Observing Outcomes in Youth Development: An Analysis of Mixed Methods

Karla A. Henderson
Gwynn M. Powell
Margery M. Scanlin

EXECUTIVE SUMMARY: Organizations such as park and recreation departments, not-for-profit youth groups, churches, and resident and day camps offer opportunities to promote positive youth development through structured recreation activities. Researchers involved with organized camp programs have documented some values of camp experiences relative to growth and development, yet more is to be learned about how camp structures and settings influence positive youth development and how the outcomes of youth development can be measured. The purpose of this analysis was to explore indicators of youth development outcomes through a comparison of two forms of data collected at six camps. The results of comparing quantitative and qualitative data sources to determine if indicators of positive youth development can be observed in a short period of time in the recreational and educational setting of summer camp is described. Quantitative data came from a study that sampled families representing American Camp Association (ACA) accredited camps from across the United States. The data included pre- and post-questionnaires given to campers to measure domains such as positive identity, social skills, positive values, and thinking and physical skills. The instrument used was called the Camper Growth Index-Children (CGI-C). Qualitative data came from on-site observations in six camps that had participated in the quantitative study the prior year. A guided outline was used as the basis for field observations and informal interviews. The observation rankings/ratings and the quantitative results were compared. Two of the three camps showing statistically positive developmental change in campers also were ranked higher in using the focused qualitative observation. Two of the three camps showing no statistically significant change were also identified as ranking lower through the observation process. Observing opportunities for adventure and exploration activities as well as leadership and independence displayed by campers were easiest to observe. Observing indicators for developmental outcomes, including environmental awareness, positive identity, and positive values, were more difficult. Although consistent agreement did not exist in this comparison of quantitative and qualitative data, the two approaches provided some convergence and complementary data. This study provided an opportunity to explore the measurement of camp experiences from an external view along with an internal self-report approach. We examined the micro data obtained from the individual campers in relation to the social environmental macro structure of the camp. Implications exist for triangulating data and validating methods in other recreation organizations to better understand how and why youth development programs in recreation and camp settings

work. Given the complexity of the desired outcomes and the individual nature of growth and development, mixed methods and systematic multi-measure approaches offer information for supporting youth development.

AUTHORS: Karla A. Henderson is a Professor in the Department of Parks, Recreation and Tourism Management, North Carolina State University, Box 8004 Biltmore, Raleigh, NC 27695-8004. E-mail: karla_Henderson@ncsu.edu. Phone: 919-513-0352A. Gwynn M. Powell is with the University of Georgia, and Margery M. Scanlin is with the American Camp Association.

Introduction

Youth development encompasses efforts to create organizations and communities that enable youth to move toward adulthood by supplying supports and opportunities necessary to go beyond problem prevention (Eccles & Gootman, 2002; Witt, 2002). Organizations such as park and recreation agencies, not-for-profit youth groups, churches, and resident and day camps offer valuable opportunities to promote positive youth development through recreation activities. Nicholson, Collins, and Holmere (2004) underlined that youth development organizations have a common commitment to young people's physical, emotional, and educational growth. Yet little has been done to document the outcomes and positive change that may occur through nonschool programs at recreation centers and camps.

Researchers (e.g., Eccles & Gootman, 2002; Gambone, Klem, & Connell, 2002; Lerner, Lerner, Almerigi, & Theokas, 2005; Roth & Brooks-Gunn, 2003) have described the outcomes associated with positive youth development, such as social relationships with peers and adults, opportunities for cognitive and physical skill building, personal identity, and physical and psychological safety. Lerner et al. (2005) suggested that these outcomes occur when a youth program has positive and sustained adult–youth relationships, youth skill-building activities, and opportunities for youth participation and leadership. Organized camp experiences sponsored by youth agencies, churches, and other independent organizations have provided these structures for decades. Camp experiences have the potential to be important venues for positive youth development. Therefore, the purpose of our study was to explore indicators of youth development outcomes through a comparison of two forms of data collected at six camps.

Background

Youth development has been inherent in the recreation movement since the start of the 20th century. Jane Addams (1998) wrote extensively about the need to keep children off the street and to provide wholesome

activities for them. Although the focus on youth development in parks and recreation and other related fields has never disappeared, it has been given renewed importance in the past 15 years (Caldwell, 2000). The research on youth development seems to have evolved through emphases on program inputs and outputs (e.g., Henderson & King, 1998; Witt & Crompton, 1996) to benefits (e.g., Allen, Stevens, & Harwell, 1996) and today to how outcomes (e.g., Baldwin, 2000) intersect with all aspects of youth programming. The concentration of youth development research in parks and recreation is in documenting the outcomes of recreation "beyond fun and games" (Witt & Crompton, 1996).

Not only has the focus of the youth development literature changed to encompass the dimensions that contribute to successful programs and youth development, but the language has also changed. The population under study in the 1990s was "youth at risk," which generally referred to young people who were likely to participate in negative behaviors (Witt, 2002). This terminology evolved to reflect notions that some youth come from "at-risk communities" where they do not get the supports and opportunities necessary to model positive behaviors. Today's philosophy suggests that all youth are at risk, although some youth are more so than others (Witt, 2002). High rates of boredom, alienation, and disconnect from meaningful challenges and activities result in a deficiency of positive development (Larson, 2000). Positive youth development is the term of the 21st century and is a perspective that provides a strength-based conceptualization of adolescence (Lerner et al., 2005). The focus is not on the problems of youth but on the ways that their resources and community resources can be developed to help young people grow into successful adulthood.

