Journal of Youth Development
Bridging Research & Practice

What's Going On?
Developing Program Theory for Evaluation

Mary E. Arnold
Department of Youth Development Education
Oregon State University
Corvallis, OR
mary.arnold@oregonstate.edu

Brooke D. Nott
Department of Youth Development Education
Oregon State University
Corvallis, OR
dolencb@onid.orst.edu

JOURNAL OF YOUTH DEVELOPMENT
bridging research and practice

Volume 5, Number 2, Summer 2010 **Article 100502RS001**

What's Going On? Developing Program Theory for Evaluation

Mary E. Arnold and Brooke D. Nott
Oregon State University

Abstract: This paper outlines the process undertaken to develop a program logic model into a program theory that could be tested through a large scale evaluation. Logic models are widely promoted to establish program intention and outcomes, thus setting the stage for effective program evaluation. However, the causal links in program logic models are often assumptive in nature, and lack the specificity of conditions needed in order for the program to achieve its planned success. The authors share a process of observation and focus groups that allowed them to probe some of the links of the program logic model. The result presented is an "enhanced" logic model that reveals possible moderators of program success, which leads to the development of evaluation questions that will be used in a subsequent large-scale evaluation. The value of using an enhanced logic model for program evaluation leading to program policy changes and improvement is underscored.

Introduction

The 4-H program is a positive youth development program that is part of the educational offerings of the Extension services of Land-Grant Universities across the county. In 2008, almost 6 million youth participated in one or more 4-H programs in the United States (Research, Education, & Economics Information System, 2009). One of the most popular 4-H projects is the horse project, in which young people learn about horses and their proper care and training. Youth participants in the 4-H horse project also have the opportunity to compete with their horse at 4-H fairs and other competitions.

Because of the popularity of the 4-H horse project we were interested in developing a program evaluation to investigate the impact of participation on youth. Very little previous research has been done on this topic, however, and therefore little theoretical information was available to guide our investigation. In addition, anecdotal information that seemed consistent from state to

state led us to believe that there may be some unique benefits and challenges within the 4-H horse program that could affect the ultimate impact of the program on youth. This paper outlines the process of using focus groups to develop a basic program logic model into more detailed program theory, which in turn allowed for the development of critical questions to be used in the program evaluation.

Review of the Literature

The research base for positive youth development (PYD) has developed significantly in recent years, and current research is revealing the particularly strong influence that the 4-H Youth Development program has on the positive development of youth. Many 4-H programs develop and use basic logic models that articulate the elements that are necessary for the program's success (cf Rennekamp & Arnold, 2006; Figure 1). These elements include:

(1) the resources, including volunteers, that make the program possible;

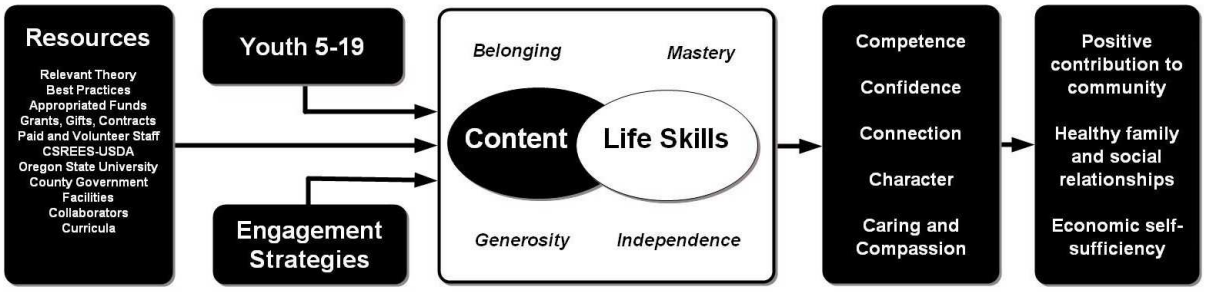
(2) the youth who participate and the types of programs in which they are engaged;

(3) the program's essential elements, which ensure that programs provide opportunities for belonging, mastery, independence, and generosity (Kress, 2004); and

(4) the content of the programs (often called a 4-H project) and the life skills that are developed through learning the content, such as leadership, responsibility, and teamwork.

