionnaire. *Psychotherapy, 33*, 275-283.

Wells, S. E. (1990). *At-risk youth: Identification, programs, and recommendations.* Engelwood, CO: Teacher Ideas Press.

Williams, J. D. (2003). Basic concepts in hierarchical linear modeling with applications for policy analysis. In C. J. Cizek (Ed.), *Handbook of Educational Policy* (pp. 473-493). San Diego: Academic Press.

Footnotes

[1] This data would be analyzed under a common set of assumptions upon which are based regression, correlation and analysis of variance procedures commonly referred to as the General Linear Model or GLM.

[2] Random sampling assures that participants selected for participation are no more likely to be in a particular group than any other. The problem with research in the field of adventure education and adventure therapy is that random sampling is rarely used, predisposing study findings to potential bias.

[3] This approach can have significant effects on the power of a study because there may be several individuals (e.g., $N = 100$) but only 8 groups. This reduces the degrees of freedom and subsequently, the power of the statistical analysis. Therefore, when there are differences to

detect across groups or programs, the power of the test may be so reduced that the analysis does not pick up the difference. This causes a Type II error, or failing to reject the null hypothesis when it is, in fact, false. This means there were differences found in the analysis that were not identified.

[4] It should be noted that this example does not include the testing of the hypothesis that the slope of the outcome measure varies at Level-1 or Level-2, and, thus, γ_{10} is treated as a constant coefficient ($\beta_{1j} = \gamma_{10} =$ a constant). While inclusion of such a hypothesis is appropriate, it makes already complex equations even more complicated and could add the presence of an interaction term between the Level-1 and Level-2 predictor variables. Despite the potential value in such an analysis, the authors agreed that a simpler, easier to understand model was most appropriate given the purposes of this paper. Those interested in learning more about varying slopes and slopes-as-outcomes models are directed to Raudenbush & Bryk (2002).

[5] (a) *Interpersonal Distress:* Assesses change in emotional distress including anxiety, depression, fearfulness, hopelessness, and self harm; (b) *Somatic:* Assesses change in somatic distress typical in psychiatric presentation, including headaches, dizziness, stomachaches, nausea, and pain or weakness in joints; (c) *Interpersonal Relations:* Assesses change in the child's relationship with parents, other adults and peers, as well as the attitude towards others, interaction with friends, aggressiveness, arguing, and defiance; (d) *Critical Items:* Assesses inpatient services where short term stabilization is the primary change sought—changes in paranoia, obsessive-compulsive behavior, hallucination, delusions, suicide, mania, and eating disorder issues; (e) *Social Problems:* Assesses changes in problematic behaviors that are socially related, including truancy, sexual problems, running away from home, destruction of property and substance abuse; (f) *Behavioral Dysfunction:* Assesses changes in a child's ability to organize tasks, complete assignments, concentrate, handle frustration, including items on inattention, hyperactivity, and impulsivity.