Positive Youth Development

For programs to be successfully implemented and evaluated based on desired outcomes, a focus on theory is needed (Peterson, 2004). Although not necessarily a single theory, a conceptualization of positive youth development provides a theoretical framework for examining developmental outcomes. The rationale for positive youth development, according to Lerner et al. (2005), emanates from contemporary developmental systems theories. These theories suggest that change is a consequence of mutually influential relationships between the developing person and aspects such as biology, psychological characteristics, family, community, and culture.

Developmental theories have guided the measurement of outcomes. One example is the work done by the Search Institute (Leffert et al., 1998). This approach has directly eschewed the problem-focused paradigm relative to youth and children and has replaced it with a focus on human development. The Search Institute has identified 40 assets or building blocks that appear to be the foundation for healthy and positive development in youth. These assets are grouped into two domains: external and internal. The internal assets include commitment to learning, positive

values, social competencies, and positive identity, while the external assets include support, empowerment, boundaries and expectations, and constructive use of time.

Other researchers (e.g., Catalano, Bergland, Ryan, Lonczak, & Hawkins, 2002; Eccles & Gootman, 2002; Gambone et al., 2002; Lerner et al., 2005; Roth & Brooks-Gunn, 2003) have identified necessary elements and protective factors that contribute to youth development in other ways in addition to the Search model. For example, Roth and Brooks-Gunn and Lerner et al. discussed five Cs of positive youth development that seem to be evident in youth studies: competence, confidence, character, connection, and caring. A report prepared by the National Research Council and the Institute of Medicine, Community Programs to Promote Youth Development, underlined the need to use a developmental framework to help youth acquire personal and social benefits (Eccles & Gootman, 2002). Among the elements described in this report were physical and psychological safety, emotional moral support, supportive adult relationships, opportunities to form close human relationships, a feeling of belonging and being valued, opportunities for skill building, personal efficacy, and opportunities to contribute to one's community. Catalano et al. (2002) described features of positive youth development programs that included such objectives as fostering self-determination, self-efficacy, clear and positive identity, and belief in the future. Gambone et al. concluded that the common outcome areas were physical and cognitive learning, social relationships, positive values, and positive identity.

Offering programs such as recreation or camp experiences, however, does not automatically result in the achievement of intended outcomes. Linkages between program elements and desired outcomes are needed (Baldwin, Caldwell, & Witt, 2005). Catalano et al. (2002) concluded that the two most important features needed for effective programs were structured program guidelines and sufficient time. A wide variety of positive youth development approaches exist and are being studied, such as after-school (e.g., Baker & Witt, 2000), not-for-profit youth groups (e.g., Carruthers & Busser, 2000), youth sports (e.g., Larson, 2000), and extracurricular activities (e.g., Eccles & Barber, 1999).

Organized camp programs are a type of recreation program that has focused on youth development in various ways for 150 years. Researchers (e.g., Bialeschki, Younger, Henderson, Ewing, & Casey, 2002; Brannan, Arick, Fullerton, & Harris, 2000; Chenery, 1991; Dworkin, 1999; Marsh, 1999; Sekine, 1994) have documented the benefits of organized camp experiences, and a good deal of anecdotal evidence supports the "good" that camps provide for young people. Many individuals, including ourselves, have been positively influenced by camp experiences. Yet more information is needed about how camp structures and programs influence positive youth development and, further, how to measure youth development outcomes.

Measuring Positive Youth Development

Lerner, Lerner, DeStefanis, and Apfel (2001) argued the need in youth development research for both individual and contextual levels of analysis, which can be accomplished by triangulation between quantitative and qualitative data. Researchers are interested in instances of change that are systematic, organized, and successive rather than random and disorganized. However, Lerner et al. (2001) noted that "Change and, especially, developmental change are difficult to study empirically" (p. 10). Issues arise related to the nature of development, units of analysis, and the role of time and temporality (history) in indexing change. In addition, measures must be sensitive to detect change, and appropriate designs should be employed to increase the validity of the measures. Lerner et al. (2001) further suggested that "a focus on process and, particularly, on the process involved in the changing relations between individuals and their contexts, is the cutting edge of contemporary developmental theory" (p. 10).

One aspect of evaluation is the application of "prevention science" to the concept of positive youth development (Bumbarger & Greenberg, 2002). This approach was embodied in the work of Catalano et al. (2002) in moving positive youth development from a vague concept to a scientifically measurable set of characteristics. Although the gold standard for evaluation measurement appears to be true experimental designs, research is also needed that focuses on the processes and qualitative dimensions associated with program effectiveness. Just as a wide variety of youth development approaches exist, a variety of ways of measuring positive youth development must continue to be explored.

The purpose of our study was to explore indicators of youth development outcomes through a comparison of two forms of data collected at six camps. An analysis was conducted to ascertain how the results of quantitative data measurements could be corroborated by qualitative observations. In other words, we were attempting to focus on what was happening in camps, and whether indicators of positive youth development could be observed. As Erzberger and Prein (1997) suggested, using a multimethod approach to evaluation can increase the validity of examining youth development outcomes in an activity such as camp.

Qualitative and Quantitative Data

The debate regarding whether qualitative or quantitative approaches are best is no longer viable in most areas of social science (Henderson, 1991). Most researchers agree that a place exists for both types of data. The focus today for many researchers pertains to how the two approaches might be used together (e.g., Sale, Lohfeld, & Brazil, 2002; Strauss & Corbin, 1990).

Various opinions exist relative to how qualitative and quantitative data can co-exist in the form of mixed methods, multimethods, linking, or triangulation. Although these words may not mean exactly the same thing, they all suggest the use of more than one method or type of data, or more than one source of data. One of the issues when triangulating methods or

data is the timing of the data collection. Data might be collected concurrently or sequentially. Sometimes data are collected in one way to provide information for a subsequent data collection such as using qualitative focus groups to generate ideas for quantitative questionnaire development.