When the aforementioned aspects of the model are in place, a predicted increase in the outcomes of PYD takes place. These important outcomes are commonly known as the five "C's" of PYD: confidence, competence, character, caring, and connection (Lerner & Lerner, 2006; Roth & Brooks-Gunn, 2004). Long-term, developing PYD outcomes leads to a successful transition to adulthood, marked by positive family and social relationships, economic self-sufficiency, and positive contributions to community (Gambone & Connell, 2004).

Figure 1
4-H Program Model



From Logic Models to Program Theory

The development of program logic models, such as the one described above, has been emphasized in recent years as a primary step in effective program evaluation (Frechtling, 2007; Knowlton & Phillips, 2009; Rogers, Petroscino, Huebner, & Hacsi, 2000). The need for program developers to understand a program's intent and articulate it in a clear manner underscores the popularity of teaching logic modeling as a first step in building evaluation capacity (Arnold,

2006). The purpose of program logic models is to articulate the causal connections between the resources that are invested in a program, what is done with those resources, who is reached as a result of the investment, and what happens in terms of short, medium, and long-term outcomes. The "logical" connections that link the elements of a logic model are often assumptive in nature, and, when used to illustrate program theory, have a distinctly predictive intention (McLaughlin & Jordan, 2004). As Chen (2004) points out, the assumptions about the causal links that describe how a program should work are crucial because the success the program's "effectiveness depends on their validity" (p. 17). One of the key aspects of program theory is that it moves beyond identifying the causal elements, to identifying the *actions* that are needed to link program elements (Braverman & Engle, 2009; Chen, 2004).

Despite their popularity in program design, logic models are often overly generalized and fall short of theoretical because there is not enough specificity in the model to understand the required conditions and actions that connect program inputs, outputs, and outcomes. One purpose of program evaluation is to test the model's theory and predictability, and in order to do so, logic models need to be prepared with a greater focus on the actions supporting the causal links. As Rossi, Lipsey, and Freeman (2004) state, the basic question underlying most program evaluation is: "is what's supposed to be happening actually happening?" (p. 93).

Articulating a program's theory is even more complex when multiple programs are conducted under the umbrella of a more general logic model. This is often the case in 4-H youth development where the participants, content, setting, and facilitators (often adult volunteers) can vary considerably between program sites and the content being learned.

In addition to being known for its popularity the 4-H horse program also stands out because of its unique "culture." County 4-H agents (professional youth development educators who manage the 4-H program at the local level) often report that adult volunteers in the 4-H horse program are passionate, committed, and dedicated to working with youth and their horses. At the same time, 4-H agents often report experiencing the greatest level of conflict within their programs with 4-H horse leaders. Because of the consistent anecdotal information regarding conflict within 4-H horse programs, we wondered about the potential negative impact the conflict could have on the program outcomes. Going further, we wondered what specific actions are necessary in light of potential conflict that could possibly moderate any negative effect of conflict, thus ensuring better program effectiveness. These questions, all variations of the basic program evaluation question posed by Rossi, Lipsey, & Freeman (2004), also highlight the reality that program theories are rarely as simple and linear as basic program logic models might suggest. Rather, actual program theory is more likely to be a complex web of events, influences, actions, interactions, and outcomes. Teasing out these relationships becomes an important aspect of more sophisticated program evaluation.

As the practice of program evaluation in Extension becomes more experienced, so has the thirst for more complex understanding of program impact and outcomes. While the basic evaluation question remains the same, understanding *if* a program is working is no longer fully satisfying in light of the recognized variations among 4-H programs. More and more, evaluation questions are focused on *how* a program is working, because it is in understanding the *how* that programs can reap the most benefit in using evaluation results to inform program policy and improvement.