Some purists (e.g., Sale et al., 2002) argue that using qualitative and quantitative data together represent different paradigms and cannot easily be linked. At best, the data can only be linked for complementary purposes and not cross-validation. Other researchers have advocated over the years that mixed methods may be the best way to do program evaluation (e.g., Henderson & Bedini, 1995; Howe & Keller, 1988; National Science Foundation, 1997; Nicholson et al., 2004). For example, Nicholson et al. noted that relative to youth development, rigorous quantitative studies allow researchers to identify change, but qualitative data help to determine what works. Domitrovich and Greenberg (2000) argued that decisions regarding program adoption should not rely solely on quantitative data regarding effectiveness but also on the quality of delivery. The practical integration of methods in most cases means that quantitative and qualitative data collection and analysis occur concurrently or sequentially in the same research setting.

The major rationale for using mixed methods is to better understand the world. All methods, however, must share a commitment to rigor. The complexity of data in an area such as youth development research requires many perspectives. Another rationale given for using mixed methods is to enhance both the reliability and validity of a research project. Haberman (2000) used the example of sitting in one's office and putting samba music on, and then filling out a rating scale that assesses the grace, rhythm, and creativity of people who are walking outside. The problem is that the categories are being imposed on the events. If enough researchers use the same instrument uniformly, that measure probably is reliable even though it may not be valid. Similarly, a researcher can measure outcomes that occur in young people as a result of participation in an activity, but that measurement may not tell what these results mean or how they occurred. By linking methods, more can be learned about a phenomenon. Christenson and Ysseldyke (1989) advocated the need to move from simple description to multiple sources to design effective interventions. Sale et al. (2002) described how mixed methods could be used for complementary purposes. They can provide an additive aspect if they are philosophically and practically sound.

The use of triangulation can be done in a variety of ways that may result in data convergence, complementarity, or dissonance depending on the results of the research. Convergence relates to confirming previous results. According to Sale et al. (2002), complementarity relates to using a macro and micro analysis that can be combined like a jigsaw puzzle. Dissonance occurs when two perspectives do not fit together. Dissonance leads to asking questions regarding sample bias, misunderstood questions used in measurement, or misinterpretation of qualitative data.

One caution of using mixed methods is that "achieving similar results may be merely a matter of perception" (Sale et al. 2002, p. 47). Both quantitative and qualitative approaches can be colored by interpretivism (Howe, 1992). One of the problems with using data sequentially is that if the results of the first data collection are known, the interpretation of the results could be biased. Erzberger and Prein (1997) suggested that if the methodologies used are rigorous and trustworthy, then the way that results fit together might be a matter of considering the underlying theoretical assumptions. Thus, using mixed methods, regardless of the relationship of the data, would be a means to raise more provocative questions about the measurement of outcomes leading to positive youth development.

Using both qualitative and quantitative data, we sought to explore indicators of youth development outcomes through a comparison of quantitative and qualitative data. We used a retrospective process of examining quantitative and qualitative data collected sequentially to ascertain the complementarity of data.

Methods

We describe this study as an experiment, although not in the sense of an experimental design, to explore whether field researchers can observe daily life at a camp to ascertain what indicators might be associated with youth development outcome measures. The data included results from a quantitative questionnaire developed to measure selected positive youth development outcomes in campers as well as qualitative field observations conducted in six of the camps where quantitative data were collected. A limitation of the study is that the data were not collected concurrently. Nevertheless, we believe, for this exploratory study, that the sequential and yet "blind" observations provided some sense of whether indicators of youth development could be measured.

Quantitative Data

Quantitative data for this national study were collected in 2002 from a sample of families representing American Camp Association (ACA) accredited camps from across the United States. The participating camps were selected from a multistage stratified random sample of camps in different parts of the country representing both day and resident camps from different sponsorships (agency, religious, independent nonprofit, and independent for-profit). The study design included pre-, post-, and follow-up questionnaires given to campers and parents, a pre- and post-observation checklist completed by staff during the session that the camper attended, and questionnaires to camp directors regarding the characteristics, outcome goals, and operations of their camps. The quantitative data used in this paper include the data from the pre- and post-test of outcomes as collected from only the campers.

The development and validation of the camper questionnaire that we called the Camper Growth Index-Children (CGI-C) along with all the other measures used in the national study, are discussed elsewhere (Ameri-

can Camp Association, 2005). An initial survey draft was developed by examining existing and widely used surveys and instruments in the field of youth development (e.g., Search Institute, Camp Fire Boys and Girls, YMCA, Girl Scouts of the USA, Boy Scouts of America). In addition, ACA staff conducted visits to 20 camps in 2000 and informally observed, talked to staff and campers, and documented what appeared to be outcomes at camp (ACA Research Interim Report, 2000). Based on this information and the review of the literature, four major domains were ascertained to contribute theoretically to positive youth development in camp settings: Positive Identity, Social Skills, Positive Values and Spiritual Growth, and Thinking and Physical Skills. TheThe CGI-C initially consisted of 130 items and was pilot tested in the spring of 2001. The final questionnaire consisted of 10 constructs representing the four domains that were measured by 52 questions using a 4-point Likert scale with 4 = *strongly agree* to 1 = *strongly disagree.*

The four domains along with the 10 constructs measured and their reliabilities included: Positive Identity: (1) positive identity (.75), (2) independence (.63); Social Skills: (3) leadership (.76), (4) making friends (.69), (5) social comfort (.66), (6) peer relationships (.71); Positive Values and Spiritual Growth: (7) positive values (.76), (8) spirituality (.81); and Thinking and Physical Skills: (9) adventure/exploration (.66), and (10) environmental awareness (.76). Nunnally and Bernstein (1994) suggest that an alpha level of at least .70-.80 is sufficient for most purposes. Some of our reliability scores were not as high as we would have liked, but we believe they are acceptable relative to the efforts taken to develop valid measures in an area that had not been previously researched on a large scale.

Campers (N = 5,281) who participated in this national study completed the CGI-C pre-test up to 4 weeks prior to coming to camp. The post-test was completed on one of the last days at camp. The camp session length ranged from 5 days to 7 weeks among the camps studied. The response rate on the post-test was 64% for a total of 3,400 campers. In addition to the analysis done on the entire sample, each camp director was also given a summary for his or her camp. Matched t-tests were used to compare the pre and post data for the campers. Along with the t-test, effect sizes were calculated to determine the magnitude of the change that occurred overall and in each camp for each of the items. Positive change occurred in all four domains with statistically significant positive change in six construct areas for the entire sample: adventure/exploration skills, making friends, positive identity, independence, leadership, and spirituality. No differences were found between changes and the length of time (i.e., 5 days up to 7 weeks) that campers spent at camp.

Qualitative Data

In-camp observations were made in July of 2003 in six of the 40 camps that had participated in the quantitative study in 2002. ACA's Executive Officer of Research and Intellectual Resources chose the six camps based on their geographic location (i.e., all camps were within a day's drive of each

other either in Pennsylvania or the New England area), the results of the quantitative data collection done the previous summer (i.e., three camps were chosen in which campers had made positive change on several of the outcome measures and three camps were chosen representing no statistically significant changes), and to represent private independent, not-for-profit agencies, and religiously affiliated camps. Ideally, the quantitative data and the qualitative data would have been collected in the same camp session. However, for purposes of this analysis, we wanted to observe contrasting camps that had shown either great change or no change using the quantitative measures the summer before. We recognized that no youth program is the same each time it is delivered, but with camp director, type of camp, camp philosophy, and program efforts held relatively constant from year to year, we believed a sequential data collection rather than the ideal concurrent collection method did not compromise the validity of this exploratory study. The type of staff, the training that they receive, and the overall program of most camps does not change markedly from year to year.

Two researchers spent a day in each of the six camps observing programs, informally interviewing staff and campers, and talking to the camp director and administrative staff about their camp programs. The researchers attended the camp activities including meals, cabin rest time, activity periods, and visits to the staff lounge. Campers were asked about their week(s) at camp and what they had enjoyed, how they were treated by staff, and what they thought they were learning at camp. The two researchers who went into the camps did not know the results of the 2002 quantitative measurements.

Each of the six camps was contacted by the Executive Officer from the ACA office and invited to participate. The camp directors were told that the visit by the researchers was a follow-up to the data collection done the previous summer. The Executive Officer indicated that the researchers would visit and observe camp programs as well as talk to staff and campers. As researchers/visitors we did not wish to be intrusive and asked only that the camp director or another camp administrator provide a tour of the camp and facilitate our opportunity to chat informally with campers and staff. We arrived shortly after breakfast at each camp and spent the entire day at the camp into the late afternoon. For some of the time we were together, observing the same things and talking to the same people, while at other times we split up to talk to different individuals and to visit different activity venues at the camp. In the camps, we were given as much freedom as we wanted to go where we wished and speak to whatever campers and staff we could find as long as we did not disrupt a program. We made field notes during the day, but primarily observed and listened. We used a 12-question outline as a basis for observations and notes. The guiding questions outlined for the qualitative observations for this research and the aspects that could be observed were taken from the work of Eccles and Gootman (2002), who examined program outcomes from a developmental framework:

1. What is the result of the alignment (or lack of same) between desired outcomes, the program planned/intended, and the methodology/delivery of intended program elements?

   This question was not easy to address, although we had access to written materials for each camp that highlighted the camp's desired outcomes. We also asked questions to campers and staff (counselors) about what outcomes they thought were occurring in the camp. Other observations included the amount and intensity of the program engagement of campers, the transparent reasoning offered from directors and staff for why the program was as it was, the quality of the program activities in terms of congruency between what staff and campers wanted, and the physical structures (e.g., activity opportunities, layout of the camp, unit organization) that enabled certain kinds of interaction.

2. What are the evidences of the efforts and results in creating an atmosphere of physical and emotional safety for campers and staff?

   Because each of these camps was accredited by ACA, some assurance existed that the camp had passed standards for physical safety. Other observed measures that we used were the amount and level of supervision that existed in program activities and free time, social interactions between staff and campers and among campers, the opportunities youth had to choose the level of participation with which they felt comfortable, and the means for travel between activities. Note: For the ratings done by the researchers, physical and emotional safety were considered two separate categories. The physical safety of campers was addressed through the objective ACA accreditation process, but emotional safety was a subjective issue.

3. How strong and frequent are the demonstrations of emotionally supportive relationships? How is the emotional competence of staff to provide a supportive environment monitored and encouraged?

   This question was one of the easier ones to observe. We listened to and observed the interactions occurring, such as hugs, easy laughter, relaxed tone of interactions, and the knowing of and calling of first names. We also noted how staff related to one another and the levels of support they offered one another as well as the feedback they received from camp administrators. Note: For purposes of the ratings, this question was divided into five categories: camper–camper, camper–staff, staff–staff, administration–staff, and opportunities to belong.

4. What differences exist in the delivery of the camp program between various age groups, and are these differences age appropriate?

This question was relatively easy to observe based on how campers were divided in terms of living arrangements and activity groups. We observed aspects such as size of activity groups, age groupings, and camper-to-staff ratios. We also observed different skill levels and the choices campers had related to the range of opportunities available.