The answer to the "how" question can sometimes be found in research and policy literature related to the issue the program is attempting to address. Some may argue that programs

should not be developed without a complete understanding of the research supporting them, but practice shows that programs often get underway first, and the theory comes later. In addition, as McLaughlin and Jordan (2004) point out, evaluation results can be used to shape program theory by providing explanations of causal relationships that are observed during the evaluation. In some situations, one may even employ specific, strategic methods designed to elicit latent program theory (Braverman & Engle, 2009; Leeuw, 2003).

The stated mission of Extension is to deliver research-based educational programming to the public (Franz & Townsend, 2008). As such, it appears that 4-H programming could be a hallmark of theory-driven programming. Indeed, there are many examples of this being the case, but in other instances the research base and theory is weak or not defined at all. At its broadest level, the 4-H program logic model was developed based on research related to best practices in positive youth development. As a theoretical model, however, it lacks detail related to the actions needed to ensure its success. At a more specific level, such as how the program may unfold for youth engaged with differing project content, the research base may be unknown, or even non-existent. Such is the case with the 4-H horse project. Research on the impact of horses on youth is minimal, and most of what has been done has focused on the use of horses in therapeutic riding programs (cf. Ewing, MacDonald, Taylor, & Bowers, 2007; Schultz, Remick-Barlow, & Robbins, 2007), with little or no quality research focusing on the benefits for youth who are not necessarily at risk for problems. Two exceptions are Lambarth's (2008) study of five 4-H alumna and their horses and Slocum's (2004) study of the impact of participation in 4-H horse events on the development of youth life skills. Despite the fact that the horse project is one of the most popular 4-H programs, we have little evidence to support its theoretical model.

What is going on? Focusing on the connections within a program logic model

The popularity of the 4-H horse program, combined with anecdotal evidence of its unique culture, piqued our interest in understanding the impact of participation on youth. We had reached a point in our development of evaluation capacity and organizational learning where we longed for something more; more detail and more understanding about how our programs work and why. In their multidisciplinary model of evaluation capacity building Preskill and Boyle (2008) outline the synergistic relationship between sustainable evaluation capacity building, sustainable evaluation practice, and the moderating effects of organizational leadership, culture, systems and communication. As an organization, 4-H has gained tremendous ground in developing evaluation capacity (Arnold, 2006) and continues to address many of the elements for sustainable evaluation practice (Preskill & Boyle, 2008). Along the way, we conducted many program evaluations based on the assumption that our program logic model was complete and accurate. However, we knew that the model (see Figure 1) was neither complete nor accurate, in that the connections between the model elements were largely unarticulated, specifically lacking detail on conditions and actions that promote success. In addition, even where more specificity was provided, there was little evidence to support how accurate the connections actually were. As Preskill and Boyle (2008) point out, one of the moderating elements for organizational learning includes leadership. The prospect of evaluating the 4-H horse program provided the opportunity for organizational leaders to press further into a more detailed understanding of program theory, which, in turn, would inform local 4-H educators and stakeholders, thus increasing organizational learning through the use of more detailed program knowledge.

With little theoretical knowledge to guide us, we decided to undertake a research project with the purpose of developing a more detailed program theory about the 4-H horse program, from

which a program evaluation could be developed. The research was conducted in two phases. The first was a series of observations at county-based 4-H horse fairs, followed by three focus groups later in the year. The goal of this work was to arrive at a more complete logic model for the 4-H horse program, highlighted by some of the actions that are necessary for the program to be effective. This enhanced program model would in turn be used as the basis for a large scale evaluation on the impact of the 4-H horse program on youth participants.

Methods

Using qualitative methods and focus groups

Qualitative methods are often able to break into the "black box" and describe the "how" and processes of theory related to practice (Zussman, 2004). Qualitative research strives to find creative ways of linking research to constructs. According to Barbour (2007), focus groups are an effective method for understanding the process of how outcomes are reached. "Focused" group interviewing as a data collection method has its beginning in the 1940's (Merten & Kendall, 1946). After some initial interest, the method largely disappeared as texts related to its use went out of print (Berg, 2004). Focus groups remained popular in marketing research, however, and by the late 1980's, made their way back to the social science realm as "group interviewing" methods (Berg, 2004).

In social science, focus groups are helpful as an independent research method, for informing various aspects of an in-depth study, such as the development of a questionnaire, and for multi-method research studies. The present study used focus groups to elucidate processes linked to the 4-H horse program as it relates to the 4-H logic model and uncover underlying theories to be tested in follow-up research about to understand the mechanisms that may impede or facilitate PYD within the 4-H horse program.

Focus groups and participants

Three focus groups were held across the state in both rural and urban sites, representing large and small horse programs. Previous research on focus groups indicates that three to five groups are adequate for saturation, though the exact number may depend on the context and research goals (Morgan, 1997). Each group consisted of eight to eleven participants; all participants were female, which is representative of the 4-H horse program in general. A total of 20 adult volunteer leaders and nine youth members of the 4-H horse program participated in the focus groups.

Leaders had been volunteering with the 4-H horse program for varying amounts of time. Three leaders reported being a leader for one year, five reported being a leader for two to four years, five reported being a leader for five to eight years; four reported being a leader for nine to 13 years, and two reported being a leader for 20 years (one adult leader did not report years in the program). Between two and four youth members were in each focus group. Youth members were between the ages of 12 and 16 (M= 15.3). Among the youth members, two had participated in the program for three years, six had participated for five to eight years, and one had participated for nine years (see Table 1).

Table 1

Focus Group Participant Characteristics

	Youth Members	Leaders
Number of Participants	9	20
Age range	12-16 (M = 15.3)	n/a
Years in 4-H Horse Project		
1 year	0	3
2-4 years	2	5
5-8 years	6	5
9-13 years	1	4
20 years or more	0	2
Participate in other 4-H projects	5	7
Participate in other horse clubs	9	15

Data collection

Participants were recruited with the help of local 4-H agents who were familiar with their 4-H horse leaders and youth. The groups convened on-site in the county connected with their 4-H horse program; two groups were held in 4-H Extension conference rooms and one in a local restaurant meeting room. All participants were given and signed consent forms; the youths' parents also signed a consent form allowing their child to participate. Participants received a $10 gift card for attending the focus group.

The moderator introduced her background with the 4-H horse program, explained the purpose of the focus group, asked the group to help self-monitor the focus group interaction, and mentioned the presence and reason for video recording during the group. The discussions were recorded through a note-taker and videotaping.

The focus groups followed a general outline of questions posed by the moderator and the discussions followed a semi-structured interview format. This composition allowed participants to interact and engage with one another in a free flowing manner. As in one-on-one interviews, the semi-structured format also allowed interviewees more freedom in what they disclosed and the researcher could highlight certain themes from the free speech of the participants (Corbin & Morse, 2003). Interview questions focused on the benefits of participation, motivation for involvement, the role of competition, parents' involvement, and the challenges leaders and members face in the 4-H horse program.

Data analysis

Because we had a broad set of themes in mind, based on previous observations at 4-H horse fairs and anecdotal evidence that we hypothesized would emerge during the focus groups, the focus group interview outline was developed on these "a-priori" themes. These themes included a heightened level of adolescent passion for and about horses, the importance of competition, and potential sources of conflict in the program. What the horse members and leaders felt about these themes and the nuances related to the themes, however, was unknown. Thus, we sought confirmation of the themes as well as emerging themes from during the focus groups (Barbour, 2007).

Information was observed and analyzed, including the participants' comments, interactions, body language, and social cues and during careful post hoc evaluation of transcripts, notes, and recordings (Morgan, 1997). After reviewing this information, the emerging themes were documented and sub grouped under headings, a method often used in focus group analysis (Barbour, 2007). We organized the results under the following broad categories:

1) effective and ineffective program factors;
2) leader characteristics, motivations, and challenges;
3) 4-H member characteristics, opportunities, and skills gained; and
4) parent characteristics.

Several themes were found pertinent to at least three of the broad categories such as passion, commitment, conflict, knowledge, life skills, and mentoring. Themes were also noted that were cited frequently within and across groups such as the higher self-esteem, confidence, patience, and leadership skills youth gain from their involvement in 4-H horse.