5. Is the program delivery well organized, with clear limits and consistent expectations that are enforced by all staff?

This question was observed through the daily schedule. Some campers had similar schedules each day, while in other camps and with different age groups, more choices were available. We also observed and talked to campers about the "rules" at camp. The relationships between the campers and staff (e.g., amount of time spent together, the way staff talked to campers, and the way campers talked about their counselors) were observed and discussed while at camp.

6. Do camp staff and camp programs demonstrate cultural competence and sensitivity to inclusion? What is the evidence that describes the sense of belonging that exists in this camp?

This question was easy to address superficially, but difficult to observe in a one-day visit. The diversity of campers in terms of nationality, race, people with disabilities, tolerance for sexual orientation, staff training, and evidence of scholarship programs available for low-income campers were external measures. The feelings that campers had about inclusion were not as easy to identify, even when campers were queried. Sense of belonging was evident through the wearing of camp clothing, the way that campers connected (e.g., e-mail, visiting one another's homes) when they were away from camp, the number of small groups together during free time, and the campers' descriptions of the importance of camp traditions.

7. What are the demonstrations that there are high, clear standards for staff behavior and youth behavior? Are there identifiable norms?

This question was not easy to observe, but a subjective "feeling" was evident as one walked around the camp. Staff and campers paying attention to one another was one possible indicator, along with the use of language and the use of behavior management with campers. The observation of campers, as well as staff, having fun was another measure.

8. To what degree do staff and campers have the opportunity to make decisions that matter or that affect them in the camp community?

This question was primarily answered by talking to campers and staff and observing the typical camp schedule. Choices in activities and of food were examples of campers having decisions that they could make.

9. What evidence exists that there is a results-oriented staff improvement process in place?

This question was difficult to observe. We asked staff about their opportunities for improvement relative to such aspects as training and feedback. Camps that had counselor-in-training programs (i.e., grow-your-own staff programs) were an example of how this question was addressed.

10. Is there a sustainable and obvious effort at self-improvement for the camp? Are there efforts to identify what changes are needed based on identified factors and with identified potential impact?

This question was also difficult to measure in a one-day visit. We asked staff what input they had to program objectives and how the camp program was determined. The use of Camper Councils and the professional involvement of administrators were other possible indicators of how self-and camp improvement could occur.

11. Do campers and staff come away from the camp experience with improved social and cultural capital? Do they learn physical, emotional, cultural, and social skills?

Elements of this question, especially as it related to some of the measurement categories in the quantitative questionnaire, could be observed. The social skills as well as the physical skills expressed were obvious in terms of what campers said about meeting friends and learning to do new activities. Note: This category was divided into the two categories of program skills and social skills when the researchers ratings were done.

12. Is there an identifiable focus on lifelong values consistent with family and community needs?

Measuring personal values was not easy to do, but some indication was seen in the appreciation for different activities, being in the outdoors, patriotism expressed at camp, and positive interactions as an expectation of the camp.

Each night after our visit to a camp, we spent 2-3 hours writing notes addressing the 12 questions and recording facts and observations about our visits. We recorded as much data as possible in the form of field notes. At that point, together we discussed broad ideas about whether the 2002 camp study might have measured particular outcomes for each camp we visited.

Upon completing the visits and returning home, we wrote narratives (memos) describing the camps. We then independently rated each camp on a matrix of high, average, or low and then ranked them 1 to 6 based on our observations and how the camps appeared to perform in relation to one

another. In addition, we rated and ranked what our determinations were of the four ACA domain outcomes of Positive Identity, Positive Values, Social Skills, and Thinking and Physical Skills based on the observations. Our purpose was not necessarily to compare the camps, but by doing the ranking, we could make sure that our subjective ranking was consistent with the ratings that we had undertaken. If disagreement existed between the two of us, it was indicated as an average between the two. In no case did we as researchers differ by more than one ranking measure. We are aware that by doing the rating and ranking, we turned qualitative data into a form of quantitative data. However, we felt this step was necessary if we were going to "see" outcomes with any reliability.

Although the methods we used were systematic and adequate for the time we had, we believe that being in each camp for only one day was clearly a limitation of this study. The ways that we were treated in the camps and the "feeling" of each camp certainly contributed to our perceptions. Although we believe our observations were valid and consistent with each other, our biases (e.g., beliefs in the value of youth-centered programming, primacy of environmental awareness as important in camp programming) could not help but enter into this evaluation. Often, as we recorded our notes each evening, we came upon topics that we wished we had explored in more depth. Nevertheless, these observations provided a way to experiment with the convergence, complementarity, or dissonance of a mixed-method approach in assessing indicators of positive youth development when we compared these observations to the objective data.

Results

After we had completed the observations and written the final camp visit report, the ACA's Executive Officer shared the results of the quantitative data analysis. The results of the paired t-tests for the outcome scales for the six camps are presented in Table 1. The table indicates that three of the camps had three or more quantitatively measured positive change constructs. It also shows three camps that quantitatively had no change, either positive or negative.

The next step was to compare the observation rankings/ratings and the quantitative results in terms of positive outcomes, as shown in Table 2. This table shows the effect sizes of the statistically significant quantitative results as well as the ratings and rankings that we generated from our observations. When the quantitative results were compared to the qualitative summary, two (Camp #2 and Camp #3) out of the three camps showing statistically positive change were also rated as the highest by qualitative observation. Two (Camp #5 and Camp #6) out of the three camps showing no change were also identified as the lowest two camps in the observation. The remaining two camps (Camp #1 and Camp #4) were rated in the middle by our observations. Camp #1, in fact, had the most measured quantitative change of all six camps, and Camp #4 showed no change.