We also looked for polarities in opinion as well as similarities in the responses. For example, contradictory and ambivalent opinions were expressed about the role of competition. In one case competition was portrayed as a positive motivator for youth to stay involved in the program, as a chance for youth to grow out of competition and focus more on helping others, and in another case as a personal struggle for youth, parents, and leaders. Competition represents a theme that had many connotations connected with it, both positive and negative, and guided us to consider further in-depth research on the role of competition.

Conflict was another theme requiring careful reflection. One group in particular commented on and disagreed about the function of rules, regulations, and professional horse trainers with heightened emotionality. This group also talked about parents undermining leaders and youth. Leaders overstepping their boundaries and conflict between members were also cited. These tensions were not mentioned in other groups with such depth and emotion. Furthermore, when asked directly about conflict the group specifically stated conflict posed a problem for them in the 4-H horse program. The two other groups did not state conflict as a problem and only mentioned minor disagreements. Leaders in these focus groups mentioned being opinionated and disagreeing with parents but never did they mention a problem with members or parents not listening to them. These two groups also mentioned peer mentoring within the program as an important positive experience for their youth. The group claiming conflict discussed peer mentoring with negativity and how youth would not cooperate with other more experienced peers. The members in this group were clearly not benefiting from the youth leadership opportunities as much compared to the other two groups.

The subtle disparities between how differing opinions were expressed and reported between the groups were an important insight into conflict among the 4-H horse programs. This led us to categorize the groups into low and high conflict groups. Further research is clearly needed in the area of conflict and how conflict impacts the effectiveness of the 4-H horse program. While all 4-H horse programs have the potential to help youth with PYD, conflict within a group and among leaders and parents was recognized as a possible negative moderator of success. Alternatively, leaders who are able to work together collaboratively and work effectively with parents provide an environment for youth to thrive and engage with the horse program in such a way that it reaches its full potential in supporting PYD.

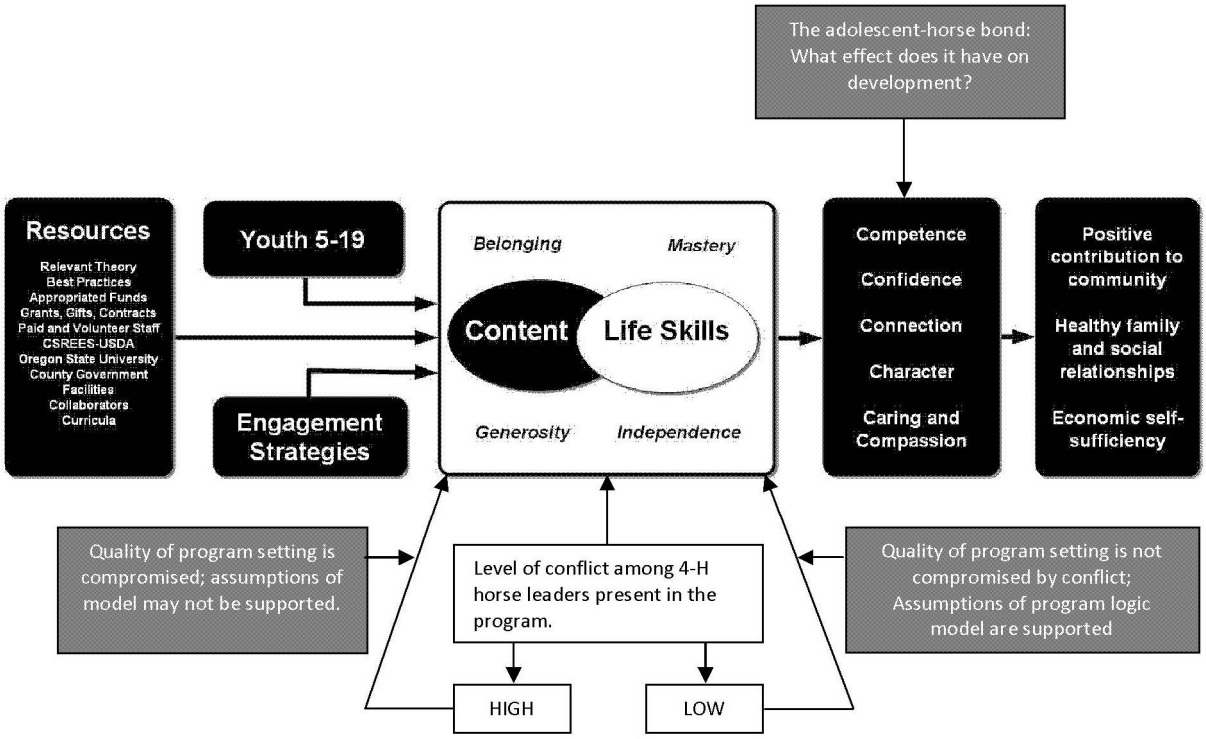
Next Steps: Understanding What's Going on