Table 1
Paired *t*-tests for the Pre- to Post-Camper
Outcomes Questionnaires for Six Camps

Outcome	Camp											
	1[a]		2[b]		3[c]		4[d]		5[e]		6[f]	
	pre	post	pre	post	pre	post	pre	post	pre	post	pre	post
Positive Identity	3.61	3.78*	3.67	3.74	3.70	3.71	3.76	3.69	3.70	3.73	3.80	3.75
Independence	3.62	3.66	3.46	3.63*	3.48	3.55	3.31	3.36	3.46	3.48	3.53	3.50
Leadership	3.05	3.24*	3.24	3.36*	3.29	3.34	3.01	3.10	3.17	3.18	3.31	3.24
Making Friends	3.43	3.70*	3.44	3.44	3.57	3.66	3.29	3.43	3.27	3.31	3.33	3.23
Social Comfort	3.34	3.42	3.20	317	3.26	3.21	3.17	3.17	3.40	3.31	3.62	3.45
Peer Relationships	3.78	3.81	3.62	3.64	3.74	3.77	3.71	3.68	3.61	3.63	3.67	3.65
Adventure/ Exploration	3.45	3.78*	3.60	3.74*	3.38	3.62*	3.69	3.76	3.53	3.36	3.32	3.35
Environmental Awareness	3.47	3.59*	3.84	3.80	3.51	3.63*	3.89	3.85	3.85	3.74	3.66	3.57
Positive Values	3.43	3.52	3.45	3.47	3.48	3.62*	3.44	3.52	3.44	3.44	3.35	3.36
Spirituality	2.89	3.02	2.90	2.97	2.90	2.90	3.49	3.35	3.25	3.28	2.28	2.58

* $p < .05$.

[a] n = 42
[b] n = 48
[c] n = 34
[d] n = 36
[e] n = 25
[f] n = 25

At Camp #3 we observed that the items we rated as *high* related to positive youth development such as psychological safety, appropriate age structure, supportive camper–camper relationships, supportive staff–camper relationships, support for efficacy, social skills, positive identity, and relationship between reported goals and intended outcomes. The quantitative data revealed that adventure/exploration, environmental awareness, and positive identity showed positive change from pre- to post-camp. In the narratives about the camp, we identified the beautiful environmental backdrop that existed for the activities undertaken at camp and the opportunities that campers had to progress from beginners to experts in various camp activities.

For Camp #6, a camp that had no change as a result of quantitative measurement, we observed many activities occurring and a safe, structured program, but little intensive interaction between the staff and campers. We did not feel uncomfortable at the camp, but we also did not see the same type of engagement occurring with campers that we saw elsewhere.

Camp #1 was the camp that seemed to have quantitative results that were the most incongruent with our observations. A great deal of positive change was measured in the quantitative data, but subjectively we did not

Table 2
Comparison of Results between Quantitative Questionnaires and Qualitative Observations

Camp	No. of Significant Items[a]	Quantitative Measures			Qualitative Observations	
		Positive Significant Item Descriptors	Effect Size[b]		No. of "High" Ratings[c]	Relative Rankings[d]
#1	5	Positive Identity	.48		6	4
		Leadership	.32			
		Making Friends	.54			
		Adventure/Exploration	.55			
		Environmental Awareness	.22			
#2	3	Independence	.36		15	2
		Leadership	.25			
		Adventure/Exploration	.36			
#3	3	Adventure/Exploration	.44		16	1
		Environmental Awareness	.27			
		Positive Values	.37			
#4	0	--			11	3
#5	0	--			1	6
#6	0	--			5	5

[a] Reflects only items that were statistically significant (Table 1 includes all items)
[b] Effect size: Small = .10-.29; Moderate = .30-.49; Large > .50.
[c] Number of "high" ratings on the observation matrix
[d] Relative rank from highest to lowest of the six camps

sense as much positive development occurring. This camp seemed to be well structured, with a good program in a beautiful wilderness setting. The youth were appropriately age grouped and had many opportunities to participate in a variety of activities in cabin groups as well as other self-selected groups. We were not surprised that positive change occurred in positive identity, leadership, making friends, adventure/exploration, and environmental consciousness, as all these aspects were evident in camp. However, we did not get an overwhelmingly positive feeling about the camp. Unlike the other camps, we did not feel a sense of warmth and belonging For example, some of the campers seemed to interact primarily in small cliques. Similarly with Camp #4, which we rated relatively high in the observation, the quantitative results did not show positive outcome gains. Camp #4 was an exciting place to be, and we felt welcomed by the staff and campers. Perhaps the sense of achievement and a safe and warm community of campers and staff were not adequately measured in the quantitative questionnaires. In addition, Camp #4 had a different seasonal director in 2003 than in 2002, so differences in elements of staff training and supervision between the time of quantitative data collection and observations the following year may have existed.

Conclusions and Discussion

The purpose of this analysis was to explore indicators of youth development outcomes through a comparison of two forms of data collected at six camps. Although perfect agreement did not exist in this mixing of quantitative and qualitative data, the two approaches provided some convergence and certainly some complementary data, as indicated in our discussion of the results. As Sale et al. (2002) noted, complementarity is related to using a macro and micro analysis that can be combined. This study provided an opportunity to explore the subjective aspects of camp experiences along with quantitative approaches. It enabled us to compare the micro data obtained from the individual campers to the social environmental macro structure of the camp. This example of data corroboration reflects the recommendations of Domitrovich and Greenberg (2000) in examining how a theoretical framework of positive youth development can be measured in relation to how a program is implemented.