The results of the focus groups provided us with insight into the processes at work in the 4-H horse program, and set the stage for a large-scale program evaluation that is currently underway. Some processes, such as the building of youth confidence, are consistent with the basic 4-H program logic model. Others, however, raised concern about possible program effects that could negatively influence the predicted program outcomes. These processes were specifically related to conflict among 4-H horse leaders, members and parents. Still other emerging processes were identified, particularly related to the impact of peer teaching on the PYD. Finally, narrative information about the special relationship that teens have with horses, and the potential positive mediating effect on adolescent development invited us to look more closely at the nature of the horse-adolescent bond.

These findings were used to develop an enhanced program logic model that added explicit conditions and actions that may be taking place (see Figure 2).

Figure 2

Enhanced Program Logic Model



This enhanced model then served as the basis for developing a set of questions to test the accuracy of the program theory:

1. What is the nature of the adolescent-horse bond, and how does this bond contribute to the well-being of adolescents in the 4-H horse project?

2. Does youth participation in the 4-H horse project contribute to the development of life and personal skills of youth? If so, in what ways?
3. What is the extent nature of conflict among 4-H horse leaders? Are there factors that mediate the presence of conflict among horse leaders?
4. Does conflict among 4-H horse leaders have a negative effect on the development of 4-H youth?

Questions one and three are exploratory evaluation questions that will provide insight in the *processes* through which the 4-H horse program achieves its goal of positive youth development. Questions two and four are predictive evaluation questions that will provide information on possible moderators that may affect the success of the program as outlined in the basic program logic model. Both types of inquiry, exploratory and predictive, are useful for guiding program evaluations, the results of which can be used to enhance basic program logic models as well as to determine program policy and practice in future programming. Enhancing program logic models to understand and examine program theory through evaluation is an important step in ensuring the likelihood of program success in the future.

As stated earlier, logic modeling is a popular and effective form of program planning for many educational organizations. This paper supports the importance of thoughtful logic modeling for developing and evaluating educational programs. At the same time, the results of this work highlight the inadequacies of simple, linear logic models for understanding the complex interactions and moderating relationships present in most educational programs. Such complexities may be even more prevalent in educational programs that rely on the use of trained volunteers to implement programs, which is the case in many youth-serving programs. Developing enhanced logic models that specifically tease out and test potential moderating effects, both positive and negative, could be useful for all youth-serving organizations.

References

Arnold, M.E. (2006). Developing evaluation capacity in Extension 4-H field faculty: A framework for success. *American Journal of Evaluation, 27*(2), 257-269.

Barbour, R. (2007). *Doing focus groups.* U. Flick (Ed.), Thousand Oaks, CA: Sage Publications.

Berg, B.L. (2004). *Qualitative research methods for the social sciences* (5th ed.). Boston: Allyn and Bacon.

Braverman, M.T., & Engle, M. (2009). Theory and rigor in Extension program planning. *Journal of Extension* [On-line], *47*(3). Available at: http://joe.org/joe/2009june/pdf/JOE_v47_3a1.pdf

Chen, H. (2004) *Practical program evaluation: Assessing and improving planning, implementation, and effectiveness*. Thousand Oaks, CA: Sage.

Corbin, J., & Janice, M. (2003). The unstructured interactive interview: Issues of reciprocity and risks when dealing with sensitive topics*. Qualitative Inquiry, 9,* 335-354.

Ewing, C.A., MacDonald, P.M., Taylor, M., & Bowers, M.J. (2007). Equine-facilitated learning for youths with severe emotional disorders: A quantitative and qualitative study. *Child Youth Care Forum, 36*, 59-72.

Franz, N.K., & Townsend, L. (2008).The nature of complex organizations: The case of cooperative extension. In M. T. Braverman, M. Engle, M. E. Arnold, & Roger A. Rennekamp (Eds.) *Program evaluation in a complex organizational system: Lessons from Cooperative Extension. New Directions for Evaluation, 120,* 5-14.

Frechtling, J.A. (2007). *Logic modeling methods in program evaluation*. San Francisco: John Wiley and Sons.

Gambone, M.A., & Connell, J.P. (2004). The community action framework for youth development. *The Prevention Researcher*, 11, 17-20.

Knowlton, L.W., & Phillips, C.C. (2009). *The logic model guidebook: Better strategies for great results*. Thousand Oaks, CA: Sage.

Kress, C. (2004). *The essential elements of 4-H Youth Development: Distillation to four elements*. Washington, DC: CSREES/USDA.

Lambarth, L. (2008). *A phenomenological study of the relationship between award-winning 4-H alumni horsewomen and their horses*. Unpublished doctoral dissertation, Gonzaga University, Spokane, WA.

Leeuw, F.L. (2003). Reconstructing program theories: Models available and problems to be solved. American Journal of Evaluation, 24(1), 5-20.

Lerner, R.M., & Lerner, J.V. (2006). Toward a new vision and vocabulary about adolescence: Theoretical, empirical, and applied bases of a "positive youth Development" Perspective. In L. Balter, & Tamis-Lemonda, C. (Eds.) Child Psychology: A Handbook of Contemporary Issues. New York: Taylor and Francis Group.

McLaughlin, J.A., & Jordan, G.B. (2004). Using logic models. In J.S. Wholey, H.P. Hatry, K.E. Newcomer (Eds.), *Handbook of practical program evaluation* (2nd ed., pp. 7-32).

Merton R.K., & Kendall P.L. (1946). The focused interview. *American Journal of Sociology*, *51*, 541-557.

Morgan, D.L. (1997). Focus groups as qualitative research. *Qualitative Research Methods Series, 16, (2nd ed).* Thousand Oaks, CA: Sage Publications.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*, 443-459.

Rennekamp, R.A., & Arnold, M.E. (2006). *The 4-H program logic model*. Corvallis, OR: Department of Youth Development Education, Oregon State University.

Research, Education & Economic Information System (REISS). (2009). *2008 National 4-H enrollment report*. Washington, DC: Author.

Rogers, P.J., Petroscino, A., Huebner, T.A., & Hasci, T.A. (2000). Program theory evaluation: Practice and problems. In P.J. Rogers, T.A. Hacsi, A. Petrosino, & T. Huebner (Eds.). *New directions for Evaluation; Program theory in evaluation: Challenges and opportunities, 87*, 3-13.

Rossi, P.H., Lipsey, M.W., & Freeman, H.E. (2004). *Evaluation: A systematic approach (7th ed.)*. Thousand Oaks, CA: Sage.

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.

Schultz, P.N., Remick-Barlow, G.A., Robbins, L. (2007). Equine-assisted psychotherapy: A mental health promotion/intervention modality for children who have experienced intra-family violence. *Health and Social Care in the Community, 15*, 265-271.

Slocum. S.S. (2004). *A comparison of leadership life skills development of youth participating in riding and non-riding competitive 4-H horse events in Mississippi*. Unpublished doctoral dissertation, Mississippi State University, Mississippi State, MI.

Zussman, R. (2004). People in places. *Qualitative Sociology* 27, 351-363.