We acknowledge limitations in this study attributable to data collection at two different points in time. As a practical matter, the opportunity to go back to the camps did not occur until the following year. Nevertheless, this exploratory comparison showed that some of the indicators of youth development, such as positive and sustained adult–youth relationships, youth skill-building activities, and opportunities for youth participation and leadership that lead to positive outcomes (Lerner et al., 2005), can be observed by others and self-reported.

Peterson (2004) articulated the need for rigor and theory in youth program evaluation studies and identified several "truisms" about youth development from existing literature that can also be applied to camp experiences. He noted, as did Catalano et al. (2002), that more contact is better. In residential camps, even though some campers may only stay for a one-week period, the interaction with the program and staff is 24 hours a day. The actual hours of contact, even in a one-week camp (e.g., 15 hours a day for 6 days = 90 hours), might be equivalent to 2 hours a week for 9 months (e.g., 2 hours a week for 40 weeks = 80 hours). Peterson also noted that structure and sophistication are better in leading toward positive youth development. We could observe camp organizations that had a clear plan that was carefully monitored, and this plan (i.e., curriculum) resulted in concomitant positive change scores. In camp programs, individual campers live within a structured environment that offers opportunities for developing and practicing particular skills and competencies, such as living together in a cabin, making decisions about individual or group activities during the day, and participating in new activities such as ropes courses. Further, Peterson advocated that youth programs work best when they are done well. Staff interaction with campers, the participation of campers in choosing daily activities, and subjective responses (e.g., smiles) were some of the indicators we observed that correlated with camp programs that worked well in terms of the potential for positive youth development outcomes.

The results obtained through the field observations may provide information that can be further refined in developing future measures of positive youth development in camp or in other youth-serving organizations. Since it is not always possible to conduct pre- and post-test designs to measure change, the development of observation protocols might be useful to consider. Translating how outcomes could be observed to enable staff and directors to further monitor and evaluate their efforts at camp is important. Many aspects of outcome evaluation are yet to be uncovered, but the use of structured observations alone or in tandem with quantitative measures offers ways to understand what seems to be occurring in a youth program that results in achieving developmental goals.

In this study, the observers entered the camps with an array of general camp experience but a limited vision of what the desired outcomes would "look like" based on the specific camp's philosophy and implementation practices. Although natural limitations occur in observing personal values or identity, as opposed to social and physical skills, merit may lie in investigating the range of outcomes perceived by people with a vested interest in them (e.g., board members, head counselors, unit leaders, program heads, camper parents) as well as by a more detached observer (e.g., new or prospective parent and/or camper). The process of inviting an outside observer or using structured observations by key staff members within a camp has potential for yielding discussion points to complement and contrast the pen-and-paper evaluations of staff, camper experience, and program structure.

A further issue to consider is the types of guiding questions that might be used to conduct observations in camps and other youth organizations. The tool used for the qualitative observations in our study modeled the work of Eccles and Gootman (2002). These questions provided a starting framework, but more research should also be undertaken relative to what guiding questions would be most useful in attempting to observe possible outcomes relative to positive youth development. The process of discussing and refining guiding questions to investigate the goals and objectives behind a specific program can serve as a tool to achieve greater mutual understanding between administrators and frontline staff members. For years, the idea of "management by walking around" has been utilized. As more emphasis is placed on supervisor training and understanding the importance of intentional programming to reach desired objectives, knowing what to observe related to positive outcomes will be useful.

The value of this exploratory study lies in examining ways to measure micro and macro explanations of youth development in a structured camp experience. The implications for combining data and validating methods in other recreation organizations have the potential for understanding how and why youth development programs in recreation and camp settings work. The interaction of the individual within the structure of the social and cultural environment is not a clear cause-and-effect relationship. Therefore, the process used to gain insight can be a powerful tool for a program to

explore its goals and implementation. In any recreation program, but specifically in seasonal programs, the consistency and intentionality in creating a structure to support desired outcomes is an ongoing challenge. The use of mixed methods and multiple measures to document youth development outcomes holds potential as a research and program evaluation tool, in terms of both examining the process and documenting changes leading to positive youth development.

References

Addams, J. (1998). *Twenty years at Hull House: With autobiographical notes.* New York: Penguin 20th Century Classics.

Allen, L., Stevens, B., & Harwell, R. (1996). Benefits-Based Management Activity Model for youth in at-risk environments. *Journal of Park and Recreation Administration, 14,* 10-19.

American Camp Association. (2005). *The Youth Development Outcomes of the Camp Experience (YDOCE) study.* Martinsville, IN: Author.

Baker, D. W., & Witt, P. A. (2000). Multiple stakeholders' views of the goals and content of two after-school enrichment programs. *Journal of Park and Recreation Administration, 1,* 68-86.

Baldwin, C. (2000). Theory, program, and outcomes: Assessing the challenges of evaluating at-risk youth recreation programs. *Journal of Park and Recreation Administration, 18*(1), 19-33.

Baldwin, C., Caldwell, L, & Witt, P. (2005). Deliberate programming with logic models: From theory to outcomes. In L Caldwell and P. Witt (Eds), *Youth development and recreation services* (pp. 219-239). State College, PA: Venture.

Bialeschki, M.D., Younger, T., Henderson, K., Ewing, D., & Casey, M., II. (2002). Happy but sad. *Camping Magazine, 75*(1), 38-41.

Brannan, S., Arick, J., Fullerton, A., & Harris, J. (2000). Inclusive outdoor programs benefit youth. *Camping Magazine, 73*(1), 26-29.

Bumbarger, B., & Greenberg, M. (2002). Next steps in advancing research on positive youth development. *Prevention & Treatment*, *5.* Posted June 24, 2002, and retrieved August 19, 2005, from http://www.journals.apa.org/prevention/volume5/pre0050016c.html

Caldwell, L. L. (2000). Introductory comments: Special issue on youth and leisure. *Journal of Park and Recreation Administration, 18*(1), 1-4.

Carruthers, C. P., & Busser, J. A. (2000). A qualitative outcome study of boys and girls club program leaders, club members, and parents. *Journal of Park and Recreation Administration, 1,* 50-67.

Catalano, R. F., Berglund, M. L., Ryan, J. A. M., Lonczak, H. S., & Hawkins, J. D. (2002). Positive youth development in the United States: Research findings on evaluations of positive youth development programs. *Prevention & Treatment*, *5*, Article 15. Retrieved August 19, 2005, from http://journals.apa.org/prevention/volume5/pre0050015a.html

Chenery, M. F. (1991). *I am somebody: The messages and methods of organized camping for youth development.* Martinsville, IN: American Camping Association.

Christenson, S. L., & Ysseldyke, J. E. (1989). Assessing student performance: An important change is needed. *Journal of School Psychology, 27,* 409-425.

Domitrovich, C. E., & Greenberg, M. T. (2000). The study of implementation: Current findings from effective programs that prevent mental disorders in school-aged children. *Journal of Educational and Psychological Consultation, 11*(2), 193-221.

Dworkin, B. S. (1999). Campers speak: New England youth share ideas on societal issues. *Camping Magazine, 72*(5), 30-34.

Eccles, J. S., & Barber, B. L. (1999). Student council, volunteering, basketball, and marching band: What kind of extracurricular involvement matters. *Journal of Adolescent Research, 14*(1), 10-43.

Eccles, J. S., & Gootman, J. A. (Eds). (2002). *Community programs to promote youth development.* Washington, DC: National Academy Press.

Erzberger, C., & Prein, G. (1997). Triangulation: Validity and empirically-based hypothesis construction. *Quality & Quantity, 31*, 141-154.

Gambone, M. A., Klem, A. M., & Connell, J. P. (2002). *Finding out what matters for youth: Testing key links in a community action framework for youth development.* Philadelphia: Youth Development Strategies, Inc., and Institute for Research and Reform in Education (http://www.ydsi.org/YDSI/pdf/Whatmatters.pdf).

Haberman, M. (2000). Urban schools: Day camps or custodial centers? *Phi Delta Kappan, 82*(3), 203-208.

Henderson, K. A. (1991). *Dimensions of choice: A qualitative approach to recreation, parks, and leisure research.* State College, PA: Venture.

Henderson, K. A., & Bedini, L. A. (1995). Linking data in therapeutic recreation. *Therapeutic Recreation Journal, 29*(2), 124-130.

Henderson, K. A., & King, K. (1998). Recreation programming for adolescent girls: Rationale and foundations. *Journal of Park and Recreation Administration, 16*(2), 1-14.

Howe, C. Z., & Keller, J. (1988). The use of triangulation as an evaluation technique: Illustrations from regional symposia in therapeutic recreation. *Therapeutic Recreation Journal, 22*(1), 36-45.

Howe, K. R. (1992). Getting over the quantitative-qualitative debate. *American Journal of Education, 100*(2), 236-257.

Larson, R.W. (2000). Toward a psychology of positive youth development. *American Psychologist, 55*(1), 170-183.

Leffert, N., Benson, P. L., Scales, P. C., Sharma, A. R., Drake, D. R., & Blyth, D. A. (1998). Developmental assets: Measurement and prediction of risk behaviors among adolescents. *Applied Developmental Science, 2*(4), 209-230.

Lerner, R. M., Lerner, J. V., & Almerigi, J., & Theokas, C. (2005). Positive youth development: A view of the issues. *Journal of Early Adolescence, 25*(1), 10-16.

Lerner, R. M., Lerner, J. V., De Stefanis, I., & Apfel, A. (2001). Understanding developmental systems in adolescence: Implications for methodological strategies, data analytic approaches, and training. *Journal of Adolescent Research, 16*, 9-27.

Marsh, P. (1999). Does camp enhance self-esteem? *Camping Magazine, 72*(6), 36-40.

National Science Foundation. (1997). *User-friendly handbook for mixed method evaluations.* Retrieved August 21, 2005, from Division of Research, Evaluation and Communication web site: http://www.ehr.nsf.gov/EHR/REC/pubs/NSF97-153?START.HTM

Nicholson, H. J., Collins, C., & Holmer, H. (2004). Youth as people: The protective aspects of youth development in after-school settings. *The Annals of the American Academy of Political and Social Science, 591*, 55-71.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw Hill.

Peterson, C. (2004). Positive development: Realizing the potential of youth. *The Annals of the American Academy of Political and Social Science, 591*, 202-220.

Roth, J. L., & Brooks-Gunn, J. (2003). What exactly is a youth development program? Answers from research and practice. *Applied Developmental Science, 7*, 94-111.

Sale, J. E. M., Lohfeld, L. H., & Brazil, K. (2002). Revisiting the quantitative–qualitative debate: Implications for mixed-methods research. *Quality & Quantity, 36*, 43-53.

Sekine, A. (1994). The effect of camp experience upon the locus of control and general self-efficacy of school children. *Bulletin of Institute of Health and Sports Sciences, 17*, 177-183.

Strauss, A. L., & Corbin, J. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Newbury Park, CA: Sage.

Witt, P. A., & Crompton, J. L. (1996). *Recreation programs that work for at-risk youth: The challenge of shaping the future.* State College, PA: Venture.

Witt, P. A. (2002). Youth development: Going to the next level. *Parks & Recreation, 37*(3), 52-59.